## ABSTRACT

LEE, SANGMI. Characterization of the Immigration Control Region of *Listeria monocytogenes*. (Under the direction of Sophia Kathariou.)

*Listeria monocytogenes* is a foodborne pathogen that causes a potentially severe disease (listeriosis). Of 13 serotypes of *L. monocytogenes*, serotype 4b has been involved with numerous outbreaks and four serotype 4b epidemic-associated clonal groups (Epidemic Clone (EC) I, II, Ia and V) have been identified. Genomic DNA from ECI strains is characteristically resistant to Sau3AI due to a Sau3AI-like restriction-modification (RM) system. In this study, we showed that the genetic region harboring this RM system ("Region 85") was inhabited by various RM systems such as Type I RM systems, two novel Type II RM systems and two Type IV RM systems (Mrr and McrB), suggesting that this is an Immigration Control Region (ICR) of *L. monocytogenes* and might be important for the protection against phage infection. Hybridizations with the DNA probes based on these RM systems revealed that Type I RM systems were overrepresented in serotype 4b and often accompany *mrr* and *mcrB*. Among serotype 4b strains, ECIa strains harbored a Type I RM system and *mrr* but were characteristically missing *mcrB*. This feature could discriminate ECIa from other serotype 4b isolates. Meanwhile, the genetic diversity in Region 85 also resulted from a lipoprotein gene, which frequently diversified in a population lineage overrepresented among animal isolates (lineages III and IV). Most of the genes in Region 85 seem to be horizontally acquired by an ancestral strain and serial deletions might result in the observed diversity.

The novel Type II RM systems (LmoJ2 and LmoJ3) and the lipoprotein gene were further characterized. LmoJ2 and LmoJ3 recognized closely-related DNA sequences (GCWGC (W=A or T) and GCNGC (N=A, T, G or C), respectively). Phages methylated by LmoJ2 exhibited significant increase in infectivity at low temperatures compared with the non-methylated counterpart whereas methylation by LmoJ3 increased infectivity at all temperatures examined from 4 to 42ºC. For the lipoprotein gene, a deletion mutant of this gene was compared with the parental strain under stress conditions and with various toxic compounds. The deletion mutant exhibited moderate or small phenotypic differences including reduction in swarming on the soft agar plates.

To identify other ICR(s) in *Listeria* species, the RM systems in *Listeria* species were retrieved from the REBASE database and the flanking regions were compared. Consequently, the second ICR was identified downstream to a multidrug resistance gene and this region harbored Type I RM systems with two specificity subunit genes separated by a bacteriophage integrase gene in *L. innocua*, *L. monocytogenes* and *L. welshimeri* genomes.

We also identified three novel clonal groups of serotype 4b *L. monocytogenes* that commonly harbor a genetic cassette (*lmo0734-lmo0739*) specific to a population lineage (lineage II) to which serotype 4b (lineage I) does not belong. These clonal groups exhibited distinct hybridization profiles for ECI and ECII-specific markers and multi-locus genotyping (MLGT) types, and all harbored multiple RM systems in Region 85. The prevalence of these novel clonal groups as well as epidemic clones was investigated among serotype 4b isolates of various sources. Our analyses suggested that three major clonal groups of ECI, II and Ia were not overrepresented among clinical isolates, suggesting that frequent human exposure

could be important in association of these clonal groups with listeriosis cases. A subset of clinical isolates collected in the United States between 2003 and 2008 was subjected to the analysis of the temporal changes of clonal groups. In this analysis, ECI was identified in all the years and the prevalence of ECIa increased between 2006 and 2008.

Characterization of the Immigration Control Region of *Listeria monocytogenes*

by
Sangmi Lee

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Food Science

Raleigh, North Carolina

2011

APPROVED BY:

| | |
|---|---|
| _____ | _____ |
| Dr. Sophia Kathariou<br>Committee Chair | Dr. James W. Brown |
| _____ | _____ |
| Dr. Illenys M. Perez-Diaz | Dr. Sue Carson |

## DEDICATION

I dedicate this dissertation to somebody, including myself in the future, who might refer to this dissertation for their research. My efforts would not be in vain if this dissertation is of any help to anybody seeking information on *Listeria monocytogenes*, one of the most intriguing microorganisms.

## BIOGRAPHY

Sangmi Lee was born and raised in Daegu, a beautiful city in the south region of South Korea. In 1998 after high school, she moved to Seoul, the capital city, to seek for her advanced education. She earned bachelor's and master's degrees in Food Science and Technology from Seoul National University. For her master's degree, she conducted research on the sugar transport system in *Escherichia coli* under the mentorship of Dr. Sangryeol Ryu. In 2007, she joined Dr. Sophia Kathariou's lab and has worked on *Listeria monocytogenes* for four and half years. In 2011, she is finishing her PhD study and expecting a new change in her career as a researcher.

## ACKNOWLEDGMENTS

In finishing my PhD study, the person that I feel most indebted to is Dr. Sophia Kathariou, who accepted me into her lab and provided an opportunity to receive education in the United States. She trained me to have curiosity and venture out in the research and, whenever I encountered an obstacle, she willingly offered encouragement and suggestions based on extensive knowledge that has accumulated through her long, prolific career as a researcher. I feel sad that I might no longer see her contemplating with her eyes with long lashes attentively blinking during our meeting. Also, I would also like to express my gratitude to my committee members, Dr. James Brown, Dr. Sue Carson and D. Illenys M. Perez-Diaz, for their support and encouragement.

Besides Dr. Kathariou and committee members, I am thankful for all former and current lab members in Kathariou group, to name a few, Ms. Robin Siletzky, Harun, Jae-won, Johari, Lola, Mira, Mohammed, Reha, Savitri, Sandra, Shakir, Sheea, Vikrant, and Ying. They stood by me and celebrated with me in every step of the PhD study and it has been an enriching experience to work harmoniously and make friends with the people from such a diverse cultural background.

However, my heart always goes across the Pacific to my parents who I feel sorry that I have not been able to see enjoying their golden ages, and my siblings who have struggled to take a step into their careers. Without help and support of any of these people, I could not have been who I am now.

## TABLE OF CONTENTS


LIST OF TABLES ..... viii
LIST OF FIGURES ..... ix
Chapter 1 Literature Review ..... 1
Introduction ..... 1
Epidemiology of listeriosis ..... 2
Recent epidemiological trends of listeriosis ..... 4
Pathogenicity of *L. monocytogenes* ..... 9
Intracellular life of *L. monocytogenes* ..... 9
*Listeria* pathogenicity islands ..... 11
Premature stop codon mutations in *inlA* ..... 13
Stress Tolerance ..... 15
Surface proteins in *L. monocytogenes* ..... 16
Lipoproteins ..... 18
Phylogeny and Genomic Comparisons of *Listeria* species ..... 20
Phylogeny of *Listeria* species ..... 20
Genomic comparisons of different *Listeria* species ..... 22
Serotyping and population structure of *L. monocytogenes* ..... 24
Lineage I ..... 25
Lineage II ..... 27
Lineage III ..... 29
Evolution of different *L. monocytogenes* serotypes ..... 31
PCR-based Serotyping of *L. monocytogenes* ..... 32
Epidemic clones ..... 34
Epidemic Clone I ..... 34
**Fragment 85** ..... 36
**Fragment 17B** ..... 37
**Fragment 133** ..... 38
Epidemic Clone II ..... 38
**Region 18** ..... 39
**Region 1168** ..... 40
**Region 0088** ..... 40
**Region 2753** ..... 41
Epidemic Clone III ..... 41
Epidemic Clone IV ..... 42
Epidemic Clone V ..... 43
Increased exposure or higher pathogenicity? ..... 44
*Listeria* phages ..... 46
Temperate *Listeria* phages ..... 47
**A118** ..... 47
**A006** ..... 49
**PSA** ..... 49

**A500** ... 50
**B025** ... 50
**B054** ... 51
**Relationships between temperate Listeria phages** ... 51
Lytic *Listeria* phages ... 52
**A511** ... 52
**P100** ... 53
**P35** ... 53
**P40** ... 53
Relationship between lytic *Listeria* phages ... 54
Influence of *Listeria* phages on the *L. monocytogenes* ecology and evolution in the food processing environment ... 54
Restriction-modification systems ... 56
Type I ... 57
Type II ... 59
Type III ... 62
Type IV ... 63
Immigration Control Region ... 64
Biological roles of RM systems ... 65
Defense against restriction ... 67
REFERENCES ... 74
Chapter 2 Characterization of Atypical *Listeria monocytogenes* Serotype 4b Strains Harboring a Lineage II-Specific Gene Cassette ... 97
ABSTRACT ... 97
INTRODUCTION ... 98
MATERIALS AND METHODS ... 99
RESULTS ... 102
DISCUSSION ... 106
REFERENCES ... 133
Chapter 3 Genetic Diversity of the Immigration Control Region of *Listeria monocytogenes* ... 139
ABSTRACT ... 139
INTRODUCTION ... 141
MATERIALS AND METHODS ... 143
RESULTS ... 146
DISCUSSION ... 167
REFERENCES ... 230
Chapter 4 Two Novel Type II Restriction-Modification (RM) Systems Occupying a Genomically Equivalent Location on the Chromosome of *Listeria monocytogenes* ... 238
ABSTRACT ... 238
INTRODUCTION ... 240
MATERIALS AND METHODS ... 242
RESULTS ... 247

DISCUSSION ..... 255
REFERENCES ..... 275
Chapter 5 Functional Characterization of a Putative Lipoprotein in an Immigration Control Region of *Listeria monocytogenes* ..... 282
ABSTRACT ..... 282
INTRODUCTION ..... 283
MATERIALS AND METHODS ..... 285
RESULTS ..... 294
DISCUSSION ..... 298
REFERENCES ..... 332
Chapter 6 Restriction-Modification Systems in *Listeria* Species and Their Phages ..... 337
ABSTRACT ..... 337
INTRODUCTION ..... 338
MATERIALS AND METHODS ..... 340
RESULTS ..... 343
DISCUSSION ..... 350
REFERENCES ..... 381
Chapter 7 Prevalence and Chronological Trends of *Listeria monocytogenes* Serotype 4b Clonal Groups ..... 390
ABSTRACT ..... 390
INTRODUCTION ..... 391
MATERIALS AND METHODS ..... 393
RESULTS ..... 395
DISCUSSION ..... 399
REFERENCES ..... 433

## LIST OF TABLES

Table 2.1 Bacterial strains used in this study ..... 111
Table 2.2 List of the probes and oligonucleotides employed in this study ..... 112
Table 3.1 Isolates examined in this study ..... 173
Table 3.2 Number of isolates belonging to different serotypes and sources ..... 195
Table 3.3 List of isolates whose Region 85 was sequenced ..... 196
Table 3.4 Primers and DNA probes used in this study ..... 197
Table 3.5 Pair-wise comparisons of lipoproteins ..... 199
Table 3.6 Distribution of the isolates positive for the Region 85 gene probes ..... 200
Table 3.7 Distribution of the different hybridization profiles of Region 85 probes ..... 201
Table 4.1 Bacterial strains used in this study ..... 262
Table 4.2 DNA primers used in this study ..... 264
Table 4.3 Frequency of LmoJ2 and LmoJ3 recognition sites (GCWGC and GCNGC, respectively) ..... 265
Table 5.1 Bacterial strains used in this study ..... 302
Table 5.2 Primers employed in this study ..... 303
Table 5.3 MIC of the wild type strain (F2365) and deletion mutant (SL5-L16-5-5) for toxic agents ..... 304
Table 6.1 List of RM systems in *Listeria* and phages retrieved from REBASE ..... 356
Table 6.2 Number of the RM systems in *Listeria* and phages retrieved from REBASE. ..... 362
Table 7.1 List of the isolates analyzed in this study and clonal group classification ..... 406
Table 7.2 Oligonucleotides and DNA probes used in this study ..... 421
Table 7.3 Clonal group prevalence among serotype 4b isolates from different sources ..... 423
Table 7.4 List of the serotype 4b clinical isolates collected in the United States, 2003-2008 ..... 424
Table 7.5 Prevalence of serotype 4b clonal groups among 142 serotype 4b isolates from human listeriosis cases in the United States, 2003-2008 ..... 432

## LIST OF FIGURES

Figure 1.1Comparison of the genetic region flanked by *prs* and *ldh* between *L. monocytogenes* F2365 and *L. innocua* CLIP11262. ..... 70
Figure 1.2 Comparison of the genetic region harboring LIPI-3 in *L. monocytogenes* F2365 (lineage I/ ECI) and the genomically equivalent region in *L. monocytogenes* HPB2262 (lineage I/ ECIa), which is deficient of LIPI-3. ..... 71
Figure 1.3 Summary of the phylogenetic relationships of *Listeria* species revealed by previous studies. ..... 72
Figure 1.4 Comparison of Region 18 between F2365 (ECI) and H7858 (ECII). ..... 73
Figure 2.1 Multiplex serotyping results of an atypical 4b isolate. ..... 113
Figure 2.2 DNA-DNA hybridization results. ..... 114
Figure 2.3 PFGE dendrogram of atypical serotype 4b. ..... 115
Figure 2.4 DNA sequence comparison. ..... 116
Figure 2.5 PCR amplicons from primers annealing flanking genes and *lmo0737*. ..... 131
Figure 3.1Genetic organization of Region 85 of different *Listeria* spp. and *L. monocytogenes* strains obtained from bioinformatics analyses and sequencing results. ..... 204
Figure 3.2 Alignment of the *hsdS* genes identified in Region 85 of J3976 (J3976 *hsdS*), HPB2262 (*LMSG_01574*) and *L. welshimeri* SLCC3954 (*lse_0280*). ..... 210
Figure 3.3 Comparison of the lipoprotein genes and homologous DNA sequences (A and B) and of the proteins encoded in the lipoprotein genes (C). ..... 213
Figure 3.4 Comparison of the upstream (A) and downstream (B) flanking regions of Region 85 in which different gene contents reside. ..... 223
Figure 3.5 Phage infection assay results. ..... 229
Figure 4.1 Organization of the genomic region harboring LmoJ2 and LmoJ3. ..... 267
Figure 4.2 LmoJ2 and LmoJ3 are present in the same genomic locations in isolates of diverse serotypes. ..... 269
Figure 4.3 PFGE dendrogram of isolates harboring LmoJ2 and LmoJ3. ..... 270
Figure 4.4 LmoJ2 and LmoJ3 are associated with resistance to digestion by restriction enzymes recognizing GCWGC and GCNGC, respectively. ..... 272
Figure 4.5 Phage infection assay results of J2479 (A) and J3115 (B). ..... 274
Figure 5.1 Genomic region harboring the putative lipoprotein gene in *L. monocytogenes.*. 305
Figure 5.2 Southern hybridization of the genomic DNA digested with EcoRI with the DNA probe (LP) amplified with primers LPF/LPR based on the lipoprotein gene. ..... 309
Figure 5.3 Construction and confirmation of the complementation mutant. ..... 310
Figure 5.4 Results of motility tests. ..... 313
Figure 5.5 Microscopic observation of the parental strain F2365 (A) and the lipoprotein deletion mutant SL5-L16-5-5 (B) after 24 hr incubation in BHI at 37ºC. ..... 315
Figure 5.6 Hemolysis (A) and growth on the agar plates (B-D). ..... 316
Figure 5.7 Growth of the spotted cultures on the agar plates. ..... 317
Figure 5.8 Growth curves and post-stationary survival in BHI and TSB. ..... 318
Figure 5.9 Growth curves and post-stationary phase viability in BHI supplemented with 7% NaCl. ..... 324

Figure 5.10 Results of disk diffusion assays after 24 hr incubation at 37ºC. ..... 327
Figure 5.11 Killing curves in the lethal level of acid (A), nisin (B) and SDS (C)..... 328
Figure 5.12 Phage infectivity assays. ..... 331
Figure 6.1 Comparison of methyltransferases (A), restriction endonucleases (B) and specificity subunits (C) comprising RM systems in *Listeria* and phages. ..... 364
Figure 6.2 Comparison of the DNA of the homologous RM systems (G3-G10 shown in A-H, respectively) with the WebACT. ..... 368
Figure 6.3 Comparison of the regions flanking the RM systems in *Listeria* and phages whose genome was sequenced. ..... 373
Figure 6.4 Genetic organizations and diversity in the additional immigration control region (ICR) (A) and inverted repeats in proximity of integrase genes (B, C and D). ..... 375
Figure 6.5 Comparisons of the homologous genes within the second ICR. ..... 380

# Chapter 1 Literature Review

## Introduction

*Listeria monocytogenes* is a rod-shaped, facultative anaerobic, non-sporeforming Gram-positive bacterium that was first discovered in 1921 in lab animals and named in honor of the taxonomist Lord Lister (130, 141, 145). This microorganism belongs to the genus of *Listeria* and is classified as the only human pathogen in this genus (141, 145). As an etiological agent, *L. monocytogenes* causes listeriosis in humans and animals (141, 145). Among humans, listeriosis is transmitted exclusively via foods (93, 121, 130, 138, 140, 141) and its clinical manifestations range from mild flu-like symptoms accompanying gastroenteritis and diarrhea to life-threatening meningitis, encephalitis, and sepsis (38, 140, 141, 145). Moreover, the infection during pregnancy can lead to abortions or stillbirths (38, 130, 141, 145). The number of listeriosis cases is relatively low in comparison to other foodborne pathogens such as *Salmonella* and *Campylobacter* (145); however, the mortality rate amounts on average to 30%, which is considered high as a foodborne pathogen given that *Salmonella enteritidis* and *Campylobacter* species result in the death rate below 0.4% (141). Apart from foods, *L. monocytogenes* is often found in the natural environments such as water and soil, and in food processing facilities (72, 141, 145). All these characteristics, i.e., the severe invasive symptoms, high mortality rate and ubiquitous presence, have rendered *L. monocytogenes* a considerable public health concern.

## Epidemiology of listeriosis

Until 1980s *L. monocytogenes* had been known as an animal pathogen which rarely infects humans (145). The first evidence of foods as a transmission vehicle was presented in 1981 as a result of the outbreak that occurred in 1981 in Nova Scotia, Canada, due to contaminated coleslaw (130, 138, 145). Subsequently, North America and Europe witnessed a number of listeriosis outbreaks that are associated with deli meats, unpasteurized milk, pâté, soft cheese, seafood and turkey frankfurters (72, 130, 138, 141, 145). These outbreaks triggered the introduction of rigorous regulations to reduce the contamination by *L. monocytogenes*. In the United States, *L. monocytogenes* has been designated as zero-tolerance bacterium in ready-to-eat (RTE) foods (38, 52, 140, 141, 145). Meanwhile, the European Union has applied the limit of 100 colony forming units (cfu) per g in the RTE foods where *L. monocytogenes* is not capable to multiply (52, 141, 145) since the contamination below 100 cfu per g is unlikely to cause listeriosis (141).

The analysis of the listeriosis outbreaks and sporadic cases to date has revealed intriguing features of human listeriosis epidemiology. Firstly, listeriosis occurs more frequently among several at-risk populations such as the elderly; the immunocompromised due to cancer, organ transplant and HIV infection; neonates; and pregnant women (38, 130, 141). In the United States, Hispanics are considered as another at-risk population since listeriosis, in particular pregnancy-associated cases, was observed among this ethnic group with a higher rate than other ethnicities (140). This higher incidence is speculated to be associated with soft cheese made from raw milk (140). The infectious dose in these at-risk

groups has not been known (130, 145) but considered to be in the range of 100-1000 cells (145).

Secondly, the incubation period ranges from 1 to 90 days (141), exceeding 30 days in many cases (130). This long incubation period creates challenges for the epidemiological investigation of listeriosis cases (130, 145).

Thirdly, several high risk foods were identified such as the above-mentioned outbreak-associated foods (141). In most cases, those implicated in clinical cases were contaminated with *L. monocytogenes* at $10^2$-$10^6$ cfu/ml or g (141).

Fourthly, out of twelve serotypes of *L. monocytogenes*, only three serotypes (141), i.e., serotypes 1/2a, 1/2b and 4b, are responsible for over 95% of clinical cases (73, 130, 140, 141). In particular, serotype 4b has been associated with a majority of listeriosis outbreaks (72, 130) in spite of the fact that serotypes 1/2a and 1/2b are more frequently isolated from foods or processing facilities (130). This feature and other epidemiological findings indicate that serotype 4b might surpass other serotypes in virulence; however, its molecular basis has not been known (130).

Finally, listeriosis revealed a seasonal peak during the late summer and fall (8, 140, 141). The US study during 1996-2003 observed that 37% of listeriosis cases among persons over 50 years occurred in the period from August to October (140). Similar seasonality was reported in France (53) and among nonpregnancy-associated listeriosis patients aged over 60 years in England and Wales (58). In another US study in 1990-2005, mortality was also increased from July through October with the highest level in September (8). The reason for this seasonal peak remains to be unclear (8, 140) although Bennion et al. suggested that this

tendency might be linked to the increased exposure to RTE meats in the summer through picnics or barbecues (8).

**Recent epidemiological trends of listeriosis**

Latest studies revealed changing dynamics of listeriosis epidemiology, which is most noticeably evidenced in the contrasting trends of incidence in two different geographic regions, the US and Europe. In the United State, listeriosis has been estimated in 2011 to result in annual 1600 cases and 250 deaths (121). Following implementation of food safety policies against *L. monocytogenes*, the incidence of listeriosis was significantly decreased in the United States (140, 141). Listeriosis incidence was reported in 1989 at 7.9 cases per million; however, it diminished to 4.2 cases per million in 1993 (140). This trend was maintained in the study conducted by Voetsch et al. analyzing the Foodborne Disease Active Surveillance Network (FoodNet) data collected from 1996 through 2003 (140). The overall incidence declined from 4.1 cases per million in 1996 to 3.1 cases per million in 2003, which suggests 24% reduction in comparison with the baseline period of 1996-1998 (140). The most recent report by CDC on the foodborne pathogen incidence showed that 3.4 cases of listeriosis occurred per million population in 2009 (16). This figure was lower by 26% than the period of 1996-1998 (16). Interestingly, the listeriosis incidence was reduced most noticeably until 2001 and has been sustained at a similar level (52) although a moderate increase has been observed in recent years (16, 58). In addition to the decrease in the incidence, mortality was lowered in the US. Bennion et al. reported that deaths by listeriosis were reduced each year by 10.74% during 1990-1996 and by 4.26% during 1996-2005,

culminating in the decrease in mortality from 0.53 to 0.18 per million during this period (8). Possibly reasons for these phenomena include strict food safety regulations (8, 130, 140); efforts by industry such as Hazard Analysis Critical Control Points (8, 130, 140) and test-and-hold systems (140); and improved communications on listeriosis with at-risk groups (8, 130).

On the other hand, listeriosis has increased significantly in the European countries (52, 58, 66, 141, 145). In England and Wales, Gillespie et al. reported annual listeriosis incidence at 3.47 cases per million persons during 2001-2004, which had increased from the baseline level of 2.13 cases per million between 1990 and 2000(53). During 2001-2007, the incidence doubled in comparison with the period between 1990 and 1999 (54). This rise was observed mainly in the sporadic cases affecting the patients aged over 60 (53, 54). In this population group, bacteremia significantly increased without the infection within the central nervous system (53, 54). In France, listeriosis decreased between 1987 through 1997 after extensive control strategies were conducted by the food industry due to large scale outbreaks (59). Subsequent to the implementation of the mandatory reporting, listeriosis incidence was further reduced from 4.5 cases per million in 1999-2000 to 3.5 cases per million in 2001-2003 (59). This level was maintained until 2005; however, the listeriosis incidence rose to 4.7 cases per million in 2006 and was estimated at 5.6 cases per million in 2007, suggesting 46% increase during 2006-2007 from the period between 2001 and 2004(58). This increase was accounted for by the rise in listeriosis patients over the age of 60 years and in bacteremia (58). In Germany, listeriosis substantially increased during 2001-2005 with marked rise in listeriosis cases among the elderly population over 60 years (81). Elevated listeriosis

incidence was also observed in other European countries such as the Netherlands (34), Switzerland, Denmark, Finland and Belgium (58).

The reason for this trend in Europe is not known (52, 53, 58, 81). Although several factors were suggested that might have produced artifacts in the listeriosis incidence in an attempt to explain these phenomena, most of them did not prove to be plausible (53, 58, 81). For instance, in the study on the listeriosis trend in England and Wales in 2001-2004, Gillespie et al. considered following possible factors that might have influenced the listeriosis incidence: altered reporting system; improvement in the *Listeria* isolation procedures; enhanced awareness for listeriosis in the medical community; dietary changes; demography; outbreaks; and emergence of a virulent *L. monocytogenes* strain (53). The increased pathogenicity was excluded from the possible scenarios since multiple subtypes of *L. monocytogenes* were involved in the increase (53). In the other aspects, they did not observe changes to such extents to trigger the observed increase in listeriosis incidence (53). Similarly, Goulet et al. suggested that the increasing listeriosis in 2006-2007 in France is likely to result from neither of improved reporting system, better diagnostic awareness, large-scale outbreaks, nor changes in the control measures implemented in the food industry (58). The rise in bacteremia might have been due to altered medical practices such as increased hospitalization, more frequent blood cultures or more sensitive blood-culture techniques; however, Goulet et al. also failed to discover evidence in these aspects that supports the upsurge in bacteremia by *L. monocytogenes* (58).

Nonetheless, the above-mentioned studies and others shed light on potential contributing factors to the increase in listeriosis (53, 54, 58). Both Goulet et al. and Gillespie

et al. suspected that the higher incidence of listeriosis might be related to higher exposure to *L. monocytogenes* (53, 58). As a scenario that could lead to increased exposure to this bacterium, Goulet et al. proposed that, in compliance to the recommendation against hypertension, less salt was used in some RTE foods; consequently, *L. monocytogenes* may have grown more readily in these foods if contaminated, thus causing listeriosis in at-risk groups in higher probability (58). A study by Gillespie et al. conducted in England and Wales attempted to elucidate medical factors underlying the listeriosis patients aged over 60 years who exhibited bacteremia without the central nervous system infections in 2001-2007 (54). These patients were associated with cancer or other conditions whose treatments suppress the secretion of stomach acid (54). All these medical conditions are likely to lead to increase in persons vulnerable to *L. monocytogenes* infections (54).

Another topic of note is the emergence of a new kind of listeriosis outbreaks predominantly manifested by febrile gastroenteritis in the 1990s and 2000s (130). Dissimilar to invasive listeriosis outbreaks, these outbreaks usually afflicted immunocompetent adults and the patients developed self-limiting fever, diarrhea, vomiting, headache and body pain after a short incubation period of 18-27 hr (130). These outbreaks were implicated with rice salad (120), chocolate milk (28), cold corn and tuna salad (5), RTE deli meats (48), cold-smoked fish, cheese (130), which were contaminated with a large number of *L. monocytogenes*, for instance, at ~$10^9$ cfu/ml in 1994 US chocolate milk outbreak (28, 130). Many of the febrile gastroenteritis outbreaks were caused by serotypes 1/2a and 1/2b although serotype 4b was also involved (22, 73, 130). After the above-mentioned chocolate milk outbreak, pulsed-field gel electrophoresis (PFGE) revealed that the outbreak strains

were associated with sporadic cases of invasive listeriosis (112), suggesting that the high levels of *L. monocytogenes* may cause febrile gastroenteritis in healthy individuals (72, 130). However, further research would still be needed to determine whether these non-invasive listeriosis outbreaks stem from the unusual characteristics of outbreak strains or from the exposure to the inoculum at high quantities (72).

Related to the invasive listeriosis, susceptible group categories for listeriosis are altering, in particular among HIV patients and Hispanic ethnic group. For instance, the percentage of the HIV patients in the deaths from listeriosis was lowered from 4.9% in 1990-1995 to 3.4% in 1996-2005 in the US, possibly due to the introduction of effective treatments against diseases caused by opportunistic organisms including *L. monocytogenes* (8). Among the Hispanic population in the US, the study by Voetsch et al. showed that listeriosis occurred at the disproportionately high incidence in 1997-2001; however, the frequency of listeriosis was considerably diminished in 2002 and 2003 (140). This change might originate from various public education campaigns or elimination of unsafe foods from the market (140).

High risk foods for listeriosis are also being reviewed by attempts to determine the food vehicles of sporadic cases of listeriosis (138), which occupy the majority of cases (138, 141). Several studies in the 1980s and 1990s linked sporadic listeriosis cases with frankfurters, chicken, milk, pâté, soft cheese and deli foods (138). The most recent study on this subject, which was conducted by Varma et al., compared 169 sporadic case patients with control subjects in the US in 2000-2003 (138). This case-control study associated melons consumed at a commercial establishment and hummus prepared at a commercial

establishment with sporadic cases, highlighting the importance of retail environments in contamination with *L. monocytogenes* (138). Given that high risk foods for listeriosis are often ready-to-eat, refrigerated and subjected to storage for a long period, Goulet deemed that hummus was more likely to belong to this category of foods than melons (57).

## Pathogenicity of *L. monocytogenes*

### Intracellular life of *L. monocytogenes*

As evidenced by the listeriosis epidemiology, a majority of cases are mediated by the foods (93, 130, 138, 140, 141). Following the intake of the contaminated foods, the infection by *L. monocytogenes* is initiated with this bacterium penetrating through the epithelial lining of the intestine (56, 104). *L. monocytogenes* spreads via blood circulation from the intestine into the liver and spleen, and grows in these organs (104). At this stage, *L. monocytogenes* is either resolved by the immune system or further proceeds into other organs (104) such as the central nervous system and fetoplacental unit (56). This infection route is mainly due to the capability of *L. monocytogenes* to invade and proliferate within diverse non-phagocytic cells (11, 26, 38, 61), for instance, enterocytes, endothelial cells, hepatocytes and fibroblasts (26, 38). Mechanism of this ability has been extensively studied, resulting in the elucidation of the important virulence factors (41, 42).

Entry into the mammalian cells is mediated by surface invasion proteins such as internalin A (InlA) and internalin B (InlB) (38). InlA plays an indispensable role in the translocation into the intestinal track (11, 38) by interacting with E-cadherin (11, 38, 42,

104), an adhesion molecule located at adherens junctions between epithelial cells (38). InlA binds to E-cadherin in such a species-specific manner that E-cadherins in rats and mice do not function as the receptor to InlA whereas those in other mammalian species including guinea pigs and gerbils interact with InlA (11). This attribute is dictated by the amino acid occupying position 16 in the first extracellular domain of the E-cadherin, which is a glutamic acid in rats and mice but a proline in other susceptible species (11, 41). Meanwhile, InlB contributes to the invasion into a wider range of cell types (11, 38, 42) such as endothelial cells, heptocytes and fibroblasts (41) due to the ubiquitous nature of its receptor, the hepatocyte growth factor receptor Met (11, 38, 42). InlB also interacts with the receptor of the globular part of the complement component Clq (gClq-R)/p32 (42).

Upon the attachment on the cells, *L. monocytogenes* is internalized within a vacuole (42), which is lysed via the synergistic actions of listeriolysin O (LLO), a pore-generating cytolysin (42, 104); and two phospholipases C (PLCs), PC-PLC (PlcA) and PI-PLC (PlcB) (38, 42, 104). The PI-PLC is produced as a proenzyme and transformed into a mature form by the zinc-dependent metalloprotease Mpl (42, 61). After lysis, *L. monocytogenes* is released into the cytosol that is permissible for replication (42). For the multiplication in the intracellular environment, *L. monocytogenes* relies on nutrients available in the cytosol, for instance, glucose-1-phosphate (42, 61), through the mediation by Hpt, a hexose phosphate transporter (42, 61, 104).

In parallel to cytosolic replication, *L. monocytogenes* disseminates intracellularly and into adjacent host cells by synthesizing filaments with the host cell actin by means of ActA protein (42, 63, 104). For cell-to-cell transfer, *L. monocytogenes* moves at the extremity of

the growing actin tails (42, 104) and creates the pseudopod-like protrusion into neighboring cells during the cell-to-cell transfer (42, 63, 104). A recent study provided evidence that the internalin C (InlC) contributes to the cell-to-cell dissemination between polarized epithelial cells by decreasing cortical tension at apical junctions so that the protrusion formation is facilitated (114). Subsequently, the protrusion is engulfed by the adjacent cells (42, 63, 104), forging double-membrane secondary vacuole (42, 104). *L. monocytogenes* is escaped from this vacuole again through the actions of LLO and two PLCs (42, 104) and repeating the cytosolic replication and actin-based movement in the neighboring cells (104).

### *Listeria* pathogenicity islands

In *Listeria* species, three virulence gene clusters have to date been identified and termed *Listeria* pathogenicity island 1-3 (LIPI-1-3) (27, 141). Since LIPI-2 is exclusively found in *L. ivanovii* that is associated with animal listeriosis (27, 141), we will only discuss in further detail LIPI-1 and LIPI-3, both harbored by *L. monocytogenes* (27, 141).

LIPI-1 refers to a genomic region flanked by *prs* and *ldh* in which several well-known virulence genes reside (Fig. 1) (122). This region is identified in three different *Listeria* species, i.e., *L. monocytogenes*, *L. seeligeri* and *L. ivanovii* (122, 141). Virulence genes within LIPI-1 are encoded in a 9-kb gene cluster of six genes, *prfA-plcA-hly-mpl-actA-plcB* (122). Among these genes, *prfA* encodes positive regulatory factor A (PrfA) that regulates the expression of its own gene as well as a majority of virulence genes identified so far, for instance, *hly*, *mpl*, *actA*, *plcA*, *plcB*, *inlA*, *inlB*, *inlC* and *hpt* (123). Genes related by PrfA are transcribed from the promoters with a PrfA box to which the PrfA dimer is bound (123).

PrfA boxes are located ~200 bp upstream to the start codon and characterized with the consequence sequence of TTAACANNTGTAAA with seven conserved underlined nucleotides (123). In addition to the auto-regulation, PrfA expression is also dependent on the sigma factor $\sigma^B$ involved in the stress response (29, 123); and a 5' untranslated region thermo-sensor whose secondary structure blocks the ribosomal binding site ≤30ºC, thus inhibiting translation (29, 47, 123). Meanwhile, the PrfA regulation of the virulence gene regulation is also mediated by transforming PrfA into an active form (29, 47, 123). PrfA activation is enhanced when the cells are inside the cytosol of the mammalian cells (29, 47, 123)or grown in the medium supplemented with activated charcoal (29, 123). On the other hand, PrfA activation is repressed by sugars such as glucose and cellobiose that are transported by the phosphoenolpyruvate-sugar phosphotransferase system (PTS) and abundant in the environment whereas such a phenomenon is not observed in the presence of non-PTS sugar hexose phosphates that are rich in the cytosol (29, 47, 123). LIPI-1 does not differ significantly in GC content from the genome average, suggesting that this cluster is not likely acquired through the horizontal gene transfer (122). The evolution of LIPI-1 will be discussed in more detail in the section on the evolution of *Listeria* species.

LIPI-3 is the most recently identified pathogenicity island, which is located between *LMOf2365_1111* and *LMOf2365_1120* genes in *L. monocytogenes* F2365 (Fig. 2) (25, 27). This pathogenicity island is composed of the eight gene cluster (*llsA*, *llsG*, *llsH*, *llsX*, *llsB*, *llsY*, *llsD*, and *llsP*) encoding listeriolysin S (LLS), a bacteriocin-like modified peptide exhibiting hemolytic and cytotoxic activities (27). LLS contributes to the virulence and intracellular survival since LLS negative mutant lacking *llsB* gene exhibited the reduced

bacterial loads in mouse model and within human polymorphonuclear neutrophils (PMNs) in comparison to the wild type (27). Deletion mutants of other genes in *llsG-llsD* than *llsB* exhibited loss of LLS-derived hemolysis, suggesting that these genes also play indispensable roles for LLS synthesis (25). In fact, *llsX* was employed as a genetic marker in the real-time PCR scheme to detect *L. monocytogenes* isolates with LIPI-3(25). In contrast to LIPI-1 that is found in all strains of *L. monocytogenes* and in other *Listeria* species, LIPI-3 is possessed only by a subset of *L. monocytogenes* lineage I, a population group of *L. monocytogenes* often associated with human listeriosis cases (27). GC content analysis hinted that this pathogenicity island might result from horizontal gene transfer (27). Based on these findings, Cotter et al. speculated that LIPI-3 might have been horizontally acquired by an ancestral strain of lineage I, and might have been lost in some descendants during the evolution process (27).

**Premature stop codon mutations in *inlA***

An intriguing feature of InlA is that many *L. monocytogenes* strains harbor premature stop codon (PMSC) mutations in *inlA* that result in truncated, secreted InlA proteins that are deficient of the C-terminal part including the membrane-anchoring LPXTG motif (70, 136, 143). To date, at least 18 such PMSC mutations in *inlA* have been identified through the surveys of the isolates collected in many countries such as the United States (136, 142, 143), France (70, 101, 113), and Japan (65). PMSCs in the *inlA* were associated with impaired in vitro virulence since strains harboring PMSCs in the *inlA* exhibited reduced invasion efficiency in the human intestinal epithelial cells in comparison with the strains expressing

the full-length InlA (65, 101, 143). Further conclusive evidence was provided by characterizing isogenic mutants with and without a PMSC in *inlA* (100). The PMSCs in *inlA* caused decrease in invasion to human hepatocyte and intestinal epithelial cells as well as attenuated virulence in the guinea pig model (100).

This causal relationship between the *inlA* PMSCs and virulence prompted many researchers to employ these mutations as a genetic marker for human virulence (39, 70, 135, 136, 142, 143). Surveys of isolates with the PMSCs in *inlA* revealed that PMSC mutations in *inlA* were more frequently identified in food isolates than among clinical isolates (70, 136). Specifically, the PMSC mutations in *inlA* were identified among 35% of food isolates collected in France (70) and 48.5% of the RTE food and processing facilities isolates collected by the U. S. Department of Agriculture (USDA) Food Safety and Inspection Service (FSIS) (143), suggesting that a large proportion of *L. monocytogenes* food isolates might be attenuated in virulence. Meanwhile, approximately 5% of human clinical isolates were reported to harbor *inlA* with PMSC mutations (70, 136).

PMSCs in *inlA* have been identified frequently among serotypes 1/2a, 1/2b and 1/2c (70, 143) that are often found among food and environmental isolates (70, 143). Several studies showed that *inlA* PMSC mutations were absent among serotype 4b (70, 143), which is underrepresented among food and environmental isolates (65, 70, 143) but overrepresented among clinical isolates (22, 70). Recently, some serotype 4b isolates belonging to epidemic-associated clones (Epidemic Clones I, Ia, II and III) were reported by van Stelten et al. to carry PMSC mutations in *inlA*; however, the study did not provide the frequency of the isolates carrying *inlA* PMSCs among serotype 4b (136).

Apart from the linkage with food isolates, truncated InlA expression was associated with the virulence attenuation that was often observed among *L. monocytogenes* isolates from the persons without symptoms of listeriosis (101). Also, severity of *L. monocytogenes* infection was lessened in guinea pig model when the animals were previously exposed to *L. monocytogenes* harboring a PMSC in *inlA*, indicating that the virulence-attenuated strains with the truncated InlA might function as a natural vaccine that elicits immune response against further *L. monocytogenes* infection (100).

## Stress Tolerance

Ubiquity of *L. monocytogenes* is owed by its capability to withstand stresses (61, 126, 141). For instance, this microorganism can survive at extreme temperatures from -0.1 to 45˚C, within the pH range of 3.0-9.5, and in the presence of salt up to 10% (141). Of particular note is the fact that *L. monocytogenes* can multiply, albeit at a low rate, at temperatures below 4ºC (17, 141). This psychrotrophic feature permits *L. monocytogenes* to grow in refrigerated foods, increasing its importance as a foodborne pathogen (130, 141).

In agreement with this stress tolerance, studies on stress responses have revealed that *L. monocytogenes* is well-equipped with effective stress response mechanisms (141). Major stress response proteins identified to date include (1) heat shock proteins (HSPs) such as GroES, GroEL, DnaK, DnaJ, HtrA, and Clp proteins for extremely high temperatures (126); (2) cold shock proteins (CSPs), for instance, OpuC, Gbu, OppuA, and Fri, and osmolyte transporters of BetL, Gbu and OpuC for the exposure to low temperatures (17); (3) acid

shock proteins, glutamate decarboxylase (GAD), arginine deiminase (ADI), and F0F1-ATPase systems for acid stress; and (4) osmolyte transporters and osmotic stress proteins such as RelA, Ctc, HtrA, KdpE, LisRK, ProBA, and BtlA for osmotic stress (126). A majority of these stress genes are regulated by the alternative sigma factor $\sigma^B$ (17, 126).

Interestingly, many of the stress response genes contribute to virulence of *L. monocytogenes* (61, 126), which increases in parallel with the ability to tolerate the stresses encountered within the host (126). In addition, adaptive response, in particular to acid, confers cross-protection against a lower pH and other stresses, suggesting that stresses outside the host will prepare *L. monocytogenes* to survive the *in vivo* stresses more effectively (126). Connection between stress responses and virulence is also supported by the finding that $\sigma^B$ regulates *inlA* and *inlB* encoding the adhesins InlA and InlB, and *prfA* encoding the pleiotropic regulator PrfA of pathogenicity genes (61).

## Surface proteins in *L. monocytogenes*

In bacteria, the cell wall directly contacts the environment and infected host (128) and anchors numerous proteins that mediate the interaction between bacteria and extracellular environments (13). Accordingly, the cell wall proteins participate in indispensable cellular functions such as adhesion (9, 13, 128), interaction with the host immune systems (13, 128), invasion (9, 13), transport of nutrients (9, 128), and motility (9).

The genome sequencing of *L. monocytogenes* EGDe revealed that this bacterium encodes 133 surface proteins, corresponding to 4.7% of the total genes (13). These surface

proteins were divided into three classes according to the anchoring mode to the cell wall or membrane: (1) proteins covalently bound to peptidoglycan cell wall via LPXTG motif located at the carboxyl terminus; (2) proteins non-covalently linked to the peptidoglycan layer or membrane by means of carboxy-terminal domains such as GW modules or hydrophobic tail; and (3) lipoproteins imbedded in the membrane through the non-covalent interaction mediated by the amino-terminal region (9, 13).

In addition to the number of the surface protein, the importance of the *Listeria* surface proteins is reflected in the fact that many major virulence factors belong to this class of proteins (9, 13). For instance, the major invasion proteins InlA and InlB possess LPXTG and GW motifs at the carboxyl terminus, respectively (9, 13). Based on the amino terminal sequences, these two proteins are members of the internalin family that is characterized with the leucine-rich repeats (LRRs) (9, 13). LRRs play a role in the interactions between proteins and contribute to adhesion, ligand-receptor interactions and signaling (13). Another example is found in the intra- and inter-cellular movement determinant ActA (42), which bears the hydrophobic tail motif at the carboxyl terminus (9, 13).

*Listeria* surface proteins have also attracted attention due to their unambiguous variability from strain to strain (12, 13, 42). Comparison of *L. monocytogenes* EGDe and *L. innocua* CLIP11262 genomes unveiled that the differences between these *Listeria* species are primarily attributable to their varying surface proteins (13). This trend was also observed when *L. monocytogenes* strains were compared (12, 42). Therefore, the diversity in surface proteins among the *Listeria* species is expected to explicate the strain-specific characteristics including the ability to cause outbreaks.

**Lipoproteins**

Lipoproteins in Gram-positive bacteria perform various cellular functions including substrate-binding and transport as a part of ABC transporters; cell signaling; sporulation; conjugation; adhesion to the host; protein folding and export; and antibiotic resistance (129). Lipoproteins possess the amino-terminal signal peptide that is characterized by a unique consensus sequence called lipobox, $Leu_{-3}Ser/Ala_{-2}AlaGly_{-1}Cys_{+1}$ (7, 13, 128). Lipoproteins are initially translated as prolipoproteins that contain the amino terminal signal peptide (50) and exported as dictated by this sequence (128). According to the studies on Gram-negative bacteria, the enzyme prolipoprotein diacyglycerol transferase (Lgt) mediates the lipidation reaction wherein an N-acyl diglyceride group is transferred from a glycerophospholipid to the thiol group of the conserved cysteine in the lipobox (13, 128). The resulting lipidated prolipoproteins lose the signal peptide by the action of the prolipoprotein signal peptidase II (Lsp) (50, 128), which recognizes and cleaves within the lipobox sequence (50, 128). The third reaction attaches an N-acyl moiety to the amino group of the lipidated cysteine at the amino terminus (50, 128); however, this reaction, which is catalyzed by the enzyme N-acyl-transferase (Lnt) (50), seems to be absent in low GC content Gram-positive bacteria since no homologs of the *lnt* gene were found in the genomes of these bacteria (50, 128). Upon the completion of the processing steps, the mature lipoproteins are fastened to the membrane through the diacyglycerol moiety linked to the amino terminal cysteine (50, 128).

The *L. monocytogenes* EGDe genome was predicted to encode 68 lipoproteins, suggesting that lipoproteins are the largest subgroup of the surface proteins (9, 13). With

regards to putative functions, 28 lipoproteins were putative substrate-binding components of ABC transport systems, 15 were predicted to perform various enzymatic and metabolic functions, and the others were of unknown function (7). However, most studies on *Listeria* lipoproteins to date concerned the processing of the lipoproteins. As a member of the low GC Gram-positive bacteria, *L. monocytogenes* is expected to maturate lipoproteins through the actions of Lgt and Lsp (50). The *lgt* deletion affected neither cell morphology of *L. monocytogenes* EGDe nor growth of the bacteria in the rich medium (7). However, the mutant displayed mild growth inhibition in minimal medium, elevated susceptibility to cationic peptides, decreased growth in both non-phagocytic and phagocytic cells, and virulence attenuation in mice (7, 90). Of particular interest was that the *lgt* deletion resulted in selective release of the 26 mature lipoproteins into the supernatant, indicating that the processing mediated by Lsp can be carried out on the prolipoproteins without the diacyglycerol moiety (7). In the case of *lsp*, deletion did not affect cellular morphology or growth in rich medium but resulted in decreased phagosomal escape and *in vivo* virulence (116). In addition, Lsp was required for the protein's translocation across the membrane since the above-mentioned release of lipoproteins was reduced when the *lgt* deletion mutant was treated with globomycin, an inhibitor of Lsp (7).

## Phylogeny and Genomic Comparisons of *Listeria* species

### Phylogeny of *Listeria* species

The genus *Listeria* is composed of eight species: *L. monocytogenes, L. seeligeri, L. innocua, L. welshimeri, L. ivanovii, L. grayi, L marthii* and *L. rocourtiae* (60, 84, 122). *L. marthii* and *L. rocourtiae* have been recently identified from the natural environment of the Finger Lakes region of New York and a lettuce from Salzburg (Austria), respectively (60, 84). Neither of these species exhibited hemolysis on the blood agar or in vitro virulence, suggesting that they are non-pathogenic *Listeria* species (60, 84). For *L. marthii*, the absence of LIPI-1 and *inlA* was confirmed with PCR and sequencing (31, 60). Among the *Listeria* species, only two are considered to be pathogenic: *L. monocytogenes* and *L. ivanovii* (122). *L. monocytogenes* causes listeriosis in humans and animals while *L. ivanovii* is mainly responsible for bovine abortions (122) although a few cases of gastroenteritis and bacteremia were reported which were caused by *L. ivanovii* (62).

To date, the phylogenetic relatedness and evolution of *Listeria* species have been investigated by means of Multilocus Enzyme Electrophoresis (MLEE) (10), sequences of 16S rRNA genes (137), and concatenated DNA sequences of select genes (30, 31, 122). These studies consistently showed that *L. grayi* is separated from the other *Listeria* species (Fig. 3); however, each study presented different relationships among the other *Listeria* species (10, 30, 31, 122, 137). The MLEE analysis by Boerlin et al. proposed that *L. monocytogenes, L. innocua,* and *L. welshimeri* form one cluster in which *L. innocua* and *L. welshimeri* are grouped as sister branches, whereas *L. seeligeri* and *L. ivanovii* compose the

other (10). Phylogenetic relationships based on the 16S rRNA gene sequences divided *Listeria* species into the same two clusters (137). However, this analysis placed *L. innocua* and *L. monocytogenes* as sister species in the group consisting of *L. monocytogenes, L. innocua* and *L. welshimeri* (137). In the phylogenetic tree based on the concatenated DNA sequence of 16S and 23S-rRNA, LIPI-1 flanking genes (*prs*, *ldh* and *vclB*) and *iap*, Schmid et al. placed *L. monocytogenes* and *L. innocua* in one group; and *L. welshimeri, L. seeligeri* and *L. ivanovii* in a separate cluster in which *L. welshimeri* was located in the deepest branch (Fig. 3) (122). Two recent studies employed concatenated DNA sequences of 10 and 100 select genes, respectively, to determine the phylogenetic relationships among *Listeria* species including *L. marthii* (30, 31). These examinations identified two clades of *Listeria* species, one consisting of *L. seeligeri* and *L. ivanovii* and the other composed of *L. welshimeri, L. innocua, L. marthii* and *L. monocytogenes* (30, 31). In the latter lade, *L. welshimeri* was placed as a basal lineage to the other species in the same group and *L. marthii* was located as a sister group of *L. monocytogenes* (Fig. 3) (30, 31). The close relationship between *L. marthii* and *L. monocytogenes* was also observed in the phylogeny based on concatenated partial *sigB, gap* and *prs* sequences (60). Another new *Listeria* species, *L. rocourtiae*, was placed separately from other *Listeria* species in the phylogram inferred from 16S rRNA gene (Fig. 3) (84).

Excluding *L. grayi*, the common ancestor of *Listeria* species seems to possess pathogenic potential (30, 30, 122). After comparing the DNA sequences of LIPI-1, Schmid et al. proposed that this pathogenicity island was inherited from the common progenitor (122). They hypothesized that LIPI-1 was lost through two deletion events leading to the current *L.*

*innocua* and *L. welshimeri*, respectively (122). More recently, den Bakker et al. reached a similar conclusion that the ancestral strain of *Listeria* species harbored LIPI-1 through the inference drawn from the phylogram based on 10-loci MLST and absence/presence of LIPI-1(30). This analysis suggested that LIPI-1 might have been deleted at least five times during the evolution of *Listeria* species, thus resulting in pathogenic and non-pathogen *Listeria* species and atypical *Listeria* strains, i.e. *L. innocua* strains with LIPI-1 and *L. seeligeri* strains without LIPI-1 (30). In another study employing whole genome comparisons, den Bakker et al. hypothesized that the ancestral strain of *Listeria* species might be equipped with LIPI-1 and many internalins including *inlA*, *inlB*, and *inlC* (31). This scenario of the virulence gene deletion is consistent to the large-scale deletions observed in *L. innocua* and *L. welshimeri* genomes in comparison with *L. monocytogenes* (64), which will be discussed further in the next section.

**Genomic comparisons of different *Listeria* species**

Investigation of *Listeria* species genomes provides essential information on genetic characteristics, fluidity and evolution of *Listeria* species. With the exception of *L. grayi* and *L. rocourtiae*, all *Listeria* species have been subjected to the comparison at the genomic level (31, 40, 56, 64, 99). Previous studies revealed that *Listeria* species possess circular genomes of the similar sizes, ranging from 2.8 to 3.2 Mbp (12, 31, 56, 64, 127), that encode approximately 2,700 to 3,000 genes (12, 31, 56, 64). The genome encoded numerous surface proteins, transporter proteins, and transcriptional regulators, which are compatible with the ability to inhabit various environments (12). The *Listeria* pan-genome is estimated to consist

of 6,494 genes with 1,994 commonly shared genes and 4,500 accessory genes (31). Analysis of the 13 *Listeria* species genomes by den Bakker et al. suggested that most of the core genes (2,032 genes) have been discovered but a large proportion of accessory genes remains unidentified (2,918 accessory genes observed) (31). The GC content was reported for *L. monocytogenes* (38%), *L. innocua* (37%), *L. seeligeri* (37.4%) and *L. welshimeri* (36.4%) (12, 56, 64, 127).

*Listeria* genomes are characterized by the conserved synteny with limited horizontal gene transfer (12, 31, 64) although inter- and intra-species recombinations have been detected (30, 40). Genetic changes involving large genome segments have not been reported and the conserved synteny was disrupted mainly by prophages, monocins, strain specific genes, and to a lesser degree by translocation (12). Prophages have been identified in sequenced *L. welshimeri*, *L. monocytogenes*, *L. innocua*, *L. seeligeri*, and *L. ivanovii* (31, 56, 64, 99, 127). Notably, *comK* encoding a natural competence protein was often disrupted by prophage integration (12, 31). Monocins refer to *Listeria*-specific bactericidal substances encoded in cryptic prophages (87). Monocins have been identified in *L. monocytogenes*, *L. innocua*, *L. welshimeri*, *L. marthii* and *L. ivanovii* (12, 31).

Strain-specific genes are dispersed in various genomic locations (12) and tend to deviate from the average GC content of the genome, suggesting the acquisition through horizontal gene transfer (99). These genes disproportionately code for the virulence- or environmental adaptability-associated proteins; hence, they are anticipated to account for the strain-specific characteristics of *Listeria* (56). Interestingly, many genes specific to *L. monocytogenes* seem to result from extensive deletions that occurred in the non-pathogenic

*L. innocua* and *L. welshimeri* genomes in the evolutionary process from a pathogenic progenitor (56, 64). Hain et al. found that in total 249 genes are commonly missing from *L. innocua* and *L. welshimeri* in comparison to *L. monocytogenes* (64). In addition, *L. innocua* exhibited 63 specific deletions whereas 233 genes were absent specifically from *L. welshimeri* (64). Most of the deletion events were observed in the same genetic location, and the deletion was usually extended in *L. welshimeri* relative to *L. innocua*, thus resulting in the smaller genome size of *L. welshimeri* (64). Notably, genes absent in *L. welshimeri* and *L. innocua* were pronounced particularly among those encoding proteins possessing LPXTG- and LRR motifs, many of whose members are involved in the adhesion and invasion into the mammalian cells (64).

Although rarely observed, translocation also takes a part in interrupting the synteny of *Listeria* genomes (12). The example of translocation was identified in *L. welshimeri* genome where a series of genes encoding the F0F1-ATP synthase were found in a different location than those in the *L. monocytogenes* or *L. innocua* genomes (64).

## Serotyping and population structure of *L. monocytogenes*

Studied conducted to date have agreed that *L. monocytogenes* population consists of several evolutionarily different lineages. However, the number of the lineages has continuously changed. Early studies identified two groups termed lineage I and II by means of MLEE, PFGE, ribotyping and sequence of a virulence gene (73, 117). An additional lineage, lineage III, was first discovered by Rasmussen et al. (117), and the presence of three

distinct lineages has received support from studies employing PCR-restriction fragment length polymorphism (PCR-RFLP), ribotyping (117), multilocus sequence of LIPI-1 gene cluster (144) and of housekeeping genes(94, 113), and a DNA array of the strain-specific genes (37). More recently, four lineages of *L. monocytogenes* were proposed by dividing lineage III into two separate evolutionary groups (30, 32, 106, 142). In this review, we will describe the population structure of *L. monocytogenes* on the basis of the three-lineage classification.

The serotyping classification divides *L. monocytogenes* into 13 serotypes according to the reactivity to somatic (O) and flagellar (H) antigens: serotypes 1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e, and 7 (51). An intriguing link between serotypes and *L. monocytogenes* population structure is reflected in the fact that all the serotypes of *L. monocytogenes* except for serotype 4b exclusively belong to one of these lineages, following the classification based on the somatic antigens (73, 117). As well as the serotypes, each lineage is dissimilar to one another in terms of the frequently associated sources, homogeneity within the lineage (73), and possibly recombination rate (32, 113).

**Lineage I**

Lineage I is comprised of serotypes 4b, 4d, 4e, 1/2b, and 3b (37). As suggested by serotypes 4b and 1/2b, which are implicated in many outbreaks and sporadic cases, lineage I is overrepresented among clinical samples (22)**.** Factors contributing to the association of lineage I to clinical cases have not been elucidated yet; however, additional pathogenicity island LIPI-3 was exclusively found in a fraction of lineage I isolates (27) and PMSC

mutations in *inlA* were rarely present among serotype 4b (70, 136, 143). Among the food isolates, serotype 1/2b is relatively often identified compared with serotype 4b (65, 70, 143). For instance, serotype 1/2b was identified at the second highest prevalence (39.5%) after lineage II serotype 1/2a (41.4%) in the survey of 501 *L. monocytogenes* isolates from RTE foods and processing environments whereas serotype 4b represented only 6.8%(143).

Lineage I exhibits a high degree of clonality in comparison to the other *L. monocytogenes* lineages (32, 94, 106, 113, 144), primarily owing to a low rate of recombination (32, 94, 106, 144) and positive selection (106). Possible molecular bases of this low changeability of the genome include effective restriction-modification (RM) systems, efficient mismatch repair systems, or decreased competency (106). In spite of this homogeneity, some studies further divided lineage I into several subgroups. For instance, two lineage I subgroups were identified through DNA-DNA hybridization between the strain-specific *Listeria* DNA array and the genomic DNA from various isolates: one consisting of serotypes 4b, 4d, and 4e (serotype 4b complex (22)); and the other including serotypes 1/2b and 3b (37).

The population history analysis has indicated that this lineage might experience a recent bottleneck either separately (94, 144) or along with lineage II (32), which is estimated to have occurred between one to one-half million years ago (94). Based on the observation that the first glaciation of the Pleistocene started in this period, den Bakker et al. suggested that the population bottleneck might be connected with reduced hosts and lower temperature, both of which bear associations with the characteristics of *L. monocytogenes* (32).

**Lineage II**

Lineage II is constituted with serotypes 1/2a, 1/2c, 3a and 3c (117, 144). Isolates belonging to this lineage are frequently found in foods and natural environments (65, 143) although serotypes 1/2a and 1/2c are also identified among clinical isolates (70) and sometimes implicated in outbreaks (55, 105).

Studies on the population structure of *L. monocytogenes* suggest that lineage II differ considerably from the other two lineages (37, 113, 144). For instance, when Doumith et al. compared the partial genome sequence of a serotype 4b strain with the complete sequence of the serotype 1/2a EGDe, 8 % of the serotype 4b genes were not found in EGDe. Since there were 9.5% of EGDe specific genes in comparison with *L. innocua*, they concluded that lineage II differs from lineage I to a degree equivalent to the interspecies level between *L. monocytogenes* and *L. innocua* (37). More recently, Ragon et al. calculated on the basis of the multilocus sequence typing (MLST) of the seven housekeeping genes that the nucleotide divergence of lineage II amounts to 4.99% relative to lineage I and to 7.57% in comparison to lineage III (113).

Genes specific to or highly conserved in lineage II are mainly related to sugar transport and metabolism (37). According to the study using the species-specific *Listeria* array, twelve out of nineteen lineage II-specific genes were putatively involved in sugar metabolism (37). Also, most of the PTS genes conserved among lineage II strains were absent in serotypes 4b and 1/2b, which belong to lineage I (37).

Lineage II represents a heterogeneous group of strains in comparison to lineage I (32, 94, 106, 113). According to Ragon et al., the diversity within lineage II was computed to be

0.61%, which was higher than that of lineage I, 0.33% (113). Although Ragon et al. reported a similarly limited recombination in all three lineages (113), more studies proposed that this heterogeneity results from a more active recombination than lineages I (32, 94, 106) in addition to positive selection (106). For instance, according to the estimation by den Bakker et al, lineage II has six times higher recombination rate in comparison to lineage I (32). Given the transfer of the short DNA fragments by recombination and presence of genes whose homologs in *Bacillus subtilis* are involved in transformation, den Bakker et al. speculated that this high recombination rate might result from transformation and that lineage II might exhibit competency under particular circumstances (32). Another hypothesis regarding the underpinnings of the fluidity of the lineage II genomes concerns frequent exposure to bacteriophages that might be abundant in the ecological niches inhabited by lineage II (106). Also, den Bakker et al. suspected that recombination might contribute to the ability to survive a wide variety of environments (32).

As expected from this diversity present in lineage II, some studies observed that lineage II contain multiple subgroups (37). Based on the genetic content revealed by the DNA array composed of strain-specific genes, Doumith et al. proposed two lineage II subgroups, in which serotypes 1/2a and 3a constitute one subgroup whereas serotypes 1/2c and 3c belong to the other (37).Evolutionary history studies, as mentioned in lineage I, proposed that only lineage I underwent a recent bottleneck (94, 144); however, den Bakker et al. suggested that the bottleneck has influenced both lineages I and II (32).

**Lineage III**

Lineage III is rarely found and comprises three serotypes, i.e., 4a, 4c, and 4b (117). This lineage has caused a few outbreaks (130) and occasional sporadic cases (144); however, compared with the other *L. monocytogenes* lineages, lineage III is the least frequently linked with human listeriosis (144). Factors culminating in this phenomenon have been suggested by several researchers. For instance, some hypothesized that lineage III isolates are impaired in the ability as a human pathogen (106). However, this hypothesis was encountered with contradicting findings from several studies. Comparison of different *Listeria* strains via the DNA array found that all *L. monocytogenes* strains tested harbored all known virulence genes regardless of phylogenetic lineages (37). Later, Roberts et al. reported that the lineage III isolates tested form plaques in the cell pathogenicity test (117); and Liu et al. showed that a majority of lineage III strains exhibited in vivo virulence in the intraperitoneal mouse infection (86). Also, no premature stop codon mutation in *inlA* has been observed among the lineage III (and IV) isolates (142). Meanwhile, lineage III was rarely found among food isolates (142-144). This finding brought about the speculation that listeriosis cases caused by lineage III are rare due to decreased exposure to this lineage through foods (144). Indeed, Roberts et al. suggested that lineage III might not be well-equipped against environmental stresses encountered in foods or food-processing plants as expected from the following observations: a long lag period at the refrigeration temperature; vulnerability to thermal treatment; impaired biofilm formation; and presence of a premature stop codon in the stress response regulator *sigB* (117). Rather with human listeriosis cases and foods, lineage III is

often associated with animal listeriosis since lineage III is overrepresented in isolates from animals with clinical listeriosis (117).

Analysis of a large set of 46 lineage III isolates utilizing partial *actA* and *sigB* sequences indicated that lineage III exhibits the most pronounced diversity among *L. monocytogenes* lineages (117). Later, this proposal was supported by a study conducted by Ragon et al. using the MLST data of the seven housekeeping genes, wherein they calculated the genetic diversity within lineage III at 1.25%, conspicuously higher than lineage I (0.33%) or lineage II (0.61%) (113). However, Ragon et al. found a highly restricted recombination in lineage III (113).

Expectedly from this noticeable heterogeneity, lineage III was further classified into several subgroups. In one study using the *Listeria* DNA array composed of strain-specific genes, lineage III was divided into two subgroups according to serotypes, i.e., one subgroup of serotype 4a and the other of serotype 4c (37). In the study by Roberts et al., the partial *actA* and *sigB* sequence-based analysis identified three subgroups within lineage III, i.e., subgroups IIIA, IIIB, and IIIC, among which subgroup IIIA is the most heterogeneous (117). Genetic transfer appears to be limited from one subgroup to another, and no association was detected with serotypes (117, 142), with isolation source, or with cytopathogenicity in a cell culture assay (117). In terms of rhamnose fermentation often used for biochemical tests for *L. monocytogenes*, subgroup IIIA utilized rhamnose as observed for typical *L. monocytogenes* isolates, whereas subgroups IIIB and IIIC did not (117). Moreover, subgroup IIIA isolates were less likely to harbor the *lmaA* gene, which is putatively involved in virulence, in comparison with the other lineage III subgroups, suggesting that some lineage III isolates

could elude the PCR-based *L. monocytogenes* detection method targeting this gene (117). Phylogenetic studies revealed that subgroup IIIB differed from other *L. monocytogenes* lineages including subgroups IIIA and IIIB that formed one cluster (32, 106, 142). Hence, lineage IIIB was termed lineage IV by Ward et al. as the fourth population lineage of *L. monocytogenes* (142).

Related to other *L. monocytogenes* lineages, Ward et al. suggested that lineage III forms a sister group of lineage I (144). Similarly, Ragon et al. concluded that lineage III is more closely related to lineage I than to lineage II since lineage III showed 5.3% nucleotide divergence against lineage I, whereas that against lineage II was 7.57% (113). In the study by Orsi et al., which identified four lineages, the phylogenetic analysis showed that lineage IIIA/C remains closely related to lineage I whereas lineage IIIB is separated from other *L. monocytogenes* groups (106). A recent 10-loci MLST on isolates representing six *Listeria* species identified many common alleles between lineages III and IV as well as some alleles shared between lineage IV and *L. innocua* (30).

**Evolution of different *L. monocytogenes* serotypes**

Accumulation of the detailed knowledge on the phylogenetic distribution of serotypes enabled several studies to attempt to elucidate the evolutionary relationships among *L. monocytogenes* serotypes (37, 113). By means of a *Listeria* DNA array spotted with strain-specific genes, Doumith et al. proposed that the divergence began as serotypes 1/2b and 1/2c were separated from a common ancestor (37). Subsequently, the evolution from serotype 1/2b proceeded into serotypes 4b, 4a, and 4c, from which *L. innocua* branched mainly

through gene loss whereas the divergence of serotype 1/2c resulted in lineage II (37). Similarly, in the sequence analysis of the seven housekeeping genes, Ragon et al. considered that serotype 1/2b, the ancestor of lineage I, was evolved into serotype 4b (113). However, they disagreed on the emergence of *L. innocua* from *L. monocytogenes* (113). Also, Ragon et al. considered serotype 1/2a as the likely ancestor of lineage II, from which serotype 1/2c is derived (113). Intriguingly, serotype 1/2a strain EGDe, an exemplary strain that has been extensively studied to date, was located in the middle of this divergence from serotype 1/2a to 1/2c (113).

**PCR-based Serotyping of *L. monocytogenes***

In order to trace back the incriminated source of *L. monocytogenes* contamination (36, 113) and study the ecological distribution (113), it is imperative to investigate the characteristics of *L. monocytogenes* strains; hence, various subtyping techniques have been developed to differentiate *L. monocytogenes* strains (36). The commonly employed subtyping methods include serotyping, MLEE, PFGE, PCR-RFLP and MLST (117). Although low in resolution, serotyping has been considered as an indispensable subtyping tool for *L. monocytogenes* since (1) it enables a rapid screening of serotypes frequently involved in clinical cases, thus reducing the number of isolates that will be further characterized with higher-resolution methods such as PFGE in the outbreak investigation ; (2) it provides information for long-term microbiological surveillance of the listeriosis cases in humans; and (3) serotyping is useful in assessing the food chain and processing plants for contaminating *L. monocytogenes* (36).

Traditionally, serotypes were determined by reactivity with antisera (51); however, this method has displayed disadvantages of high cost, low accessibility, requirement of experienced personnel and occasional inconsistency (36). As an attempt to overcome these limitations, PCR-based serotyping methods were developed (18, 36). For instance, Doumith et al. described a multiplex PCR that divides *L. monocytogenes* isolates into four groups employing primers annealing to *Listeria* genus-specific *prs* and the genes that are specific to phylogenetic lineages or their subgroups within a lineage of *L. monocytogenes* (36). To illustrate, *prs* confirms the *Listeria* genus; then, two genes (ORF2819 and *lmo0737*) specific to lineages I and II, respectively, distinguish these lineages (36). Based on the serotype 4b complex-specific gene (ORF2110), lineage I is further divided into serotype 4b complex and a group comprised of serotypes 1/2b, 3b, and 7 (36). The gene specific for serotypes 1/2c and 3c (*lmo1118*) separates serotypes belonging to lineage II into serotypes 1/2a and 3a; and serotypes 1/2c and 3c (36).

Recently, a group in Australia reported that some serotype 4b isolates produce an extra band corresponding to *lmo0737* in this multiplex PCR, challenging the specificity of the genes targeted in the multiplex PCR (69). Our lab also identified such atypical serotype 4b isolates and these isolates will be characterized in Chapter 2. Also, our routine multiplex PCR serotyping revealed a few cases of incongruence between the serotypes determined by this PCR scheme and serological method and between the multiplex PCR-based serotypes and lineage designation determined by other molecular subtyping methods such as multilocus genotyping (MLGT (39, 142)). For the two additional *Listeria* species identified after this multiplex PCR was designed, *L. marthii* generated only *prs* band whereas *L. rocourtiae*

failed to amplify either of the target genes (60, 84). Indeed, *L. rocourtiae* reacted neither of the antisera against known somatic and flagella antigens, suggesting that this species belongs to a novel serotype (84).

## Epidemic clones

The comparison of the strains implicated in outbreaks showed that a few clonal groups of *L. monocytogenes* have been involved with the majority of the listeriosis outbreaks which differ from each other geographically and chronologically (22, 72, 73). Initially discovered by means of MLEE, this finding was also supported by PFGE, ribotyping (22, 72), multiple-locus variable-number tandem-repeat analysis (MLVA) (124), multi-virulence-locus sequence typing (MVLST) (21, 89) and MLST(32). These groups of epidemic-associated strains are referred to as epidemic clones (22, 72). To date, five different epidemic clones have been reported and termed Epidemic Clones (ECs) I-V (22). Consistent to the prevalence of serotype 4b in outbreak strains, all ECs belong to serotype 4b with the only exception of ECIII, which is of serotype 1/2b (22).

### Epidemic Clone I

Epidemic Clone I (ECI) has attributed to numerous major outbreaks that occurred in North America and Europe (22, 72, 73) such as Nova Scotia outbreak in 1981, which first showed the link between listeriosis and foods, in this case, contaminated coleslaw (22, 130, 138, 145); California outbreak in 1985 associated with Mexican-style cheese; soft cheese

outbreak in Switzerland in 1983-1987; and jellied pork tongue outbreak in France in 1992(22). Early identification and epidemiological significance prompted this clonal group to be subjected to the most extensive investigation among ECs; thus, several ECI-features have been revealed (22, 72). For instance, a distinct restriction fragment length polymorphism profile was observed among ECI strains in a genomic region required for the growth at 4˚C (22, 72) and in a putative mannitol transporter gene (22, 72, 134). In addition, the genomic DNA of ECI strains shared resistance to the restriction endonuclease Sau3AI but susceptibility to the isoschizomer MboI (43). Since resistance to Sau3AI and MboI is determined by methylation at cytosines and adenines, respectively (92), this attribute suggests methylation of cytosine in GATC sites (43).

ECI-specific DNA sequences were first discovered by Herd and Kocks in 2001 (67). In this study, they conducted subtractive hybridization against the serotype 1/2a EGDe of the genome of F4565, a ECI strain involved in the 1985 Mexican style cheese outbreak in Los Aangeles, in order to identify DNA fragments present only in the outbreak strain (67). Then, these fragments were hybridized to multiple *L. monocytogenes* strains encompassing different serotypes and isolation sources as well as *L. innocua* and *L. ivanovii* (67). Hybridization patterns revealed several ECI-specific fragments; some shared among ECI and other serotype 4b isolates; and others shared among ECI and serotypes 1/2b and 3b (67). Subsequently, another ECI strain, F2365, implicated in the 1985 LA cheese outbreak, was fully sequenced and compared with other *L. monocytogenes* genomes and *L. innocua* CLIP11262 (99). This comparison at the genome level confirmed the findings by Herd and Kocks (22).

Some of these fragments identified by Herd and Kocks have been employed as the genetic markers to distinguish ECI strains (147). In this review, three of these fragments will be discussed in detail: Fragments 85, 17B, and 133.

***Fragment 85***

Fragment 85 was identified by Herd and Kocks as an ECI-specific DNA fragment (67). Protein sequence analysis revealed that the deduced amino acid sequence possesses homology to a methyltransferase in *Lactococcus lactis* which contains the target recognition domain similar to that of M.Sau3AI (67). The whole genome sequence of F2365 revealed that the gene in Fragment 85 corresponds to *LMOf2365_0327*, which was annotated as C5-cytosine-specific methyltransferase (99). Upstream to this methyltransferase lies a gene encoding a putative Sau3AI-like restriction endonuclease (*LMOf2365_0325*) (99). Between two genes, a putative DNA-binding protein gene is located (99), which might function as a control (C) protein involved in the regulation of the Sau3AI-like RM system (98). Presence of this putative Sau3AI-like RM system was associated with the resistance to Sau3AI digestion that is observed in ECI strains (22, 147). A recent study provided evidence that the mutant lacking the methyltransferase gene in this Sau3AI-like RM system is susceptible to the Sau3AI digestion (146). The GC content of these three genes ranges from 28 to 31%, which is considerably lower than the *L. monocytogenes* genome average of 38%, suggesting that this RM system might be obtained via horizontal gene transfer (147).

The Sau3AI-like RM system, in particular the methyltransferase gene, has been used as a ECI-specific genetic marker (probe 85M designed by Yildirim et al. (147)) (23, 46,

147).However, this RM system has been also identified in some strains of other serotypes of *L. monocytogenes* than serotype 4b. For instance, Chen and Knabel reported serotype 1/2a and 4a *L. monocytogenes* isolates that are positive for the methyltransferase-targeting PCR (18). Indeed, Yildirim et al. provided evidence through *in silico* analysis and sequencing that some serotype 4a and 1/2a strains harbored a highly homologous RM system to the Sau3AI-like RM system identified in ECI at the genomically equivalent location (146). Taken together, the Sau3AI-like RM system is not exclusively present in the serotype 4b ECI strains; however, this would not hamper the utility of this genetic region as an ECI-specific genetic marker if employed for serotype 4b *L. monocytogenes*.

***Fragment 17B***

Fragment 17B is a 1.7 kb DNA fragment discovered by Herd and Kocks that hybridized only with ECI strains but whose function was not known (67). Subsequently, Yildirim et al. included more ECI isolates for hybridization with a DNA probe (17B) based on Fragment 17B and confirmed that probe 17B is conserved among ECI strains but not among other serotype 4b isolates, thus establishing its status as an ECI-specific genetic marker (147). The below-average GC content of ca. 30% indicates that a horizontal gene transfer might have contributed to the acquisition of Fragment 17B (147). In F2365 genome, Fragment 17B corresponds to *LMOf2365_2798* and this gene was employed as the target for ECI in the multiplex PCR that discriminates epidemic clonal groups (18). Two downstream genes (*LMOf2365_2799* and *LMOf2365_2800*) were identified to be ECI-specific via genome comparison (18).

### *Fragment 133*

Fragment 133 was found by Herd and Kocks as a gene fragment with unknown function that generated hybridization signal with ECI strains and those of serotypes 1/2b or 3b (67). In parallel with Fragments 85 and 17B, Yildirim et al. verified the presence of Fragment 133 in ECI members but not in other serotype 4b isolates through hybridization with the DNA probe originating from this fragment (133) (147). However, PCR-based screening by Franciosa et al. and Southern hybridization by Eifert et al. revealed that some serotype 4b isolates outside ECI, including ECV, also harbor probe 133(43, 46). Taken together, probe 133 does not appear to be strictly linked to ECI even within serotype 4b, thus limiting its use to an auxiliary marker to probes 85M or 17B.

Similar to the other ECI-specific fragments, the GC content of Fragment 133 amounts to approximately 30% (147). This figure is significantly below the genome average, thus suggesting a possible lateral transfer (147). In F2365, this region of 748 bp (accession number: AJ410381) includes parts of *LMOf2365_2628* and *LMOf2365_2629*.

## Epidemic Clone II

Epidemic Clone II (ECII) is associated with two multi-state outbreaks in the US, the 1998-1999 hot dog outbreak (22, 72-74) and 2002 turkey deli meat outbreak (22, 72, 74). Strains of these two outbreaks shared the unique PFGE pattern (74) and ribotype (DUP-1044 and DUP-1044A) (21, 22). Also, these outbreak strains commonly exhibited a diversification in a genetic region, Region 18, which is conserved among non-ECII serotype 4b *L. monocytogenes* (45, 74). These features contributed to the assignment of the strains

implicated in the 1998-1999 and 2002 outbreaks into an independent epidemic clonal group (22).

Diversified Region 18 in ECII enabled several DNA probes to be utilized as ECII-specific markers (23, 45, 46, 74). Meanwhile, whole genome sequencing was conducted with H7858, a food isolate of the 1998-1999 outbreak (99). Comparison of this genome with other *Listeria monocytogenes* strains triggered the identification of more ECII-specific markers (23, 74). In this review, frequently used ECII-specific probes will be discussed in more detail.

***Region 18***

Region 18 was first identified by Herd and Kocks as two DNA fragments (Fragments 7 and 18) whose functions are not known and that bound to all serotype 4b isolates examined (67). Evans et al. found that Fragments 7 and 18 constitute parts of the same genomic region and that ECII strains are not positive for the genetic markers derived from Fragments 7 and 18 (4bSF7 and 4bSF18) in PCR and DNA hybridization (45). Thus, this genomic region including Fragments 7 and 18 was called Region 18 and defined through DNA comparison as the region flanked by serotype 4b-specific *wap* gene encoding a putative wall-associated protein (*LMOh7858_0479*) and *inlA* gene (*LMOh7858_0498*) (Fig. 4) (22, 23).

Apart from 4bSF7 and 4bSF18, four other Region 18-derived markers have been described: H1 (or H18RP9/10), H2 (or H18RP11/12) (23, 43, 74), ECIIC-WAP and non-ECIIC-WAP (23). ECII is negative for probes 4bSF7 and 4bSF18 (23, 45), and non-ECIIC-WAP but positive for H1, H2 and ECIIC-WAP; meanwhile, non-ECII strains such as ECI

exhibit the opposite results (23). However, neither of these genetic markers discriminates ECII unambiguously (23). Probes 4bSF18 and non-ECII-WAP do not hybridize with some of non-ECII strains whereas probes H1, H2 and ECIIC-WAP show signal with non-ECII strains in the DNA-DNA hybridization (23). In particular, ECV exhibited the same pattern as ECII for Region 18-based probes (23), implying that this region is shared between ECII and ECV(22). Further diversification is expected in Region 18 among *L. monocytogenes* since a few non-ECI, non-ECII strains hybridized all of probes H1, H2, and 4bSF18 (23, 74).

***Region 1168***

Region 1168 was identified as a strain-specific region through the comparison of the ECII H7858 genome with F2365 (ECI), F6854 (ECIII), serotype 1/2a EGDe and *L. innocua* CLIP11262 (23). This genomic region corresponds to a gene cassette consisting of *LMOh7858_1161-LMOh7858_1174*, which is adjacent to the listeriolysin S gene cluster (23). In this cassette, *LMOh7858_1168* was used to design DNA probe Hsp001219 (or Hsp01219 (74)) (23). This probe was first described by Kathariou et al. (74) and found to be unique to ECII strains (23).

***Region 0088***

Region 0088 was another H7858-specific region identified in parallel with Region 1168 (23). In the genome of H7858, this region corresponds to *LMOh7858_0088* (23). Probe Hsp00121 derived from this gene was first designed by Kathariou et al. simultaneously to

probe Hsp001219 (74). This probe hybridized with ECII and some serotype 4b isolates such as ECIa, suggesting that Region 0088 is not exclusively conserved among ECII strains (23).

***Region 2753***

Region 2753 corresponds to a gene cassette flanked by a putative lipoprotein gene (*LMOh7858_2752*) and *rpsI* gene (*LMOh7858_2765*) (23). This gene cassette was reported to confer to ECII resistance to phage infection at temperatures below 37ºC (23, 75). Initially, this region was identified as a H7858-specific region through genome comparison (23). Four DNA probes derived from this region (Hsp2753, HspDelcon54, Hsp2759 and Hsp2764) hybridized exclusively with ECII, suggesting that Region 2753 is unique to ECII (23).

**Epidemic Clone III**

Epidemic Clone III (ECIII) was first identified through the outbreak investigation of the turkey deli meat listeriosis outbreak that encompassed eleven states in 2000 in the US (22, 72, 73, 102). The PFGE pattern of this outbreak strain was observed for a strain that was associated with a sporadic case in 1988 linked to turkey frankfurters (22, 72, 73). Of particular note is that the foods implicated in this sporadic case and the multistate outbreak were produced in the same processing facility, suggesting that this strain habituated the factory for the twelve years (22, 72, 73).

Dissimilar to other ECs that are of serotype 4b, ECIII strains belong to serotype 1/2a within lineage II (21, 22, 72, 73). Apart from the above-mentioned PFGE, ECIII share the same ribotype of DUP-1053A and MVLST (21). Currently, the genome sequence of four

ECIII strains has been finished: F6954 (99) and F6900 (105), the 1988 sporadic case isolates each isolated from the contaminated frankfurter and patient, respectively; and J0161 and J2818, the human and food isolates, respectively, from the 2000 outbreak (105). Comparison of the 1988 and 2000 isolates revealed notable diversification in the prophage integrated into *comK* between these two groups and presence of the plasmid only in the 2000 human isolate (105).

Through the genome comparison with other strains, Chen and Knabel identified 16 ECIII-specific genes out of which eight genes are located in a single locus (18). One of these ECIII-specific genes, *LMOf6854_2463.4*, was chosen as a target for ECIII in a multiplex PCR that detects epidemic clonal groups of *L. monocytogenes* (18).

**Epidemic Clone IV**

Epidemic Clone Ia (ECIa) was first discovered on the basis of ribotype (DUP-1042 and DUP-1042B), MLEE and PFGE profiles shared among the serotype 4b *L. monocytogenes* strains associated with the Boston vegetable outbreak in 1979; Boston milk outbreak in 1983; and UK pâté outbreak in 1989 (21, 22). All three subtyping methods suggested the similarity of this new clonal group to ECI; hence, it was termed Epidemic Clone Ia (ECIa) (22). Later, MVLST analysis by Chen et al. revealed that the Boston milk outbreak strain differs sufficiently from those involved in the other two strains to be excluded from this clonal group (21). Also, their MVLST results indicated that the Boston vegetable and UK pâté outbreak strains are more closely related to ECII rather than to ECI; hence, they re-named the group of these outbreak strains Epidemic Clone IV (ECIV) (21). However, this

new definition has not been widely accepted and ECIa has still been more commonly used in the scientific literatures. Hence, we will follow the ECIa designation that includes the strains of the aforementioned two Boston outbreaks and UK pate outbreak.

ECIa lacks the characteristics of ECI such as unique RFLP patterns, resistance to Sau3AI digestion and presence of ECI-specific genetic markers (22). However, characteristics of ECIa have not been identified yet at genetic and phenotypic levels (22) and no specific genetic markers for ECIa have been developed to our knowledge. However, ECIa strains revealed a specific hybridization profile for a series of ECI- and ECII-specific probes used by Cheng et al. (23). Currently, the whole genome sequencing is available forHPB2262, the ECIa strain involved in the febrile gastroenteritis outbreak in Italy, 1997, linked with corn (32) and the ECIa strain J1-220 responsible for an outbreak linked with vegetables (20). Analysis of these genomes will facilitate the elucidation of features unique to ECIa.

**Epidemic Clone V**

Epidemic Clone V (ECV) was implicated in a single outbreak, the Mexican-style soft cheese outbreak that occurred in North Carolina in 2000 (22). Strains involved in this outbreak was designated into an independent clonal group mainly due to a unique PFGE pattern, which was unprecedented in the clinical isolates deposited in the PFGE profile database (PulseNet) (22). Later, the North Carolina outbreak strain, WS1, was found to be characteristically susceptible to Sau3AI but resistant to MboI (43). The status of a separate clonal group was reinforced by the distinct combination of the genetic traits each shared by another EC. Among the ECI-specific markers shown above, this clonal group harbors probe

133(43), which is conserved among ECI strains and also possessed in some non-ECI strains (43, 46, 67). Also, Region 18 seems to be diversified in ECV into the form homologous to ECII (22), thus being positive for ECII-specific genetic markers derived from this genomic region, whereas generating the same signal as non-ECII strains for other genetic markers that distinguish ECII (23). With ECIa strains, ECV shares the ribotype of DUP-1042B (21, 22). Although only one outbreak has been identified that was caused by ECV, putative members of this clonal group were detected among environmental samples from a turkey processing plant (43) and sporadic case isolates (21, 23, 32).

**Increased exposure or higher pathogenicity?**

The reason for the recurrent association of ECs with outbreaks has been speculated to result from enhanced virulence or ubiquitous presence in foods and processing environments (22). To investigate these two hypotheses, several studies have been conducted to examine the prevalence of ECs, mainly focusing on the frequency of ECI, Ia and II among serotype 4b isolates (43, 46, 143, 147). In a survey of food and environmental isolates of (putative or confirmed) serotype 4b *L. monocytogenes*, Yildirim et al. discovered that 20 out of 34 isolates (20/34, 59%) belonged to ECI (147). In another study, 18 serotype 4b isolates were examined that were obtained from the environmental and raw product samples of two turkey processing facilities in the US (43). Among these isolates, four (4/18, 22%) and ten (10/18, 56%) were positive for the ECI- and ECII-specific markers, respectively (43). Also, Franciosa et al. surveyed clinical, food and environmental serotype 4b isolates, which mostly originate from Italy, with the PCR targeting ECI- and ECII-specific markers (46). Out of a

total of 89 isolates, 16 (16/89, 18%) and 15 (15/89, 16.8%) belonged to ECI and ECII, respectively, both exhibiting a similar prevalence of 14-20.6 % in all three categories of sources (46). In the MLGT analysis of the 501 isolates (including 34 serotype 4b isolates) from RTE foods and processing facilities, Ward et al. discovered that a large proportion of serotype 4b belonged to ECI (9/34, 26.5%), Ia (14/34, 41.2%) and II (9/34, 26.5%) (143). All these findings suggest that ECI, Ia and II might be more advantageous than non-EC serotype 4b strains in inhabiting various sources including food processing facilities and foods. General abundance of ECs could increase human exposures and thus the possibility to cause an outbreak in comparison with non-EC serotype 4b (143).

Indeed, previous studies have revealed a few unique phenotypes of ECs that might promote the survival in the food processing facilities (75). For instance, Kim and Kathariou reported that ECII strains are characteristically resistant to *Listeria* phages at low temperatures (75). Also, ECII strains displayed resistance to a disinfectant used in the processing facilities (benzalkonium chloride, BC) due to the BC resistance cassette on the plasmid pLM80 (44, 96).

However, increased pathogenicity cannot be excluded from consideration in explaining the frequent association of ECs with outbreaks. For instance, den Bakker et al. applied the MLST targeting four housekeeping genes, two virulence genes, and one stress response gene to 180 lineage I and II isolates (including other serotypes as well as serotype 4b) from various sources in the New York state (32). Although ECI and ECII were found in similar frequencies between clinical and non-clinical isolates, ECIa isolates were overrepresented among clinical isolates in comparison with non-clinical isolates (32). Based

on these findings, they proposed that ECIa might represent a cluster of more virulent isolates than other lineage I and II *L. monocytogenes*(32).However, it should be mentioned that this clinical vs. non-clinical comparison was not limited to serotype 4b *L. monocytogenes* isolates (32); thus, the over-representation among clinical isolates should be regarded as the results of both unique attributes of ECIa and common features of serotype 4b that could be responsible for the over-representation of serotype 4b isolates in general. Also, strains tested positive for LIPI-3 included several strains belonging to ECI, Ia and II (27). Although not exclusively harbored by the isolates of these epidemic clonal groups, presence of this additional pathogenicity island might increment virulence potential of ECs.

## ***Listeria*** **phages**

Bacteriophages specific for *Listeria* species have been employed as typing and cloning tools (87) and more recently for biocontrol against *L. monocytogenes* (14, 87). Currently, no less than 400 *Listeria* phages have been isolated from various sources (87). All the *Listeria* phages described to date possess double-stranded DNA genomes of various sizes whose GC contests range from 34.7 to 40.8% (87). With regard to the taxonomy, a majority of *Listeria* phages are classified in the *Siphoviridae* family within the order *Caudovirales* that is characterized by isometric capsids and long, non-contractile tails (87). Other phages belong to the *Myoviridae* family in the same order that bears the isometric capsids and long, contractile tails (87). On the basis of the life cycle, *Listeria* phages are divided into the lytic phages whose host range can cover all *Listeria* serotypes and species; and the temperate

phages that are specific to serotypes and often equipped with generalized transduction (87). The lysogeny state is prevalent among *Listeria* species (35, 87, 88), conferring the resistance to the infection by the phages of the similar immunity groups (87, 88). Although prophages can be induced with UV light or mitomycin in *Listeria* species, many of the prophages seem to be cryptic (87). Depending on the *Listeria* species, infective phages are estimated to be produced only by 6 to 37% of the strains (87). Also, five phage-like elements were identified in the genome of *L. innocua* CLIP11262; nonetheless, only one was induced into infective virions (88). Cryptic prophages in *Listeria* species result in the production of inhibitory substances only against *Listeria* strains (monocins) (87). The lethality by monocins derives from the intact lysis genes encoded in the cryptic prophages (87).

Some representative *Listeria* phages were subjected to genome sequencing and thorough molecular biological investigations (14, 35, 77, 88, 148). In the next section, we will describe these *Listeria* phages whose genome sequences are available after dividing them according to life cycle.

## Temperate *Listeria* phages

### *A118*

A118 was obtained by UV light-mediated induction from a serotype 1/2a *L. monocytogenes* isolate from a Camembert soft cheese (87, 88). This phage is classified into the *Siphoviridae* family, infects only serotype ½ *L. monocytogenes* and possess the ability to conduct general transduction (87, 88). A118 is the first *Listeria* phage whose genome was

completely sequenced (87). Outside the host, the genome of A118 consists of a linear double-stranded DNA that is circularly permuted and terminally redundant (87, 88), but A118 forms a circular DNA of the identical unit length through recombination within cells (88). On average, the linear form of the genome amounts to 43.3 kbp in size (87, 88) and to 36.1% in GC content (88). In this genome, permutation occurs in 22.0 kb (88) and terminal redundancy was found in 2.5 kb, thus accounting for 6% of the unit genome of 40.8 kb (87, 88). Bioinformatics analysis identified 72 open reading frames (ORFs) in the A118 genome whose organization consists of three gene clusters (87, 88) oriented in the opposite direction from the neighboring sections (88). These gene clusters represent functional modules necessary for the life cycle of temperate phages: (1) the left-arm cluster consisting of the late genes involved in synthesis of structural proteins, packaging, and lysis; (2) the middle cluster contributing to establishing and maintaining lysogeny; and (3) the right-arm cluster composed of the early genes responsible for replicating, modifying and recombining the phage DNA before lysis (88). This organization was recurrently observed in other temperate *Listeria* phages (35, 148).

As a prophage, A118 integrates into the chromosome at the *attB* site within *comK* homolog (87, 88), which is often inhabited by a prophage in common lab strains of *L. monocytogenes* such as EGDe (87, 88). This *attB* site exhibits 3 bp of homology (GGA) with the A118 *attP* site located downstream to the integrase gene and its occupation by A118 seems to inactivate the ComK protein (88). The exceptionally short homology agrees with the serine integrase encoded in A118, which is capable of recognizing a short DNA

sequence, in contrast to those of other *Listeria* phages belonging to the tyrosine recombinases (35).

***A006***

A006 is another temperate *Siphoviridae* phage that infects serotype ½ (35). This phage possesses the terminally redundant, circularly permuted genome of 38.1 kb whose GC content amounts to 35.5% (35). In this genome, 62 putative ORFs were identified (35). Inside the host, A006 integrates into the 3' end of $tRNA^{Arg}$ that harbors the 16 bp homologous sequence (35). After the integration of A006, the $tRNA^{Arg}$ gene seems to maintain its functionality, which is the case for attachment sites of *Listeria* phages within other tRNA genes (35, 148).

***PSA***

PSA (phage of Scott A) is a *Siphoviridae* prophage that was induced by UV radiation from serotype 4b *L. monocytogenes* ScottA (148) and displays host specificity to serotype 4 *L. monocytogenes* strains (87, 148). PSA bears a linear DNA genome which is 37.6 kbp in length and 34.8% of GC content (148). Both termini of this linear DNA consist of 3'-protruding, sticky ends of 10 nucleotides (87, 148), which are ligated intracellularly, thus converting the linear PSA genome into a circular form (87). This constant genome structure complies with the inability of PSA to mediate general transduction (148). On the PSA genome, a total of 57 ORFs were discovered *in silico* (148). Proteomics research revealed that translational +1 frameshift occurs at the 3' end of the transcripts for *cps* (major capsid

protein) and *tsh* (major tail protein), each resulting in two protein species that possess the identical N-terminus but differ in length and C-terminus (148). PSA is maintained in the cells as a prophage by integrating into the *attB* site positioned at the 3'-end of single-copy $tRNA^{Arg}$ gene by means of 17 nucleotide-homology with *attP* (87).

***A500***

A500 is classified into the *Siphoviridae* family and infects serotype 4 (35). This phage features a circularly permuted, terminally redundant linear DNA genome that is 38.9 kb in length and 36.7% in GC content (35). A total of 64 ORFs was predicted through the bioinformatics search (35). In order to integrate into the chromosome, A500 relies on 45 bp homology region at the 3' end of $tRNA^{Lys}$ gene (35). Similar to PSA, *cps* and *tsh* genes of A500 were each translated into two protein species of different sizes due to ribosomal frameshifts at the 3' end, -1 for *cps* and +1 for *tsh* (35).

***B025***

The *Siphoviridae* phage B025 infects *L. innocua* whose serotypes 5 and 6 (35). B025 owns a linear DNA genome of 42.7 kb and 35.1% GC content that contains 3'-protruding single-stranded cohesive ends of 10 nucleotides (35). *In silico* analysis predicted 65 ORFs in this genome (35). Interestingly, B025 integrates into the same site as PSA within $tRNA^{Arg}$ in spite of different host specificities (35).

***B054***

B054 is the only *Myoviridae* temperate *Listeria* phage whose genome has been sequenced (35). B054 infects serotypes 5 and 6, and harbors a linear DNA genome which is characterized by terminal redundancy and circular permutation (35). On average, this genome amounts to 42.7 kb in size and 36.2% in GC content and features 80 putative ORFs (35). B054 is integrated at the *attP* site that was located in the 3' end of *tsf* gene coding for the translation elongation factor EF-Ts with the homologous sequence of 16 bp, not disrupting the function of the *tsf* gene (35).

***Relationships between temperate* Listeria *phages***

Genome-wide comparison revealed the close relationship among A118, A006, and A500 (35). A118 is reported to resemble at the genomic level the cryptic prophage ΦEGDe harbored by *L. monocytogenes* (87). When the deduced protein sequences are compared, A118 and A006 shared a high similarity in the early gene cluster whereas A118 and A500 exhibited homology in the structural genes (35). This relatedness in the defined functional region agrees with the model in which the phage genome is essentially a 'patchwork' of various genetic modules (87, 88).

Another group was identified through the genome sequence comparison that consists of PSA and B025 (35). This relationship was also supported on the level of the amino acid sequence of the predicted proteins since the PSA contains the homologs to the approximately half of the gene products of the B025 genome although the B025 genome features a considerable mosaic character (35). This close relationship is noteworthy since PSA and

B025 differ completely in their host ranges (35). This observation suggests that ancestors of *Listeria* phages might have experienced simultaneous evolution and adaption in parallel with their hosts (35).

On the other hand, B054 belongs to neither of the above-mentioned groups; however, this phage revealed similarity both at the genome and predicted protein sequences to the *L. innocua* Clip11262 prophage 7/12, the only prophage that was induced upon UV radiation (35). In addition, B054 was distantly located from other *Myoviridae Listeria* phages described to date, i.e., the SPO1-like phages of A511 and P100 (35).

## Lytic *Listeria* phages

### *A511*

A511 is a *Myoviridae* lytic phage whose host range covers *L. ivanovii* and serotypes ½ and 4 *L. monocytogenes* (77). The A511 genome of 137.6 kbp contains terminal repeated sequence of 3.1 kb that harbors nine putative ORFs (77). Excluding this terminal repeat, the informational genome of 134.5 kbp amounts to 36.1% in GC content and encodes 190 putative ORFs and a cluster of 16 tRNA genes (77). These genes are organized into three regions transcribed in the opposite direction from each other and into two functional clusters involved in synthesis of structural proteins and DNA replication, respectively (77).

***P100***

The *Myoviridae* phage P100 was isolated in 1997 from a sewage sample obtained in a German dairy facility (14). This phage exhibits a broad host range including serotype 5 of *L. ivanovii* and serotype ½ and 4 *L. monocytogenes*, which is similar but not identical to that of A511 (14). The genome of P100 is composed of a linear DNA of 137 kb with terminal redundancy of 6 kb (77). *In silico* analysis revealed that the informational genome contains a total of 174 putative ORFs and 18 clustered tRNA genes (14).

***P35***

P35 is a lytic *Siphoviridae* phage that infects serotype ½ *Listeria* strains (35). This phage features the terminally redundant, circularly permuted genome of 35.8 kb in size and 40.8% in GC content where 56 putative ORFs reside (35). Genome sequencing revealed that the lysogeny gene cluster is completely missing in P35, confirming the obligatory lytic lifestyle of this phage (35).

***P40***

P40 is another lytic *Siphoviridae* phage infecting *Listeria* species and exhibits a wide host range including serotypes ½, 4, 5, and 6 (35). The P40 genome amounts to 35.6 kb in length and 39.3% in GC content; displays terminal redundancy and circular permutation; and bears 62 putative ORFs (35). Similar to P35, the lysogeny-related genes are absent in P40, indicating that this phage is strictly lytic (35).

**Relationship between lytic *Listeria* phages**

The *Myoviridae* phages A511 and P100 are closely related to each other (14, 77) and both classified into SPO1-like phages of the *Myoviridae* family, a group of lytic phages that infect low GC-content, G+ bacteria, possess similar, large genomes, and resemble in morphology (77). Comparison between A511 and P100 genomes revealed that genes encoding structural components exhibited homology between A511 and P100; however, each had unique genes involved in DNA replication and unknown functions, reflecting an adaptation specific to host (77). Meanwhile, on the basis of genome sequence, the *Siphoviridae* phages P35 and P40 form a group distantly located from A511 and from other *Siphoviridae* temperate phages (35). In line with this finding, half of the gene products of P35 and P40 did not match any database entries (35).

**Influence of *Listeria* phages on the *L. monocytogenes* ecology and evolution in the food processing environment**

Ecological and evolutionary importance of *Listeria* phages has been evidenced by several recent studies (19, 76, 105, 139). Kim et al. determined the host range of *Listeria* phages isolated from the turkey processing facilities in the US (76). Consequently, they obtained two wide-host-range *Listeria* phages from separate plants that are capable of infecting the same hosts and whose partial genome sequences share similarity with P100 and A511, indicating the presence of a closely related group of *Listeria* phages that frequently inhabit food processing facilities (76). Among different serotypes of the environmental *L. monocytogenes* samples, these phages failed to infect serotypes 1/2a (or 3a) and 1/2b (or 3b)

in a higher frequency than 4b complex (76). This resistance differential was speculated to be associated with the predominance of serotypes 1/2a and 1/2b in environmental and food isolates of *L. monocytogenes* (76, 130).

In another study on the *Listeria* phages as an ecological determinant, Orsi et al. compared at the genomic level the ECIII strains that originated from the same processing plant but were involved in two different clinical cases that occurred twelve years apart (105). Their analysis revealed that the genetic differences between these isolates reside mainly in two prophage sequences incorporated into *comK* (extensive diversification) and tRNA-Thr-4 (one single nucleotide polymorphism (SNP)), suggesting that phages might actively contribute to the evolution of *L. monocytogenes* in this environment (105). The genetic changes in the *comK* prophages among closely related *L. monocytogenes* strains were also supported by two studies that discriminated different outbreak strains belonging to the same epidemic clone based on the DNA sequences from the internal part or junction fragments of *comK* prophages (19, 139). Verghese et al. discovered spontaneous induction and release of defective *comK* phages, whose transduction might be responsible for diversification of the *comK* prophages (139). Another intriguing finding on the *comK* prophages is that strains without *comK* prophages (ECI and lineage III) exhibited less growth on meat- and poultry-conditioning films whereas strains harboring *comK* prophages (ECII, III and 2008 Canadian outbreak strain [termed ECV in the paper by Verghese et al. (139)]) grew better under the same condition and serotype 1/2a ECIII and 2008 Canadian outbreak strain formed biofilm (139). This observation in addition to variety in *comK* prophages prompted Verghese et al. to propose that the *comK* prophages might encode genes ("adaptons") that facilitate *L.*

*monocytogenes* to adapt to and persist in individual food processing facilities by promoting attachment, growth and biofilm formation on different food conditioning films (139). In this context, they named the defective *comK* phages "rapid adaptation islands (RAI)" (139). As reflected in these studies, only recently was appreciated the role of phages in the evolution and ecology of *L. monocytogenes* and more studies are warranted to enhance our understanding on this subject. Such research will be particularly relevant now that the influence of *Listeria* phages is expected to increase due to *Listeria* phages that have been applied against food contamination by *L. monocytogenes*(91).

## Restriction-modification systems

Restriction-modification (RM) systems refer to the gene cassettes that selectively degrade the extracellular DNA, which are marked with the absence of a chemical identification signature, usually methylation, at a particular sequence (132, 133). Generally, the RM systems consist of the methyltransferase that attaches a methyl group to adenine or cytosine bases within the recognition site; and the restriction endonuclease that cleaves the unmethylated DNA (133). The methyltransferase of the RM system methylates the host DNA after replication, thus protecting the intracellular DNA from degradation; meanwhile, the phages and plasmids are missing the methylation and targeted by the restriction endonuclease of the RM system (68, 133).

RM systems are characterized with rapid evolution (68) and dissemination through the horizontal gene transfer (68, 71, 79, 80). Also, RM systems are often harbored by mobile

genetic elements such as plasmids (79, 80, 132) and phages (71, 80), transposons and prophages (79, 80). Interestingly, RM systems vary in number from one microorganism to another (80). Some species of bacteria are marked with the abundance of the RM systems (71). Such species include *Neisseria gonorrhoeae*, *Helicobacter pylori*, *Methanococcus jannaschii* (71, 79, 80), *Haemophilus influenzae*, *Neisseria meningitidis* and *Xylella fastidiosa* (79, 80). However, the near or complete absence of RM systems was observed in some prokaryotic species such as *Rickettsia prowazekii*, *Treponema pallidum*, *Chlamydia* and *Buchnera* (80).

RM systems have been divided into four groups, Types I-IV (68, 133), that differ from each other in terms of the activities and co-factor requirements of the enzymes; the nature of the target sequence; and the relative location of the restriction site against the recognition site (97, 133). In the next section, an individual division of RM systems will be discussed in detail.

**Type I**

Type I RM systems consist of three subunits, namely, Hsd (host specificity of DNA) R, HsdM, and HsdS, that are required for restriction, modification, and DNA specificity determination, respectively (97, 133). These subunits are encoded by adjacent genes, *hsdR*, *hsdM*, and *hsdS* (97). The *hsdR* is transcribed from an independent promoter whereas *hsdM* and *hsdS* are transcribed from the common promoter (97). Typically, two HsdR, two HsdM, and one HsdS subunits form a hetero-oligomer that methylates hemi-methylated DNA and

restricts unmethylated DNA (97, 133). Another form of oligomers is comprised of two HsdM and one HsdS subunits and only capable of methylation (97).

Type I systems recognize asymmetric, bipartite DNA sequences that consist of two components of 3-4 bp and 4-5 bp between the unspecific spacer of 6-8 bp (97). Methylation requires S-adenosyl methionine (SAM) as a methyl donor (97) and results in the $N^6$-methylated adenine (133). Restriction is preceded by the DNA translocation that utilizes ATP as the energy source while two enzyme complexes are bound to their target sequences (133). The stalling of this process culminates in the cleavage in a location ~1000 bp distant from the recognition site (97, 133). Apart from ATP, restriction requires SAM and $Mg^{2+}$ (97, 133).

The recognition site of the Type I system is dictated by HsdS for both methylation and restriction (97). The HsdS subunit contains four domains: two conserved domains, i.e., a conserved central domain and a conserved C-terminus; and two variable, target recognition domains (TRDs), i.e., the amino-proximal recognition domain (ARD) and the carboxy-proximal recognition domain (CRD) (95, 97). Each TRD is responsible for recognizing one component of the recognition site; hence, changes in even one TRD sequence are expected to confer a different specificity to a Type I system (97). New specificities are acquired mainly via two mechanisms (97). One mechanism of specificity diversification depends on a combination of TRDs through domain shuffling (97) as exemplified in *Lactococcus lactis* wherein the recombination between the conserved central domains of an *hsdS* within a Type I RM cassette and another solitary *hsdS* created two novel *hsdS* hybrid genes, which displayed different specificities to those of their parental genes (107). Meanwhile, alterations in the

spacing between the TRDs also enhance specificity diversity by changing the length of the spacer connecting two components of the recognition site (97).

At this point, given the nature of the RM system, one might anticipate that, upon the generation of the new specificity, the host genomic DNA might not be protected from its resident RM system. Type I RM systems circumvent this potential disaster through the restriction alleviation (RA), the ClpXP-dependent regulation of the restriction activity, which would provide time for the host genomic DNA to be methylated at a new recognition site (68).

To date approximately 600 putative Type I RM enzymes have been identified in prokaryotes (133). Type I RM systems are distributed in an inconsistent manner with the bacterial phylogenetic relationships (6, 115), suggesting that they might have been acquired through horizontal gene transfer (6). Depending on the amino acid and DNA similarities, subunit complementation, and immunological cross-reaction (97), Type I RM systems are allocated into five families, Types IA-IE (97, 133). Recently, two additional families, Types IAB and IF, have been identified through the comparison of the Type I RM systems of 73 *C. jejuni* strains (95). Members within the same family resemble each other except for TRDs of HsdS subunits whereas those belonging to different families exhibit low similarities in all subunits (97).

**Type II**

Type II RM systems are comprised of the restriction endonuclease and the methyltransferase, both of which operate separately (97, 133). These enzymes recognize the

identical nucleotide sequence, in most cases, a palindrome of 4-8 bp and the DNA cleavage occurs inside or adjacent to the target sequence , generating DNA products with the 3'-OH and 5'-phosphate (133). To date, approximately more than 3650 Type II RM systems have been identified (133).

Type II methyltransferases are involved in the methylation of a cytosine at either the N4 or C5 position or of an adenine at the N6 position (133) on both strands of the recognition site (97). For this methylation, SAM is required as a methyl donor (97, 133). Type II methyltransferases function as monomers (97, 133) and bear conserved motifs (83, 133).

Typical Type II restriction endonucleases form symmetrical homo-dimers, wherein each subunit is responsible for the cleavage on one strand of the double-stranded DNA (133). A majority of Type II restriction endonucleases that were subjected to crystallography possess PD…(D/E)XK motif in the catalytic centers, one in an individual subunit (103, 110, 111). This motif is important in the co-factor acquisition since two carboxylates in this motif bind to $Mg^{2+}$ (111), the co-factor required for Type II restriction endonucleases (97, 110, 133). Depending on structural similarities, these Type II restriction endonucleases harboring PD…(D/E)XK are grouped into the EcoRI and EcoRV branches (110, 111). Those within the EcoRI branch access the DNA from the major groove and cut DNA into fragments obtruding at 5' end (110, 111). Meanwhile, the enzymes belonging to the EcoRV branch advance to the DNA from the minor groove and generate 3'-sticky or blunt ends (110).

In spite of these general features, Type II restriction endonucleases exhibit considerable diversity compared to the methyltransferases in terms of behaviors or amino acid sequences; thus, these enzymes were further divided into 11 overlapping classes (133).

For instance, some function as monomers or tetramers (110, 133). Also, some require the binding to multiple recognition sites prior to the restriction (133). Even the PD…(D/E)XK motif was found only in 69% of experimentally verified Type II restriction endonucleases and this number decreased to 48% when putative enzymes were included (103). Other protein folds without this motif were identified in Type II restriction endonucleases, for instance, phospholipase D, half-pipe, GLY-YIG, and HNH folds (103).

Some Type II RM systems contain the repressor-like third component, C (control) protein (97, 133), which shares similarities to helix-turn-helix (HTH) proteins (78, 98). The C protein gene usually resides upstream of the restriction endonuclease gene and both genes are transcribed from the same promoter (98). The C protein regulates the expression of the restriction endonuclease by binding as a homo-dimer to a C box, the DNA sequence up to 30 bp located upstream of this common promoter of the C protein gene and restriction endonuclease gene (98). The C box consists of two pairs of inverted repeats and the C protein-mediated regulation is determined by which inverted repeats are bound by the C protein (98). When the C protein exists at a low concentration, this protein occupies only the inverted repeat further from the transcription start due to its high affinity (98). This interaction stimulates the expression from the common promoter of the C protein and restriction endonuclease genes; thus, the concentration of both proteins increases (98). The accumulation of the C protein causes both repeats to be occupied by the C proteins and this binding negatively influences the activity of the promoter dictating the transcription of the C protein and restriction endonuclease (98).

**Type III**

Type III RM systems resemble Type I systems in that they form hetero-oligomers that are capable of both modification and restriction (97, 133). However, unlike Type I RM systems, the hetero-oligomers of Type III systems are comprised of only two components, Mod and Res subunits (97, 125, 133). Mod recognizes the 5-6 bp asymmetric sequence and attaches a methyl group to the N6 position of adenine on one strand, thus producing the hemi-methylated DNA (125, 133). Mod can function without the involvement with its corresponding Res (125, 133). On the other hand, Res cleaves the DNA sequence only as a part of the hetero-oligomer consisting of two Mod subunits and two Res subunits (133). For Type III RM systems, SAM is required as a co-factor of modification whereas $Mg^{2+}$ and ATP are essential for restriction (97, 133).

DNA restriction by Type III RM systems is initiated by the interaction between two enzyme complexes and two inversely located copies of the recognition sites (125, 133). DNA is translocated while these hetero-oligomers remains bound to their recognition sequences (125, 133). Delayed DNA translocation and collision of two complexes culminate in the cleavage, which occurs 25-27 bp apart from one of the two recognition sites (125, 133). During the cleavage process, each complex cuts one DNA strand (133).

More than 140 Type III RM systems have been identified, in most cases, among phages and Gram-negative bacteria (125, 133). Genes encoding Mod and Res are usually located adjacent to each other (68, 133) and display a considerable similarity in DNA sequence between different systems (133). Variability among Type III RM systems, in

particular in sequence specificity, seems to be achieved through restriction alleviation as demonstrated by RecE encoded by the Rac prophage (68).

**Type IV**

Type IV RM systems are restriction endonucleases that selectively cut modified DNAs, for instance, those methylated, hydroxymethyled and glucosyl-hydroxymethylated (118, 133). To date, a total of 227 putative Type IV RM system enzymes have been identified, among which *Escherichia coli* K-12 Mcr (Modified cytosine restriction) BC system was most extensively studied (115, 133). McrBC requires $Mg^{2+}$ as a co-factor (133) and recognizes two copies of a dinucleotide RmC, a purine followed by a $N^4$- or $C^5$-methylated cytosine, between which the distance ranges from 40-3000 bp (118, 133). The substrate of McrBC includes hemi-methylated DNA and the location of the hemi-methylation did not affect the restriction (115). The cleavage occurs at 30 bp apart from one of the recognition sites (118, 133). Reminiscent to the Type I RM systems, McrBC translocates DNA while adhering to the restriction sites and the arrest of translocation leads to cleavage (133). However, unlike Type I systems, GTP provides energy in this process (133) whereas the ATP inhibits cleavage (115).

McrBC system consists of two components, McrB and McrC (33). The McrB contains a GTP-binding motif in the carboxyl terminal part (33, 109), which is conserved in its homologs (109). In addition, McrB performs an indispensable role in DNA binding (82) and cleavage (115). McrC is not involved in the DNA binding process (82) but is required for cutting some substrates of the McrBC system (33, 115). The amino acid mutation experiment

revealed that the PD....(D/E)XK motif in McrC forms the catalytic center for DNA cleavage (108). Also, this protein increases the GTP hydrolysis activity by McrB (109).

The GC content of the *mcrBC* genes was below the genome average, suggesting that this system might have been introduced through a recent horizontal gene transfer (33). This was supported by the phylogenetic analysis with 199 McrBC homologs identified via BLAST searches by Fukuda et al. (49). This analysis revealed that the McrBC homologs are dispersed in various bacteria and archaea, i.e., approximately 20% of 567 completely sequenced genomes analyzed (49). Except for a few cases, McrB and McrC were located together (49). This close linkage is consistent to the finding that McrB and McrC appear to have co-evolved through complex mechanisms consisting of frequent joint horizontal gene transfer and vertical inheritance and loss (49). Beside this association between McrB and McrC, Fukuda et al. discovered that the McrBC systems often accompany DNA methyltransferases, frequently of Type I RM systems (49). Meanwhile, individual McrBC homologs displayed considerable variance in sequence, which Fukuda et al. attributed to the diversity in target recognition (49).

**Immigration Control Region**

The Immigration Control Region (ICR) refers to a specific genetic location on the chromosome where multiple RM systems reside (68). This region is highly variable among strains of the same species (68, 97, 115); and genetically mosaic, suggesting that horizontal gene transfer might be involved in the gene acquisition (68). In *E. coli* K-12, the ICR encompasses the Type I EcoKI system flanked by two Type IV systems Mrr and McrBC (97,

115). These RM systems do not seem to be acquired simultaneously since hybridization pattern and unique GC content suggest that McrBC system might be acquired through a separate horizontal gene transfer from the EcoKI system (115). Similar ICRs were identified in other *E. coli* strains (97, 115) and *S.* Typhimurium (68, 97). Comparisons of various ICRs in enteric bacteria revealed that these regions are occupied by at least three families of Type I systems (IA, IB and ID) (97, 133). Similarly, Mrr and Mcr systems in ICRs were suspected to form multiple families (115).

**Biological roles of RM systems**

Traditionally, RM systems have been described as a major defense mechanism against the bacteriophage invasion (68, 71, 79, 80, 97, 133) that confers $10^2$-$10^4$-fold protection to microorganisms (71). Also, the RM system decreases horizontal gene transfer from one bacterium to another by reducing the efficiency of conjugation (4, 132) and transformation (3, 4, 132). This feature of the RM systems was exemplified by *H. pylori* strains into which the introduction of plasmid (3) and DNA fragments (4) was inhibited by Type II RM systems. This prohibition of the gene transfer was proposed to contribute to establishing genetic isolation, which is required for speciation (71).The genetic isolation caused by mutually exclusive RM systems was reported in the comparison of two *H. pylori* strains, where all strain-specific RM systems exhibited activity in contrast to those commonly found that were generally non-operational (85). Another example was provided from the comparison of three RM systems NmeAI, NmeBI and NmeDI in distribution among *N. meningitidis* strains, which revealed lineage-specific features of the RM systems (24). In

this analysis, NmeBI was possessed by the lineages linked with serogroup B disease; serogroup A strains contained NmeAI; and the co-residence of NmeAI and NmeDI was observed among the lineages associated with serogroup C disease (24).

At the same time, RM systems are considered to stimulate homologous recombination (71, 80, 133) by generating short DNA fragments through the incomplete restriction (71). For instance, RecBCD-mediated recombination is induced by the DNA fragments cut with a Type II endonuclease (80, 133). This phenomenon might facilitate acquiring and maintaining advantageous genetic information (133).

RM systems have also been classified as selfish genes by some researchers (71, 79, 80, 97). In this hypothesis, RM systems might not be considerably beneficial to the host microorganisms (71) but they force the host microorganism to maintain them by killing the cells that have lost the RM systems (68, 79, 80). Upon the elimination of an RM system from the host cell, the methyltransferase is diluted as the cell multiplication proceeds; hence, the genomic DNA of the progeny cells is not sufficiently methylated to be protected from the restriction endonuclease (79, 80). In this case, even a few restriction endonuclease molecules would be fatal since one breakage of the genomic DNA could result in the cell death (79, 80). Post-segregational lethality anticipated from this scenario was demonstrated for many Type II systems (79). The host death by the Type IV *E. coli* McrBC system was observed when a PvuII methyltransferase gene was introduced into the cells (49). Also, Mebrhatu et al. observed cell death when the Type IV RM system Mrr of *E. coli* MG1655 was expressed in *S.* Typhimurium LT2 and the random mutagenesis screening revealed that this phenomenon is due to the Type III RM system StyLTI of *S.* Typhimurium LT2 (131). Corresponding to

these results, the homologs of *mrr*$^{MG1655}$ and *mod*$^{LT2}$ tend to be distributed in the mutually exclusive manner (131).

Recently, Aertsen et al. reported for the first time the linkage between RM systems and bacterial stress response (1). They demonstrated that, in *E. coli* MG1655, the Type IV RM system Mrr restriction activity was induced by high hydrostatic pressure without presence of the properly methylated DNA, causing the SOS response (1). However, the induction by high pressure was not observed for the Mrr of *S.* Typhimurium LT2 (2), suggesting that not all Mrr proteins are involved in the stress response for high pressure. Roles of other RM systems in bacterial stress response have not been reported yet; however, further studies on this subject will facilitate our understanding on novel functions of RM systems in bacterial physiology.

**Defense against restriction**

Several mechanisms to date have been elucidated to counteract the restriction by RM systems in phages, plasmids and bacterial genomes (71, 79, 80, 133). One of these mechanisms replies on the modification of the DNA (71, 79, 133). This mechanism could be mediated through the action of other RM systems or solitary methyltransferases with the same recognition site (71, 79, 80, 87, 133, 148). Alternatively, some phages stimulate the methyltransferases encoded in the host genome or temporarily escape the restriction attack with the accessory proteins binding to their genomes in order to obtain protective modification (133).

Another mechanism relates the reduction of eligible restriction sites (71). For instance, some phage genomes contain unusual bases, which render these genomes out of the available substrates of restriction endonucleases (133). Selective elimination of the restriction sites not only facilitates escape from restriction but also bestows additional advantages of expediting full methylation and reducing epigenetic gene expression problems (80, 119). This phenomenon called restriction avoidance (79, 119) is often observed in phages (119, 133), plasmids (133) and bacterial genomes (79, 80, 119). Compared with the phages, bacterial genomes exhibited this tendency more extensively, suggesting that bacteria are asserted under a stronger selective pressure than phages to eradicate restriction sites (119). Among the host chromosomes, restriction avoidance was pronounced among the targets of the resident (80, 119).

The third strategy against restriction concerns disrupting the restriction endonuclease from being fully functioning (133). Phage T3 harbors a gene encoding SAM hydrolase, which degrades SAM, the essential co-factor of Types I and III RM systems (133). Another example is provided by the phages (133) and plasmids (132, 133) that are equipped with genes whose products inhibit Type I RM systems. In the phage T7, the Ocr (overcome classical restriction) protein occupies the DNA binding site of the Type I enzyme complex by imitating the properties of the target DNA (133). This interaction is stronger in affinity than that between the Type I complex and DNA to such a degree that this binding disables both restriction and modification functions of the Type I RM system (133). On the other hand, plasmid-encoded ArdA (alleviation of restriction of DNA A) and ArdB block restriction endonucleases of Type I RM systems for a limited time (133). Although transient, the early

expression of these proteins contributes to the plasmids avoiding the restriction and obtaining methylation (133).

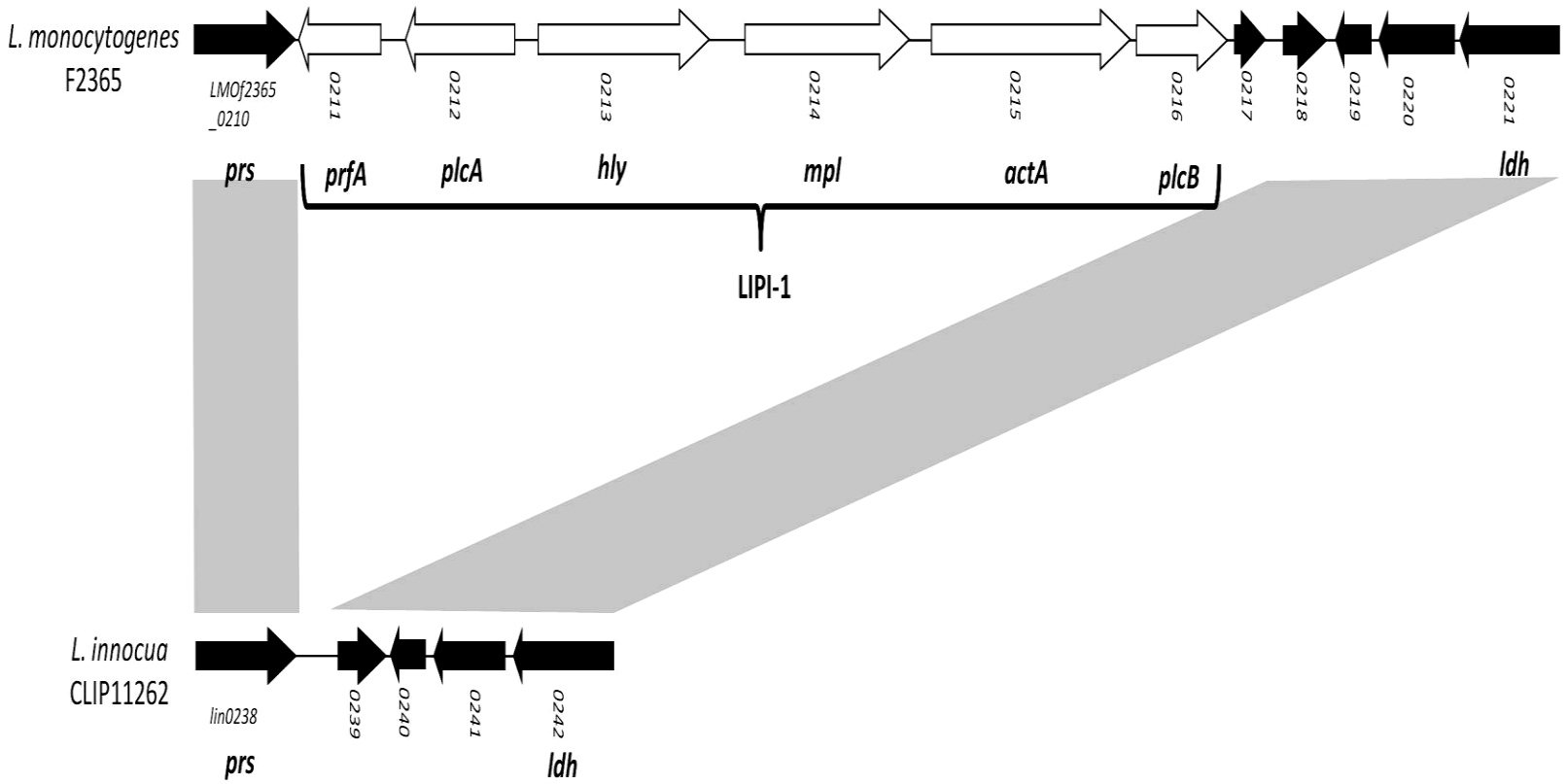



**Figure 1.1Comparison of the genetic region flanked by *prs* and *ldh* between *L. monocytogenes* F2365 and *L. innocua* CLIP11262.** Genes comprising the LIPI-1 are marked in white and other genes are colored in black. The grey shadow indicates homologous region identified with WebACT (http://www.webact.org/WebACT/home) (15).

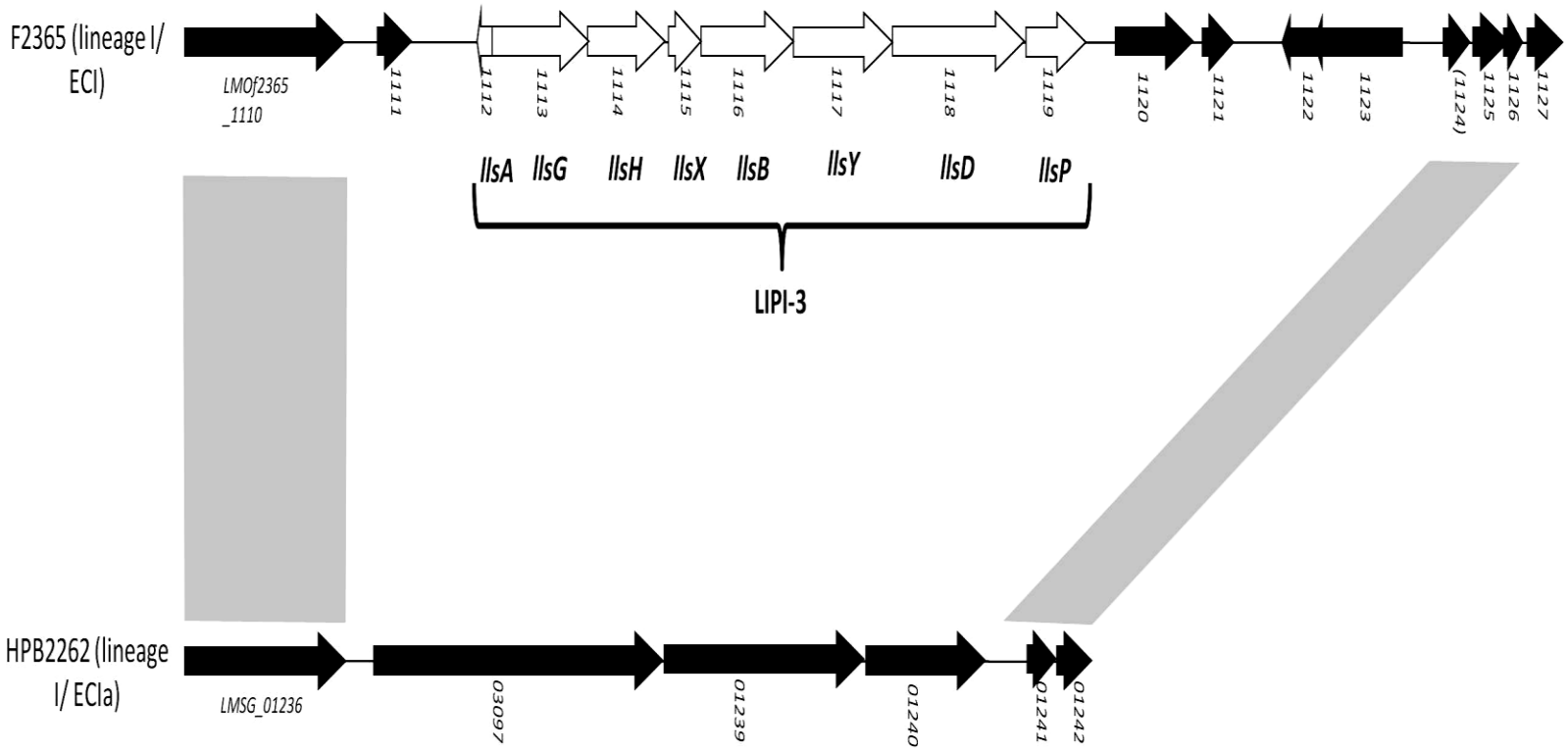



**Figure 1.2 Comparison of the genetic region harboring LIPI-3 in *L. monocytogenes* F2365 (lineage I/ ECI) and the genomically equivalent region in *L. monocytogenes* HPB2262 (lineage I/ ECIa), which is deficient of LIPI-3.** White genes represent LIPI-3 whereas all other genes are marked in black. The name of a pseudogene (*LMOf2365_1124* in F2365) is surrounded by parentheses. The grey shadow signifies homologous region found with WebACT.

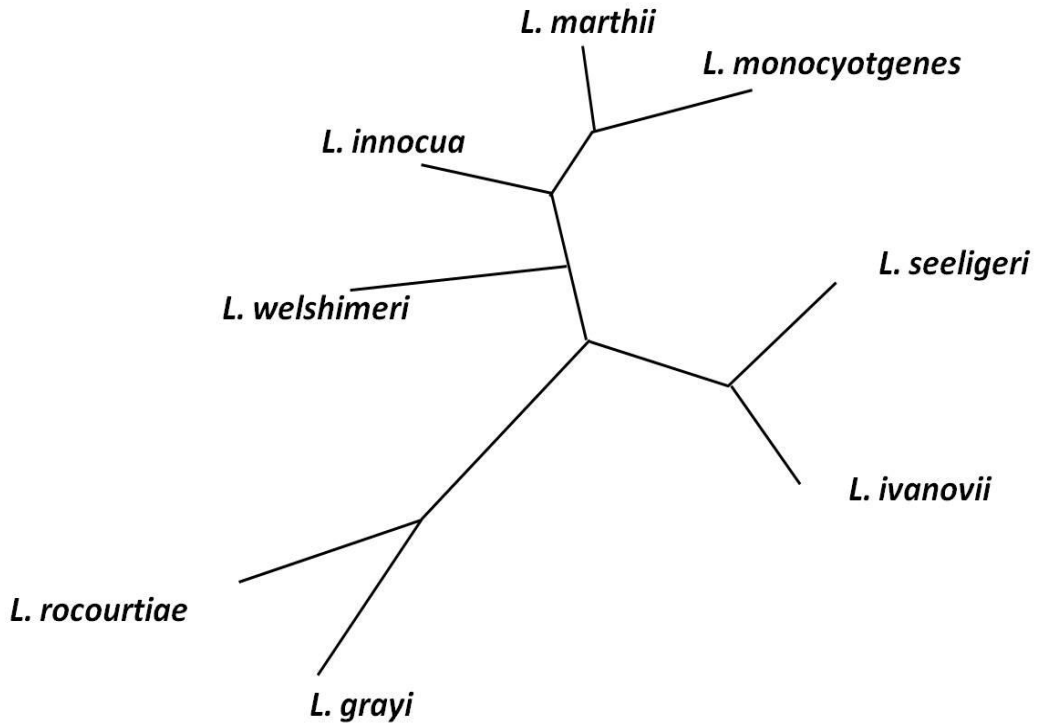



**Figure 1.3 Summary of the phylogenetic relationships of *Listeria* species revealed by previous studies.**

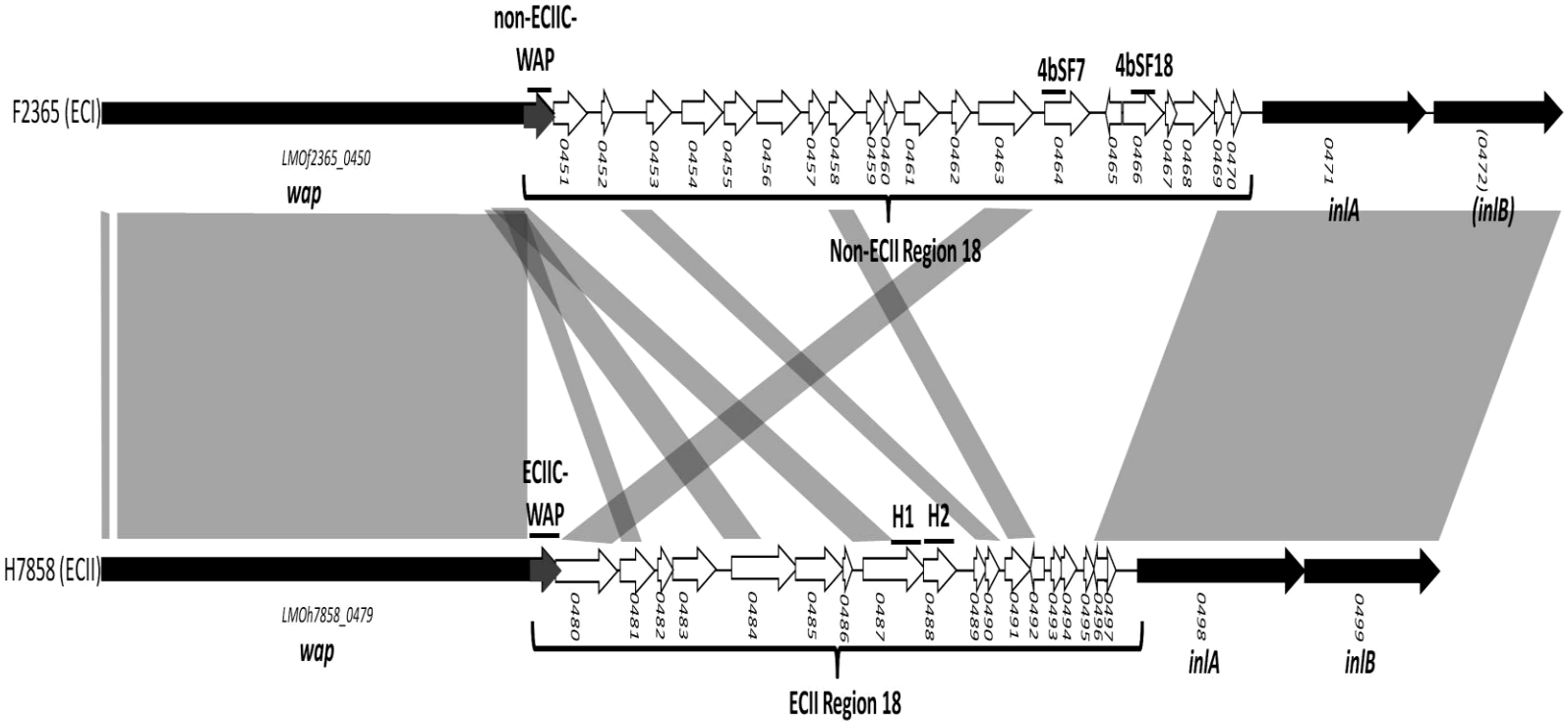



**Figure 1.4 Comparison of Region 18 between F2365 (ECI) and H7858 (ECII).** Black genes represent the flanking genes of Region 18 (*wap*, *inlA* and *inlB*) and white genes signify genes in Region 18. Part of *wap* divergent between F2365 and H7858 is colored in dark grey. The name of a pseudogene (*inlB* in F2365) is marked with parentheses. Homologous region determined with WebACT is represented by the grey shadow.

## REFERENCES


1. **Aertsen, A., and C. W. Michiels.** 2005. Mrr instigates the SOS response after high pressure stress in *Escherichia coli*. Mol. Microbiol. **58:**1381-1391.
2. **Aertsen, A., M. Tesfazgi Mebrhatu, and C. W. Michiels.** 2008. Activation of the *Salmonella* Typhimurium Mrr protein. Biochem. Biophys. Res. Commun. **367:**435-439.
3. **Ando, T., Q. Xu, M. Torres, K. Kusugami, D. A. Israel, and M. J. Blaser.** 2000. Restriction-modification system differences in *Helicobacter pylori* are a barrier to interstrain plasmid transfer. Mol. Microbiol. **37:**1052-1065.
4. **Aras, R. A., A. J. Small, T. Ando, and M. J. Blaser.** 2002. *Helicobacter pylori* interstrain restriction-modification diversity prevents genome subversion by chromosomal DNA from competing strains. Nucleic Acids Res. **30:**5391-5397.
5. **Aureli, P., G. C. Fiorucci, D. Caroli, G. Marchiaro, O. Novara, L. Leone, and S. Salmaso.** 2000. An outbreak of febrile gastroenteritis associated with corn contaminated by *Listeria monocytogenes*. N. Engl. J. Med. **342:**1236-1241.
6. **Barcus, V. A., A. J. Titheradge, and N. E. Murray.** 1995. The diversity of alleles at the *hsd* locus in natural populations of *Escherichia coli*. Genetics. **140:**1187-1197.
7. **Baumgartner, M., U. Karst, B. Gerstel, M. Loessner, J. Wehland, and L. Jansch.** 2007. Inactivation of Lgt allows systematic characterization of lipoproteins from *Listeria monocytogenes*. J. Bacteriol. **189:**313-324.
8. **Bennion, J. R., F. Sorvillo, M. E. Wise, S. Krishna, and L. Mascola.** 2008. Decreasing listeriosis mortality in the United States, 1990-2005. Clin. Infect. Dis. **47:**867-874.

9. **Bierne, H., and P. Cossart.** 2007. *Listeria monocytogenes* surface proteins: from genome predictions to function. Microbiol. Mol. Biol. Rev. **71:**377-397.

10. **Boerlin, P., J. Rocourt, and J. C. Piffaretti.** 1991. Taxonomy of the genus *Listeria* by using multilocus enzyme electrophoresis. Int. J. Syst. Bacteriol. **41:**59-64.

11. **Bonazzi, M., M. Lecuit, and P. Cossart.** 2009. *Listeria monocytogenes* internalin and E-cadherin: from structure to pathogenesis. Cell. Microbiol.

12. **Buchrieser, C.** 2007. Biodiversity of the species *Listeria monocytogenes* and the genus *Listeria*. Microbes Infect. **9:**1147-1155.

13. **Cabanes, D., P. Dehoux, O. Dussurget, L. Frangeul, and P. Cossart.** 2002. Surface proteins and the pathogenic potential of *Listeria monocytogenes*. Trends Microbiol. **10:**238-245.

14. **Carlton, R. M., W. H. Noordman, B. Biswas, E. D. de Meester, and M. J. Loessner.** 2005. Bacteriophage P100 for control of *Listeria monocytogenes* in foods: genome sequence, bioinformatic analyses, oral toxicity study, and application. Regul. Toxicol. Pharmacol. **43:**301-312.

15. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics. **21:**3422-3423.

16. **Centers for Disease Control and Prevention (CDC).** 2010. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food - 10 states, 2009. MMWR Morb. Mortal. Wkly. Rep. **59:**418-422.

17. **Chan, Y. C., and M. Wiedmann.** 2009. Physiology and genetics of *Listeria monocytogenes* survival and growth at cold temperatures. Crit. Rev. Food Sci. Nutr. **49:**237-253.

18. **Chen, Y., and S. J. Knabel.** 2007. Multiplex PCR for simultaneous detection of bacteria of the genus *Listeria*, *Listeria monocytogenes*, and major serotypes and epidemic clones of *L. monocytogenes*. Appl. Environ. Microbiol. **73:**6299-6304.

19. **Chen, Y., and S. J. Knabel.** 2008. Prophages in *Listeria monocytogenes* contain single-nucleotide polymorphisms that differentiate outbreak clones within epidemic clones. J. Clin. Microbiol. **46:**1478-1484.

20. **Chen, Y., E. A. Strain, M. Allard, and E. W. Brown.** 2011. Genome Sequences of *Listeria monocytogenes* Strains J1816 and J1-220, Associated with Human Outbreaks. J. Bacteriol. **193:**3424-3425.

21. **Chen, Y., W. Zhang, and S. J. Knabel.** 2007. Multi-virulence-locus sequence typing identifies single nucleotide polymorphisms which differentiate epidemic clones and outbreak strains of *Listeria monocytogenes*. J. Clin. Microbiol. **45:**835-846.

22. **Cheng, Y., R. M. Siletzky, and S. Kathariou.** 2008. Genomic Divisions/Lineages, Epidemic Clones, and Population Structure, p. 337-358. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

23. **Cheng, Y., J. W. Kim, S. Lee, R. M. Siletzky, and S. Kathariou.** 2010. DNA probes for unambiguous identification of *Listeria monocytogenes* epidemic clone II strains. Appl. Environ. Microbiol. **76:**3061-3068.

24. **Claus, H., A. Friedrich, M. Frosch, and U. Vogel.** 2000. Differential distribution of novel restriction-modification systems in clonal lineages of *Neisseria meningitidis*. J. Bacteriol. **182:**1296-1303.

25. **Clayton, E. M., C. Hill, P. D. Cotter, and R. P. Ross.** 2011. Real-time PCR assay to differentiate Listeriolysin S-positive and -negative strains of *Listeria monocytogenes*. Appl. Environ. Microbiol. **77:**163-171.

26. **Corr, S. C., and L. A. O'Neill.** 2009. *Listeria monocytogenes* infection in the face of innate immunity. Cell. Microbiol.

27. **Cotter, P. D., L. A. Draper, E. M. Lawton, K. M. Daly, D. S. Groeger, P. G. Casey, R. P. Ross, and C. Hill.** 2008. Listeriolysin S, a novel peptide haemolysin associated with a subset of lineage I *Listeria monocytogenes*. PLoS Pathog. **4:**e1000144.

28. **Dalton, C. B., C. C. Austin, J. Sobel, P. S. Hayes, W. F. Bibb, L. M. Graves, B. Swaminathan, M. E. Proctor, and P. M. Griffin.** 1997. An outbreak of gastroenteritis and fever due to *Listeria monocytogenes* in milk. N. Engl. J. Med. **336:**100-105.

29. **de las Heras, A., R. J. Cain, M. K. Bielecka, and J. A. Vazquez-Boland.** 2011. Regulation of *Listeria* virulence: PrfA master and commander. Curr. Opin. Microbiol. **14:**118-127.

30. **den Bakker, H. C., B. N. Bundrant, E. D. Fortes, R. H. Orsi, and M. Wiedmann.** 2010. A population genetics-based and phylogenetic approach to understanding the evolution of virulence in the genus *Listeria*. Appl. Environ. Microbiol. **76:**6085-6100.

31. **den Bakker, H. C., C. A. Cummings, V. Ferreira, P. Vatta, R. H. Orsi, L. Degoricija, M. Barker, O. Petrauskene, M. R. Furtado, and M. Wiedmann.** 2010. Comparative

genomics of the bacterial genus *Listeria*: Genome evolution is characterized by limited gene acquisition and limited gene loss. BMC Genomics. **11:**688.

32. **den Bakker, H. C., X. Didelot, E. D. Fortes, K. K. Nightingale, and M. Wiedmann.** 2008. Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. BMC Evol. Biol. **8:**277.

33. **Dila, D., E. Sutherland, L. Moran, B. Slatko, and E. A. Raleigh.** 1990. Genetic and sequence organization of the *mcrBC* locus of *Escherichia coli* K-12. J. Bacteriol. **172:**4888-4900.

34. **Doorduyn, Y., C. M. de Jager, W. K. van der Zwaluw, W. J. Wannet, A. van der Ende, L. Spanjaard, and Y. T. van Duynhoven.** 2006. First results of the active surveillance of *Listeria monocytogenes* infections in the Netherlands reveal higher than expected incidence. Euro Surveill. **11:**E060420.4.

35. **Dorscht, J., J. Klumpp, R. Bielmann, M. Schmelcher, Y. Born, M. Zimmer, R. Calendar, and M. J. Loessner.** 2009. Comparative genome analysis of *Listeria* bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. J. Bacteriol. **191:**7206-7215.

36. **Doumith, M., C. Buchrieser, P. Glaser, C. Jacquet, and P. Martin.** 2004. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. J. Clin. Microbiol. **42:**3819-3822.

37. **Doumith, M., C. Cazalet, N. Simoes, L. Frangeul, C. Jacquet, F. Kunst, P. Martin, P. Cossart, P. Glaser, and C. Buchrieser.** 2004. New aspects regarding evolution and

virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. Infect. Immun. **72:**1072-1083.

38. **Drevets, D. A., and M. S. Bronze.** 2008. *Listeria monocytogenes*: epidemiology, human disease, and mechanisms of brain invasion. FEMS Immunol. Med. Microbiol. **53:**151-165.

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

40. **Dunn, K. A., J. P. Bielawski, T. J. Ward, C. Urquhart, and H. Gu.** 2009. Reconciling ecological and genomic divergence among lineages of *Listeria* under an "extended mosaic genome concept". Mol. Biol. Evol. **26:**2605-2615.

41. **Dussurget, O.** 2008. New insights into determinants of *Listeria monocytogenes* virulence. Int. Rev. Cell. Mol. Biol. **270:**1-38.

42. **Dussurget, O., J. Pizarro-Cerda, and P. Cossart.** 2004. Molecular determinants of *Listeria monocytogenes* virulence. Annu. Rev. Microbiol. **58:**587-610.

43. **Eifert, J. D., P. A. Curtis, M. C. Bazaco, R. J. Meinersmann, M. E. Berrang, S. Kernodle, C. Stam, L. A. Jaykus, and S. Kathariou.** 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog. Dis. **2:**192-200.

44. **Elhanafi, D., V. Dutta, and S. Kathariou.** 2010. Genetic characterization of plasmid-associated benzalkonium chloride resistance determinants in a *Listeria monocytogenes* strain from the 1998-1999 outbreak. Appl. Environ. Microbiol. **76:**8231-8238.

45. **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:**2383-2390.

46. **Franciosa, G., C. Scalfaro, A. Maugliani, F. Floridi, A. Gattuso, S. Hodzic, and P. Aureli.** 2007. Distribution of epidemic clonal genetic markers among *Listeria monocytogenes* 4b isolates. J. Food Prot. **70:**574-581.

47. **Freitag, N. E., G. C. Port, and M. D. Miner.** 2009. *Listeria monocytogenes* - from saprophyte to intracellular pathogen. Nat. Rev. Microbiol. **7:**623-628.

48. **Frye, D. M., R. Zweig, J. Sturgeon, M. Tormey, M. LeCavalier, I. Lee, L. Lawani, and L. Mascola.** 2002. An outbreak of febrile gastroenteritis associated with delicatessen meat contaminated with *Listeria monocytogenes*. Clin. Infect. Dis. **35:**943-949.

49. **Fukuda, E., K. H. Kaminska, J. M. Bujnichi, and I. Kobayashi.** 2008. Cell death upon epigenetic genome methylation: a novel function of methyl-specific deoxyribonucleases. Genome Biology. **9:**R163.1-R163.22.

50. **Garcia-del Portillo, F., and P. Cossart.** 2007. An important step in *Listeria* lipoprotein research. J. Bacteriol. **189:**294-297.

51. **Gasanov, U., D. Hughes, and P. M. Hansbro.** 2005. Methods for the isolation and identification of *Listeria* spp. and *Listeria monocytogenes*: a review. FEMS Microbiol. Rev. **29:**851-875.

52. **Gerner-Smidt, P., and J. M. Whichard.** 2008. Foodborne disease trends and reports. Foodborne Pathog. Dis. **5:**365-367.

53. **Gillespie, I. A., J. McLauchlin, K. A. Grant, C. L. Little, V. Mithani, C. Penman, C. Lane, and M. Regan.** 2006. Changing pattern of human listeriosis, England and Wales, 2001-2004. Emerg. Infect. Dis. **12:**1361-1366.

54. **Gillespie, I. A., J. McLauchlin, C. L. Little, C. Penman, P. Mook, K. Grant, and S. J. O'Brien.** 2009. Disease presentation in relation to infection foci for non-pregnancy-associated human listeriosis in England and Wales, 2001 to 2007. J. Clin. Microbiol. **47:**3301-3307.

55. **Gilmour, M. W., M. Graham, G. Van Domselaar, S. Tyler, H. Kent, K. M. Trout-Yakel, O. Larios, V. Allen, B. Lee, and C. Nadon.** 2010. High-throughput genome sequencing of two *Listeria monocytogenes* clinical isolates during a large foodborne outbreak. BMC Genomics. **11:**120.

56. **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:**849-852.

57. **Goulet, V.** 2007. What can we do to prevent listeriosis in 2006? Clin. Infect. Dis. **44:**529-530.

58. **Goulet, V., C. Hedberg, A. Le Monnier, and H. de Valk.** 2008. Increasing incidence of listeriosis in France and other European countries. Emerg. Infect. Dis. **14:**734-740.

59. **Goulet, V., C. Jacquet, P. Martin, V. Vaillant, E. Laurent, and H. de Valk.** 2006. Surveillance of human listeriosis in France, 2001-2003. Euro Surveill. **11:**79-81.

60. **Graves, L. M., L. O. Helsel, A. G. Steigerwalt, R. E. Morey, M. I. Daneshvar, S. E. Roof, R. H. Orsi, E. D. Fortes, S. R. Milillo, H. C. den Bakker, M. Wiedmann, B. Swaminathan, and B. D. Sauders.** 2010. *Listeria marthii* sp. nov., isolated from the natural environment, Finger Lakes National Forest. Int. J. Syst. Evol. Microbiol. **60:**1280-1288.

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

62. **Guillet, C., O. Join-Lambert, A. Le Monnier, A. Leclercq, F. Mechai, M. F. Mamzer-Bruneel, M. K. Bielecka, M. Scortti, O. Disson, P. Berche, J. Vazquez-Boland, O. Lortholary, and M. Lecuit.** 2010. Human listeriosis caused by *Listeria ivanovii*. Emerg. Infect. Dis. **16:**136-138.

63. **Hain, T., C. Steinweg, and T. Chakraborty.** 2006. Comparative and functional genomics of Listeria spp. J. Biotechnol. **126:**37-51.

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

65. **Handa-Miya, S., B. Kimura, H. Takahashi, M. Sato, T. Ishikawa, K. Igarashi, and T. Fujii.** 2007. Nonsense-mutated *inlA* and *prfA* not widely distributed in *Listeria monocytogenes* isolates from ready-to-eat seafood products in Japan. Int. J. Food Microbiol. **117:**312-318.

66. **Hedberg, C.** 2006. *Listeria* in Europe: the need for a European surveillance network is growing. Euro Surveill. **11:**75-76.

67. **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:**3972-3979.

68. **Hoskisson, P. A., and M. C. Smith.** 2007. Hypervariation and phase variation in the bacteriophage 'resistome'. Curr. Opin. Microbiol. **10:**396-400.

69. **Huang, B., N. Fang, K. Dimovski, X. Wang, G. Hogg, and J. Bates.** 2011. Observation of a new pattern in serogroup-related PCR typing of *Listeria monocytogenes* 4b isolates. J. Clin. Microbiol. **49:**426-429.

70. **Jacquet, C., M. Doumith, J. I. Gordon, P. M. Martin, P. Cossart, and M. Lecuit.** 2004. A molecular marker for evaluating the pathogenic potential of foodborne *Listeria monocytogenes*. J. Infect. Dis. **189:**2094-2100.

71. **Jeltsch, A.** 2003. Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems? Gene. **317:**13-16.

72. **Kathariou, S.** 2003. Foodborne outbreaks of listeriosis and epidemic-associated lineages of *Listeria monocytogenes*, p. 243-256. *In* M. E. Torrence and R. E. Isaacson (ed.), Microbial food safety in animal agriculture-current topics, 1st ed. Iowa State Press, Ames, Iowa.

73. **Kathariou, S.** 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J. Food Prot. **65:**1811-1829.

74. **Kathariou, S., L. Graves, C. Buchrieser, P. Glaser, R. M. Siletzky, and B. Swaminathan.** 2006. Involvement of closely related strains of a new clonal group of *Listeria monocytogenes* in the 1998-99 and 2002 multistate outbreaks of foodborne listeriosis in the United States. Foodborne Pathog. Dis. **3:**292-302.

75. **Kim, J. W., and S. Kathariou.** 2009. Temperature-dependent phage resistance of *Listeria monocytogenes* epidemic clone II. Appl. Environ. Microbiol. **75:**2433-2438.

76. **Kim, J. W., R. M. Siletzky, and S. Kathariou.** 2008. Host ranges of *Listeria*-specific bacteriophages from the turkey processing plant environment in the United States. Appl. Environ. Microbiol. **74:**6623-6630.

77. **Klumpp, J., J. Dorscht, R. Lurz, R. Bielmann, M. Wieland, M. Zimmer, R. Calendar, and M. J. Loessner.** 2008. The terminally redundant, nonpermuted genome

of *Listeria* bacteriophage A511: a model for the SPO1-like myoviruses of gram-positive bacteria. J. Bacteriol. **190:**5753-5765.

78. **Knowle, D., R. E. Lintner, Y. M. Touma, and R. M. Blumenthal.** 2005. Nature of the promoter activated by C.PvuII, an unusual regulatory protein conserved among restriction-modification systems. J. Bacteriol. **187:**488-497.

79. **Kobayashi, I.** 2004. Restriction-modification systems as minimal forms of life. Nucleic Acids and Molecular Biology. **14:**19-62.

80. **Kobayashi, I.** 2001. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. **29:**3742-3756.

81. **Koch, J., and K. Stark.** 2006. Significant increase of listeriosis in Germany--epidemiological patterns 2001-2005. Euro Surveill. **11:**85-88.

82. **Kruger, T., C. Wild, and M. Noyer-Weidner.** 1995. McrB: a prokaryotic protein specifically recognizing DNA containing modified cytosine residues. EMBO J. **14:**2661-2669.

83. **Kumar, S., X. Cheng, S. Klimasauskas, S. Mi, J. Posfai, R. J. Roberts, and G. G. Wilson.** 1994. The DNA (cytosine-5) methyltransferases. Nucleic Acids Res. **22:**1-10.

84. **Leclercq, A., D. Clermont, C. Bizet, P. A. Grimont, A. Le Fleche-Mateos, S. M. Roche, C. Buchrieser, V. Cadet-Daniel, A. Le Monnier, M. Lecuit, and F. Allerberger.** 2010. *Listeria rocourtiae* sp. nov. Int. J. Syst. Evol. Microbiol. **60:**2210-2214.

85. **Lin, L. F., J. Posfai, R. J. Roberts, and H. Kong.** 2001. Comparative genomics of the restriction-modification systems in *Helicobacter pylori*. Proc. Natl. Acad. Sci. U. S. A. **98:**2740-2745.

86. **Liu, D., M. L. Lawrence, M. Wiedmann, L. Gorski, R. E. Mandrell, A. J. Ainsworth, and F. W. Austin.** 2006. *Listeria monocytogenes* subgroups IIIA, IIIB, and IIIC delineate genetically distinct populations with varied pathogenic potential. J. Clin. Microbiol. **44:**4229-4233.

87. **Loessner, M. J., and C. E. D. Rees.** 2005. *Listeria* phages: basics and applications, p. 362-379. *In* K. M. Waldor, D. I. Friedman, and S. L. Adhya (ed.), Phages: their role in bacterial pathogenesis and biotechnology. ASM Press, Washington, DC.

88. **Loessner, M. J., R. B. Inman, P. Lauer, and R. Calendar.** 2000. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. Mol. Microbiol. **35:**324-340.

89. **Lomonaco, S., Y. Chen, and S. J. Knabel.** 2008. Analysis of additional virulence genes and virulence gene regions in *Listeria monocytogenes* confirms the epidemiologic relevance of multi-virulence-locus sequence typing. J. Food Prot. **71:**2559-2566.

90. **Machata, S., S. Tchatalbachev, W. Mohamed, L. Jansch, T. Hain, and T. Chakraborty.** 2008. Lipoproteins of *Listeria monocytogenes* are critical for virulence and TLR2-mediated immune activation. J. Immunol. **181:**2028-2035.

91. **Mahony, J., O. McAuliffe, R. P. Ross, and D. van Sinderen.** 2011. Bacteriophages as biocontrol agents of food pathogens. Curr. Opin. Biotechnol. **22:**157-163.

92. **McClelland, M., M. Nelson, and E. Raschke.** 1994. Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases. Nucleic Acids Res. **22:**3640-3659.

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

94. **Meinersmann, R. J., R. W. Phillips, M. Wiedmann, and M. E. Berrang.** 2004. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes reveals clonal and recombination histories of three lineages. Appl. Environ. Microbiol. **70:**2193-2203.

95. **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:**337-351.

96. **Mullapudi, S., R. M. Siletzky, and S. Kathariou.** 2010. Diverse cadmium resistance determinants in *Listeria monocytogenes* isolates from the turkey processing plant environment. Appl. Environ. Microbiol. **76:**627-630.

97. **Murray, N. E.** 2002. 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. Microbiology. **148:**3-20.

98. **Nagornykh, M. O., E. S. Bogdanova, A. S. Protsenko, M. V. Zakharova, A. S. Solonin, and K. V. Severinov.** 2008. Regulation of gene expression in type II restriction-modification system. Genetika. **44:**606-615.

99. **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:**2386-2395.

100. **Nightingale, K. K., R. A. Ivy, A. J. Ho, E. D. Fortes, B. L. Njaa, R. M. Peters, and M. Wiedmann.** 2008. *inlA* premature stop codons are common among *Listeria monocytogenes* isolates from foods and yield virulence-attenuated strains that confer protection against fully virulent strains. Appl. Environ. Microbiol. **74:**6570-6583.

101. **Olier, M., F. Pierre, S. Rousseaux, J. P. Lemaitre, A. Rousset, P. Piveteau, and J. Guzzo.** 2003. Expression of truncated Internalin A is involved in impaired internalization of some *Listeria monocytogenes* isolates carried asymptomatically by humans. Infect. Immun. **71:**1217-1224.

102. **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:**962-967.

103. **Orlowski, J., and J. M. Bujnicki.** 2008. Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. Nucleic Acids Res. **36:**3552-3569.

104. **Orndorff, P. E., T. S. Hamrick, I. W. Smoak, and E. A. Havell.** 2006. Host and bacterial factors in listeriosis pathogenesis. Vet. Microbiol. **114:**1-15.

105. **Orsi, R. H., M. L. Borowsky, P. Lauer, S. K. Young, C. Nusbaum, J. E. Galagan, B. W. Birren, R. A. Ivy, Q. Sun, L. M. Graves, B. Swaminathan, and M. Wiedmann.** 2008. Short-term genome evolution of *Listeria monocytogenes* in a non-controlled environment. BMC Genomics. **9:**539.

106. **Orsi, R. H., Q. Sun, and M. Wiedmann.** 2008. Genome-wide analyses reveal lineage specific contributions of positive selection and recombination to the evolution of *Listeria monocytogenes*. BMC Evol. Biol. **8:**233.

107. **O'Sullivan, D., D. P. Twomey, A. Coffey, C. Hill, G. F. Fitzgerald, and R. P. Ross.** 2000. Novel type I restriction specificities through domain shuffling of HsdS subunits in *Lactococcus lactis*. Mol. Microbiol. **36:**866-875.

108. **Pieper, U., and A. Pingoud.** 2002. A mutational analysis of the PD...D/EXK motif suggests that McrC harbors the catalytic center for DNA cleavage by the GTP-dependent restriction enzyme McrBC from Escherichia coli. Biochemistry. **41:**5236-5244.

109. **Pieper, U., T. Schweitzer, D. H. Groll, F. U. Gast, and A. Pingoud.** 1999. The GTP-binding domain of McrB: more than just a variation on a common theme? J. Mol. Biol. **292:**547-556.

110. **Pingoud, A., M. Fuxreiter, V. Pingoud, and W. Wende.** 2005. Type II restriction endonucleases: structure and mechanism. Cell Mol. Life Sci. **62:**685-707.

111. **Pingoud, A., and A. Jeltsch.** 2001. Structure and function of type II restriction endonucleases. Nucleic Acids Res. **29:**3705-3727.

112. **Proctor, M. E., R. Brosch, J. W. Mellen, L. A. Garrett, C. W. Kaspar, and J. B. Luchansky.** 1995. Use of pulsed-field gel electrophoresis to link sporadic cases of invasive listeriosis with recalled chocolate milk. Appl. Environ. Microbiol. **61:**3177-3179.

113. **Ragon, M., T. Wirth, F. Hollandt, R. Lavenir, M. Lecuit, A. Le Monnier, and S. Brisse.** 2008. A new perspective on *Listeria monocytogenes* evolution. PLoS Pathog. **4:**e1000146.

114. **Rajabian, T., B. Gavicherla, M. Heisig, S. Muller-Altrock, W. Goebel, S. D. Gray-Owen, and K. Ireton.** 2009. The bacterial virulence factor InlC perturbs apical cell junctions and promotes cell-to-cell spread of *Listeria*. Nat. Cell Biol. **11:**1212-1218.

115. **Raleigh, E. A.** 1992. Organization and function of the *mcrBC* genes of *Escherichia coli* K-12. Mol. Microbiol. **6:**1079-1086.

116. **Reglier-Poupet, H., C. Frehel, I. Dubail, J. L. Beretti, P. Berche, A. Charbit, and C. Raynaud.** 2003. Maturation of lipoproteins by type II signal peptidase is required for phagosomal escape of *Listeria monocytogenes*. J. Biol. Chem. **278:**49469-49477.

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

118. **Roberts, R. J., M. Belfort, T. Bestor, A. S. Bhagwat, T. A. Bickle, J. Bitinaite, R. M. Blumenthal, S. K. Degtyarev, D. T. Dryden, K. Dybvig, K. Firman, E. S. Gromova, R. I. Gumport, S. E. Halford, S. Hattman, J. Heitman, D. P. Hornby, A. Janulaitis, A. Jeltsch, J. Josephsen, A. Kiss, T. R. Klaenhammer, I. Kobayashi, H. Kong, D. H. Kruger, S. Lacks, M. G. Marinus, M. Miyahara, R. D. Morgan, N. E. Murray, V. Nagaraja, A. Piekarowicz, A. Pingoud, E. Raleigh, D. N. Rao, N. Reich, V. E. Repin, E. U. Selker, P. C. Shaw, D. C. Stein, B. L. Stoddard, W. Szybalski, T. A. Trautner, J. L. Van Etten, J. M. Vitor, G. G. Wilson, and S. Y. Xu.** 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucleic Acids Res. **31:**1805-1812.

119. **Rocha, E. P., A. Danchin, and A. Viari.** 2001. Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. Genome Res. **11:**946-958.

120. **Salamina, G., E. Dalle Donne, A. Niccolini, G. Poda, D. Cesaroni, M. Bucci, R. Fini, M. Maldini, A. Schuchat, B. Swaminathan, W. Bibb, J. Rocourt, N. Binkin, and S. Salmaso.** 1996. A foodborne outbreak of gastroenteritis involving *Listeria monocytogenes*. Epidemiol. Infect. **117:**429-436.

121. **Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin.** 2011. Foodborne illness acquired in the United States-major pathogens. Emerg. Infect. Dis. **17:**7-15.

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

123. **Scortti, M., H. J. Monzo, L. Lacharme-Lora, D. A. Lewis, and J. A. Vazquez-Boland.** 2007. The PrfA virulence regulon. Microbes Infect. **9:**1196-1207.

124. **Sperry, K. E., S. Kathariou, J. S. Edwards, and L. A. Wolf.** 2008. Multiple-locus variable-number tandem-repeat analysis as a tool for subtyping *Listeria monocytogenes* strains. J. Clin. Microbiol. **46:**1435-1450.

125. **Srikhanta, Y. N., K. L. Fox, and M. P. Jennings.** 2010. The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. Nat. Rev. Microbiol. **8:**196-206.

126. **Stack, H. M., C. Hill, and C. G. M. Gahan.** 2008. Stress Responses, p. 61-96. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

127. **Steinweg, C., C. T. Kuenne, A. Billion, M. A. Mraheil, E. Domann, R. Ghai, S. B. Barbuddhe, U. Karst, A. Goesmann, A. Puhler, B. Weisshaar, J. Wehland, R. Lampidis, J. Kreft, W. Goebel, T. Chakraborty, and T. Hain.** 2010. Complete genome sequence of *Listeria seeligeri*, a nonpathogenic member of the genus Listeria. J. Bacteriol. **192:**1473-1474.

128. **Sutcliffe, I. C., and D. J. Harrington.** 2002. Pattern searches for the identification of putative lipoprotein genes in Gram-positive bacterial genomes. Microbiology. **148:**2065-2077.

129. **Sutcliffe, I. C., and R. R. Russell.** 1995. Lipoproteins of gram-positive bacteria. J. Bacteriol. **177:**1123-1128.

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

131. **Tesfazgi Mebrhatu, M., E. Wywial, A. Ghosh, C. W. Michiels, A. B. Lindner, F. Taddei, J. M. Bujnicki, L. Van Melderen, and A. Aertsen.** 2011. Evidence for an evolutionary antagonism between Mrr and Type III modification systems. Nucleic Acids Res.

132. **Thomas, C. M., and K. M. Nielsen.** 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. **3:**711-721.

133. **Tock, M. R., and D. T. Dryden.** 2005. The biology of restriction and anti-restriction. Curr. Opin. Microbiol. **8:**466-472.

134. **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.

135. **Van Stelten, A., and K. K. Nightingale.** 2008. Development and implementation of a multiplex single-nucleotide polymorphism genotyping assay for detection of virulence-

attenuating mutations in the *Listeria monocytogenes* virulence-associated gene *inlA*. Appl. Environ. Microbiol. **74:**7365-7375.

136. **Van Stelten, A., J. M. Simpson, T. J. Ward, and K. K. Nightingale.** 2010. Revelation by single-nucleotide polymorphism genotyping that mutations leading to a premature stop codon in *inlA* are common among *Listeria monocytogenes* isolates from ready-to-eat foods but not human listeriosis cases. Appl. Environ. Microbiol. **76:**2783-2790.

137. **Vaneechoutte, M., P. Boerlin, H. V. Tichy, E. Bannerman, B. Jager, and J. Bille.** 1998. Comparison of PCR-based DNA fingerprinting techniques for the identification of *Listeria* species and their use for atypical *Listeria* isolates. Int. J. Syst. Bacteriol. **48 Pt 1:**127-139.

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

139. **Verghese, B., M. Lok, J. Wen, V. Alessandria, Y. Chen, S. Kathariou, and S. Knabel.** 2011. *comK* prophage junction fragments as markers for *Listeria monocytogenes* genotypes unique to individual meat and poultry processing plants and a model for rapid niche-specific adaptation, biofilm formation, and persistence. Appl. Environ. Microbiol. **77:**3279-3292.

140. **Voetsch, A. C., F. J. Angulo, T. F. Jones, M. R. Moore, C. Nadon, P. McCarthy, B. Shiferaw, M. B. Megginson, S. Hurd, B. J. Anderson, A. Cronquist, D. J. Vugia, C.**

**Medus, S. Segler, L. M. Graves, R. M. Hoekstra, P. M. Griffin, and Centers for Disease Control and Prevention Emerging Infections Program Foodborne Diseases Active Surveillance Network Working Group.** 2007. Reduction in the incidence of invasive listeriosis in foodborne diseases active surveillance network sites, 1996-2003. Clin. Infect. Dis. **44:**513-520.

141. **Wagner, M., and J. MacLauchlin.** 2008. Biology, p. 3-25. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

142. **Ward, T. J., T. F. Ducey, T. Usgaard, K. A. Dunn, and J. P. Bielawski.** 2008. Multilocus genotyping assays for single nucleotide polymorphism-based subtyping of *Listeria monocytogenes* isolates. Appl. Environ. Microbiol. **74:**7629-7642.

143. **Ward, T. J., P. Evans, M. Wiedmann, T. Usgaard, S. E. Roof, S. G. Stroika, and K. Hise.** 2010. Molecular and phenotypic characterization of *Listeria monocytogenes* from U.S. Department of Agriculture Food Safety and Inspection Service surveillance of ready-to-eat foods and processing facilities. J. Food Prot. **73:**861-869.

144. **Ward, T. J., L. Gorski, M. K. Borucki, R. E. Mandrell, J. Hutchins, and K. Pupedis.** 2004. Intraspecific phylogeny and lineage group identification based on the *prfA* virulence gene cluster of *Listeria monocytogenes*. J. Bacteriol. **186:**4994-5002.

145. **Warriner, K., and A. Namvar.** 2009. What is the hysteria with *Listeria*? Trends in Food Science & Technology. **20:**245-254.

146. **Yildirim, S., D. Elhanafi, W. Lin, A. D. Hitchins, R. M. Siletzky, and S. Kathariou.** 2010. Conservation of genomic localization and sequence content of Sau3AI-like

restriction-modification gene cassettes among *Listeria monocytogenes* epidemic clone I and selected strains of serotype 1/2a. Appl. Environ. Microbiol. **76:**5577-5584.

147. **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:**4158-4164.

148. **Zimmer, M., E. Sattelberger, R. B. Inman, R. Calendar, and M. J. Loessner.** 2003. Genome and proteome of *Listeria monocytogenes* phage PSA: an unusual case for programmed + 1 translational frameshifting in structural protein synthesis. Mol. Microbiol. **50:**303-317.

## Chapter 2 Characterization of Atypical *Listeria monocytogenes* Serotype 4b Strains Harboring a Lineage II-Specific Gene Cassette


### ABSTRACT

*Listeria monocytogenes* is the etiological agent of listeriosis, a potentially lethal foodborne illness. The population of *L. monocytogenes* is divided into four lineages (I-IV). Serotype 4b in lineage I has been involved in numerous outbreaks and four serotype 4b epidemic-associated clonal groups (ECI, Ia, II, and V) have been identified. In this study, we characterized a panel of strains of serotype 4b that produced atypical results with a serotype-specific multiplex PCR. The isolates possessed the *lmo0734-lmo0739* gene cassette that had been thought to be specific to lineage II. The cassette was harbored in a genomically equivalent locus in these isolates and in lineage II strains. Three novel clonal groups (Group 1-3) were identified among these isolates. Pulsed-field gel electrophoresis assigned Group 1 into one cluster and Groups 2 and 3 into a separate cluster; however, multilocus genotyping (MLGT) classified each clonal group into distinct types. Groups 1 and 2 were composed of clinical isolates from multiple states in the United States whereas Group 3 consisted of clinical and environmental isolates solely from North Carolina. In addition, all Group 3 isolates had DNA that was resistant to MboI, suggesting methylation of adenines of GATC sites. These findings suggest that the lineage II-specific *lmo0734-lmo0739* cassette has been acquired via horizontal gene transfer by at least three diverse clonal groups of *L. monocytogenes* serotype 4b, one of which exhibits restrictions in its regional distribution.

## INTRODUCTION

*Listeria monocytogenes* is a Gram-positive bacterium that is ubiquitously present in the environment. Human listeriosis is clinically manifested as various symptoms from febrile gastroenteritis to invasive meningitis, encephalitis, septicemia, stillbirths and abortions. In the United States, listeriosis is estimated to cause 1600 cases and 250 deaths in a year, suggesting that listeriosis occurs rarely but with a high mortality rate (16%) as a foodborne pathogen (27).

Serological reactions with somatic (O) and flagellar (H) antigens designate *L. monocytogenes* into thirteen serotypes: 1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e, and 7 (11). Epidemiologically, some serotypes are disproportionately associated with listeriosis: serotypes 1/2a, 1/2b and 4b are associated with over 95% of clinical cases (28). Studies on population structure of *L. monocytogenes* have revealed serotype-associated lineages designated lineages I-IV (14, 26, 30). Lineage I consists of serotypes 4b, 1/2b, and 3b (7) and is characterized with pronounced clonality (5, 20, 23, 24, 31), and lineage II is composed of serotypes 1/2a, 1/2c, 3a and 3c (26, 31). Lineage III includes isolates of 4a, and 4c, as well as certain strains of serotype 4b (26, 31). Recently, a subgroup of lineage III has been assigned into an independent population lineage (IV) (30).

In order to investigate differences in genomic content among different lineages, Doumith et al. hybridized 113 *Listeria* strains to a DNA array that included genes of *L. innocua* CLIP 11262, *L. monocytogenes* EGDe (serotype 1/2a) and *L. monocytogenes* CLIP 80459 (serotype 4b) (7). The information on the species-, lineage-, and serotype-specific genes that was obtained from this hybridization was utilized to devise a PCR-based method

to determine serotype designations as an attempt to simplify the canonical serotyping method that employs different H and O antibodies (6). In this method, five primer pairs were included in the multiplex PCR that divided *L. monocytogenes* into four serotype-associated groups: serotypes 1/2a or 3a; serotypes 1/2b or 3b; 4b complex; and serotype 1/2b or 3b.

During routine determination of serotype designations of serotype 4b *L. monocytogenes* by the multiplex PCR, we noted certain isolates with the atypical profile that included the amplicons expected for serotype 4b and the band anticipated for lineage II strains (Fig. 1). This PCR profile was the same as that recently described for four serotype 4b isolates which were isolated from human listeriosis cases in Australia and which harbored the gene (*lmo0737*) from which the lineage II-specific primers were derived (13). This gene is a member of a gene cassette (*lmo0734-lmo0739*) until now considered to be unique to lineage II strains and absent from those of lineage I (7, 21).

The objectives of the current study were to characterize the prevalence and genotypic characteristics of serotype 4b isolates exhibiting the atypical profile in the multiplex PCR serotyping that were collected in the United States. Also, we attempted to determine whether such strains harbor the entire gene cassette and whether the cassette is harbored in a genomically equivalent location in such isolates and lineage II strains.

## MATERIALS AND METHODS

**Bacterial strains and culture conditions.** The serotype 4b *L. monocytogenes* isolates with atypical multiplex PCR results are listed in Table 1. Bacterial cells were routinely grown at 37˚C overnight either in brain heart infusion (BHI; Becton, Dickinson and

Company, Sparks, MD) broth or on BHI plates containing 1.2% agar (Becton, Dickinson and Company).

**Genomic DNA extraction, enzyme digestion, polymerase chain reaction (PCR) and multiplex PCR serotyping.** The genomic DNA was isolated from the overnight liquid or plate cultures with DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA) following the manufacturer's recommendations. Genomic DNA was treated with Sau3AI and MboI (New England Biolabs, Ipswich, MA) and compared with undigested genomic DNA on Tris/Borate/EDTA (TBE) gels containing 0.8% agarose (Apex BioResearch Products, Research Triangle Park, NC).

PCR was conducted with exTaq (Takara, Otsu, Shiga, Japan) in a thermocycler (Biometra, Goettingen, Germany) under the conditions described in Chapter 4. The primers used in this study were synthesized by Eurofins MWG Operon (Huntsville, AL) and are shown in Table 2. The PCR amplicon was run on the agarose gels and visualized under the UV light. Multiplex PCR serotyping was performed according to the protocol developed by Doumith et al. (6).

**Sequencing and analysis.** The template for sequencing reaction was obtained by purifying the gel fragment containing the PCR amplicon with QIAquick gel extraction kit (QIAGEN) and sequencing was conducted by GeneWiz (NJ, USA). Sequencing data were manually combined into a contiguous strand and the open reading frames (ORFs) were identified and annotated with the ORF finder (www.ncbi.nlm.nih.gov/projects/gorf/), BLASTp (1) and conserved domain search (18) provided by the National Center for Biotechnology Information (NCBI). Sequences were compared with EMBOSS Align

(http://www.ebi.ac.uk/Tools/emboss/align/), ClutalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html) (4) or the online Artemis Comparison Tool (WebACT, http://www.webact.org/WebACT/home) (2). To determine whether a DNA mutation is synonymous, the DNA sequence was translated into a protein sequence with EMBOSS Transeq (http://www.ebi.ac.uk/Tools/emboss/transeq/) (25). GC contents were calculated with Oligonucleotide Properties Calculator (http://www.basic.northwestern.edu/biotools/oligocalc.html) (16).

**DNA-DNA hybridization.** The DNA probes were labeled and hybridized to bacterial genomic DNA immobilized on the nylon membrane (GE Water &Process Technologies, Trevose, PA) with the DIG system (Roche, Florence, SC) as described in Chapter 4. The hybridization signal was visually scrutinized after the exposure to X-ray films for various lengths of time. The DNA probes employed in this study are tabulated in Table 2.

**Pulse-Field Gel Electrophoresis (PFGE), targeted multilocus genotyping (TMLGT) and multilocus genotyping (MLGT).** PFGE was conducted with AscI (New England Biolabs) and ApaI (Roche) following the procedures depicted by Graves et al. (12) and analyzed with BioNumerics (Applied Maths, Austin, TX). TMLGT and MLGT were conducted as described (8, 32).

**Nucleotide sequence accession number.** The nucleotide sequences of *lmo0731-lmo0740* region in the atypical serotype 4b isolates 2007-454 and 2007-0904 have been submitted to GenBank (accession nos. JN236006 and JN236007, respectively).

## RESULTS

**The atypical serotype 4b isolates are divided into three clonal groups.** The serotype 4b isolates investigated in this study were first identified through their atypical multiplex PCR serotyping results: the multiplex PCR amplicons included the bands expected for serotype 4b strains as well as an extra band of the same size as the band generated with lineage II-specific *lmo0737* primers (Fig. 1). The presence of *lmo0737* was confirmed with hybridizations using an internal portion of this gene as DNA probe (data not shown). To collect further information on the distribution of *lmo0737*, we examined *L. monocytogenes* isolates of various serotypes in terms of their reactivity with this probe. A total of 463 *L.* monocytogenes isolates were hybridized with the *lmo0737* DNA probe, including 210 of serotype 4b, 82 of serotype 1/2b (or 3b), 114 of serotype 1/2a (or 3a), 11 of serotype 1/2c (or 3c) and 46 of other serotypes such as 4a and 4c. Our results revealed that 23 serotype 4b isolates hybridized with the *lmo0737* probe (Table 1), amounting to 11 % of the total serotype 4b isolates analyzed. Of these 23 serotype 4b isolates, 19 were of clinical origin (9 % of the 215 clinical isolates tested) and the remaining four were of environmental origin (2 % of the 170 environmental or food isolates tested).

All tested serotype 1/2b (or 3b) isolates were negative with the *lmo0737* probe. On the other hand, *lmo0737* was present in all tested lineage II strains except for three serotype 1/2a (or 3a) isolates (OLM 5, OLM 6 and OLM 75). Of these isolates, the animal isolates OLM 5 and OLM 6 were closely related on the basis of PFGE and OLM 5 was classified into lineage III (Lm3.16), suggesting that OLM 6 might also belong to lineage III. Meanwhile,

MLGT revealed that the clinical isolate OLM 75 belonged to serotype 1/2a within lineage II (2.12_1/2a).

Genomic DNAs of the 23 serotype 4b isolates that were positive with the *lmo0737* probe were also hybridized with DNA probes specific for two major serotype 4b epidemic-associated clonal groups, ECI and ECII (ECI-specific 85M and ECII-specific Hsp001219). Genomic DNA was also digested with Sau3AI and MboI (Table 1) since DNA from ECI strains and from ECV strains associated with an outbreak of listeriosis in North Carolina (17) are known to be resistant to Sau3AI and MboI, respectively (9, 33). None of the 23 isolates hybridized with ECI or ECII-specific probes (data not shown). Genomic DNA of all isolates was susceptible to Sau3AI; however, nine isolates had DNA that was resistant to MboI (Table1), suggesting methylation of adenines at GATC sites on the genome of these isolates (19).

Hybridizations with additional DNA probes and the MboI digestion results grouped the 23 isolates into three well-defined clusters (Table 1 and Fig. 2), which will be referred to as Group 1, 2, and 3 in this study. Group 1 isolates (n=7) were positive with H18RP9/10, a probe derived from the ECII region adjacent to *inlAB* ("region 18" (3, 10)) even though, as mentioned above, they were negative with the ECII-specific probe. Group 1 isolates were negative with probes derived from the counterpart of region 18 in non-ECII strains (4bSF18 and Non-ECIIC-WAP). In contrast, Group 2 isolates (n=7) hybridized with both probe 4bSF18 and probe H18RP9/10, derived from each of these two different sequence types of region 18 (Table 2, Fig. 2). DNA from all Group 1 and Group 2 isolates was susceptible to MboI digestion (Table 1). All isolates of Group 1 and 2 were of clinical origin and isolated

from various states in the United States (Table 1). Group 3, on the other hand, was comprised of nine isolates from a single state, North Carolina, among which five were isolated from patients whereas four originated from the food processing plant environment. Genomic DNA of all isolates of Group 3 was resistant to MboI and the isolates shared a common hybridization profile distinct from those of Group 1 or 2, i.e., positive results for probes Non-ECIIC-WAP and 4bSF18, which are based on the region 18 of non-ECII strains (Table 1, Fig. 2). TMLGT and MLGT analyses revealed that all were members of lineage I and that isolates in each group had the same genotype (1.17_4b, 1.46_4b and 1.60_4b for Groups 1, 2 and 3, respectively) (Table 1).

The 23 isolates were grouped into two clusters based on analysis of PFGE profiles with AscI and ApaI. Group 1 strains formed one cluster whereas Groups 2 and 3 were allocated into another and could not be clearly differentiated by PFGE (Fig. 2).

**The *lmo0734* -*lmo0739* gene cassette is highly conserved and in a genomically equivalent region in *lmo0737*-positive serotype 4b isolates and lineage II strains.** In serotype 1/2a EGDe, *lmo0737* is a part of the lineage II-specific gene cluster encompassing *lmo0734* to *lmo0739* (7) (Figure 3). With the exception of *lmo0734*, the genes in this cluster encode proteins putatively related to sugar and ribose metabolism. The GC content of *lmo0735*-*lmo0739* ranged from 36 to 39%, which did not deviate considerably from the *L. monocytogenes* genome average of 38%. The *lmo0734* gene codes for a protein belonging to LacI transcriptional regulator family and the GC content (34%) was considerably lower than the genome average.

In the EGDe genome, the above-mentioned lineage II-specific genes are flanked by *lmo0733*, encoding a putative DNA binding protein, and *lmo0740*, encoding a hypothetical protein that contains a cAMP-binding protein domain. In the serotype 4b strain F2365, these flanking genes (*LMOf2365_0769* and *LMOf2365_0770*, respectively) are conserved and adjacent to each other, although part of *lmo0739* was present in the intergenic region (Fig. 4A).

In the proximity of the *lmo0734-lmo0739* lineage II-specific cluster, we identified a gene harbored by both F2365 (*LMOf2365_0768*) and EGDe (*lmo0732*) and encoding an internalin-family protein with a LPXTG motif (Fig. 4A). Interestingly, there was marked diversity between *lmo0732* and its counterpart in F2365 (77% and 75% identities at the nucleotide and deduced protein level, respectively), especially in the internal portions. *In silico* analysis confirmed that this gene was divergent not only between EGDe and F2365 but between lineage I and II isolates generally (data not shown). To determine whether serotype 4b isolates harboring *lmo0737* also harbored the other genes in the lineage II-specific cassette, and whether the putative internalin gene was lineage II-like as well, we sequenced the DNA region between the *lmo0731* and *lmo0740* homologs of strains 2007-0904 (Group 1) and 2007-454 (Group 3). Sequencing results showed that these strains harbored the entire cassette in the same genomic region as lineage II strains (Figure 3). The *lmo0734-lmo0739* gene cassette was highly conserved (99% identity at the nucleotide level) between 2007-0904 and 2007-454, and between these strains and EGDe (99% identity at the DNA level).

Close scrutiny of the sequence alignment of the *lmo0734-lmo0739* cassette revealed 30 single nucleotide polymorphisms (SNPs) among 2007-0904, 2007-454 and EGDe,

including 9 non-synonymous SNPs (Fig. 4B). In comparison with EGDe, 28 SNPs were shared between 2007-454 and 2007-0904 (Fig. 4B). Also, the *lmo0737* gene of 2007-454 and 2007-0904 was compared with the sequences of the Australian strains (13). This analysis exhibited three SNPs shared between the strains of the atypical serotype 4b PCR profile compared with the lineage II strains, one SNP specific to 2007-454 and 2007-0904, and one SNP specific to Australian strains (Fig. 4C). None of the SNPs resulted in changes in amino acid sequence.

Sequences flanking this cassette in 2007-0904 and 2007-454 were also highly conserved between these strains, and among these strains, EGDe and F2365. For instance, the percent identity of *lmo0733* homologs amounted to 98% between 2007-0904 and 2007-454 and to 96-99% among these atypical serotype 4b isolates, EGDe and F2365. The putative internalin family protein of 2007-0904 and 2007-454 was lineage I-like, exhibiting 99.9% identity at the nucleotide sequence level with its homolog in F2365 (*LMOf2365_0768*).

For the other *lmo0737*-posisitive serotype 4b isolates that were not subjected to the sequencing, the location of the *lmo0737* homolog was verified with PCR using primer sets annealing to *lmo0737* and the flanking gene (Fig. 5). All isolates produced amplicons of the size expected for EGDe, suggesting that the lineage II-specific cassette was located in the same genomic region in these isolates.

## DISCUSSION

In this study, we describe atypical serotype 4b isolates that harbor a lineage II-specific gene cassette composed of homologs of *lmo0734-lmo0739* in *L. monocytogenes*

EGDe (serotype 1/2a). Putative functions of *lmo074-lmo0739* products imply that this gene cassette is associated with PTS and sugar metabolism. However, the biological consequences of the presence of this gene cassette in serotype 4b strains are currently unclear. In fact, the possible functions of these genes in lineage II strains have also not been identified. A deletion mutant of this cassette in EGDe was found to have normal growth *in vitro*, including under the ribose-deficient conditions, and was not significantly impaired in invasion and intracellular growth (22). However, tiling microarray assessments in EGDe revealed that all genes in this cassette were up-regulated in the stationary phase at 37ºC. Furthermore, *lmo0735*, *0736*, *0738*, and *0739* were up-regulated in the intestine, and *lmo0739* was down-regulated in the blood of infected mice (29). Isogenic constructs of serotype 4b strains with and without this cassette would be useful in assessments of the functional impact of these genes for strains of this serotype.

In the genome of EGDe and other lineage II strains, the lineage II-specific cassette was located adjacent to a gene encoding a putative internalin protein with a LPXTG cell wall anchoring motif, and exhibiting noticeable diversity between lineage I and II strains. However, sequence analysis revealed that, in serotype 4b strains harboring the lineage II cassette, the sequence of the putative internalin protein gene was typical of those of other serotype 4b strains. Thus, the presence of the lineage II cassette in these serotype 4b strains does not appear to be accompanied by lineage II-like genetic content in the variable internalin gene in the vicinity of this cassette.

Hybridization with a panel of DNA probes and MLGT data revealed that serotype 4b isolates harboring the *lmo0734-lmo0739* lineage II-specific cassette could be divided into

three groups. Group 3 was unusual with regards to its geographical source and its MboI resistance. Group 3 consisted of clinical and environmental isolates from a single state, North Carolina, and DNA from all group 3 isolates were resistant to MboI, suggesting methylation of adenine at GATC sites (19). In contrast, Groups 1 and 2 were comprised of clinical isolates from multiple states, and DNA from these isolates could be digested by MboI. The unique epidemiological features of Group 3 might suggest the presence of a clonal group that may have emerged in North Carolina relatively recently, and that may need time to spread to geographically distant locations in the United States. Further studies will be needed to accurately monitor possible trends in distribution of strains of this group in different regions.

Interestingly, Groups 2 and 3 could not be differentiated by PFGE. These similarities in PFGE between genetically distinct strains may reflect different bands of size too similar to be adequately differentiated, and has been noted before in our laboratory during analysis of certain other strains (R. M. Siletzky and S. Kathariou, unpublished data). However, the MLGT analysis relying on the SNPs identified in various genomic locations (8) suggested that each clonal group could be identified into three distinct lineage I types, supporting that the atypical isolates consisted of three clonal groups.

None f the serotype 1/2b isolates that were screened by us and others (7) harbored this cassette. Hence, outside of lineage II, the *lmo0734-lmo0739* gene cassette seems to be exclusively harbored by certain serotype 4b isolates within lineage I. It is possible that this cassette was acquired by certain serotype 4b strains through horizontal gene transfer from lineage II strain. However, it still remains unclear how this cassette was obtained by multiple clonal groups. An alternative hypothesis would be that the cassette was present in an

ancestral lineage of *L. monocytogenes*, and has been maintained only in lineage II and certain clonal groups of serotype 4b within lineage I. However, the *lmo0734-lmo0739* cassette harbored by the atypical serotype 4b isolate exhibited a higher similarity with serotype 1/2a EGDe (percent identity 99%) than the genome average between serotypes 4b and 1/2a, which was estimated by comparing the concatenated housekeeping genes targeted by multilocus sequence typing (24)of F2365 (serotype 4b) and EGDe (percent identity 96%). This finding provides evidence favoring the horizontal gene transfer hypothesis from lineage II strains.

Isolates that exhibited a similar PCR profile in the multiplex PCR scheme (6) have been reported by Huang and colleagues (13). Although this group did not provide evidence for the presence of the entire *lmo0734-lmo0739* cassette, the available *lmo0737* gene sequences revealed common SNPs shared between the Australian isolates and U.S. isolates in comparison with lineage II, suggesting that these atypical serotype 4b strains might resemble each other. We also identified SNPs specific to the isolates of each geographical location, which is indicative of genetic changes between them. It remains unclear whether the Australian isolates belong to either of the clonal groups described in this study or whether these isolates represent another clonal group(s). Characterization and comparison of the isolates collected from different continents could shed further light into the evolution of the atypical serotype 4b strains harboring the *lmo0734-lmo0739* cassette.

In conclusion, we have described three novel clonal groups of serotype 4b *L. monocytogenes* that harbor a lineage II-specific cassette putatively involved in the sugar and ribose metabolism. Notably, one clonal group was unique in that the strains were only identified in one state and included several environmental isolates. These intriguing genetic

and epidemiological characteristics warrant further investigation to deepen our understanding of these clonal groups of serotype 4b *L. monocytogenes*.

**Table 2.1 Bacterial strains used in this study**

| Classification | Strain | Isolation year | Origin | State | Sau3AI | MboI | MLGT | Source[a] |
|---|---|---|---|---|---|---|---|---|
| Group 1 | 2007-0904 | 2007 | Clinical | NC | + | + | 1.17_4b | NCDHHS |
| | J2255 | 2003 | Clinical | GA | + | + | 1.17_4b | CDC |
| | J3139 | 2004 | Clinical | WI | + | + | 1.17_4b | CDC |
| | J3913 | 2006 | Clinical | IN | + | + | 1.17_4b | CDC |
| | J4016 | 2006 | Clinical | SC | + | + | 1.17_4b | CDC |
| | J4500 | 2007 | Clinical | NE | + | + | 1.17_4b | CDC |
| | J5000 | 2008 | Clinical | VA | + | + | 1.17_4b | CDC |
| Group 2 | 2007-618 | 2007 | Clinical | NC | + | + | 1.46_4b | NCDHHS |
| | J3026 | 2004 | Clinical | CT | + | + | 1.46_4b | CDC |
| | J3053 | 2004 | Clinical | MI | + | + | 1.46_4b | CDC |
| | J3195 | 2004 | Clinical | OH | + | + | 1.46_4b | CDC |
| | J4458 | 2007 | Clinical | MN | + | + | 1.46_4b | CDC |
| | J4490 | 2007 | Clinical | SC | + | + | 1.46_4b | CDC |
| | J4953 | 2008 | Clinical | OH | + | + | 1.46_4b | CDC |
| Group 3 | 2001-7R | 2001 | Clinical | NC | + | - | 1.60_4b | NCSU |
| | 2001-8R | 2001 | Clinical | NC | + | - | 1.60_4b | NCSU |
| | 2006-296 | 2006 | Clinical | NC | + | - | 1.60_4b | NCDHHS |
| | 2007-454 | 2007 | Clinical | NC | + | - | 1.60_4b | NCDHHS |
| | 2008-894 | 2008 | Clinical | NC | + | - | 1.60_4b | NCDHHS |
| | 18-2a | 2003 | Processing plant (floor) | NC | + | - | 1.60_4b | NCSU |
| | 128b-1 | 2004 | Processing plant (drain) | NC | + | - | 1.60_4b | NCSU |
| | 363b-1 | 2005 | Processing plant (drain) | NC | + | - | 1.60_4b | NCSU |
| | 491a-5 | 2006 | Processing plant (chiller re-work table) | NC | + | - | 1.60_4b | NCSU |

[a] CDC, Centers for Disease Control and Prevention; NCSU, *Listeria* Strain collection at North Carolina State University; and NCDHHS, North Carolina Department of Health and Human Services.

**Table 2.2 List of the probes and oligonucleotides employed in this study**

| Probe | Primer | Sequence (5' to 3') | Target | Reference |
|---|---|---|---|---|
| | lmo0731F | TACTTGCTGCTCTTTCTGGC | *lmo0731* | This study |
| | lmo0740Rl | GTGGAAGGTTATGCTGCAGTTG | *lmo0740* | This study |
| *lmo0737* | lmo0737F | AGGGCTTCAAGGACTTACCC | *lmo0737* | (6) |
| | lmo0737R | ACGATTTCTGCTTGCCATTC | *lmo0737* | (6) |
| H18RP9/10 | H18RP9/10F | CTGGATTTGCAGCTTATGAT | *LMOh7858_0487* | (15) |
| | H18RP9/10R | CCTATTCTTTCCATAAGTAAT | *LMOh7858_0487* | (15) |
| ECIIC-WAP | ECIIC-WAPF | GGGAACTTTCCATTAGCC | *LMOh7858_0479* | (3) |
| | ECIIC-WAPR | TTAAATGGGATATGATGT | *LMOh7858_0479* | (3) |
| Hsp001219 | Hspec01219F | GAGGCTATCGAAATTGCTCG | *LMOh7858_1168* | (3) |
| | Hspec01219R | AGGATTCGGAATTTCATCCA | *LMOh7858_1168* | (3) |
| 4bSF18 | 4bSF18_F | ACGGGCGTTTTATATTAAATGGG | *LMOf2365_0466* | (10) |
| | 4bSF18_R | AATATCTCGAAAACTCCGAGT | *LMOf2365_0466* | (10) |
| Non-ECIIC-WAP | ECIC-WAPF | ATGGAAATTGGGCATGGC | *LMOf2365_0450* | (3) |
| | ECIC-WAPR | GTAGTTCCAGTGGACATG | *LMOf2365_0450* | (3) |
| 85M | F2365_85MF | AATATATTTTCAATGTTTGATGGT | *LMOf2365_0327* | (34) |
| | F2365_85MR | GCTAATTCAATCCCTATTCT | *LMOf2365_0327* | (34) |

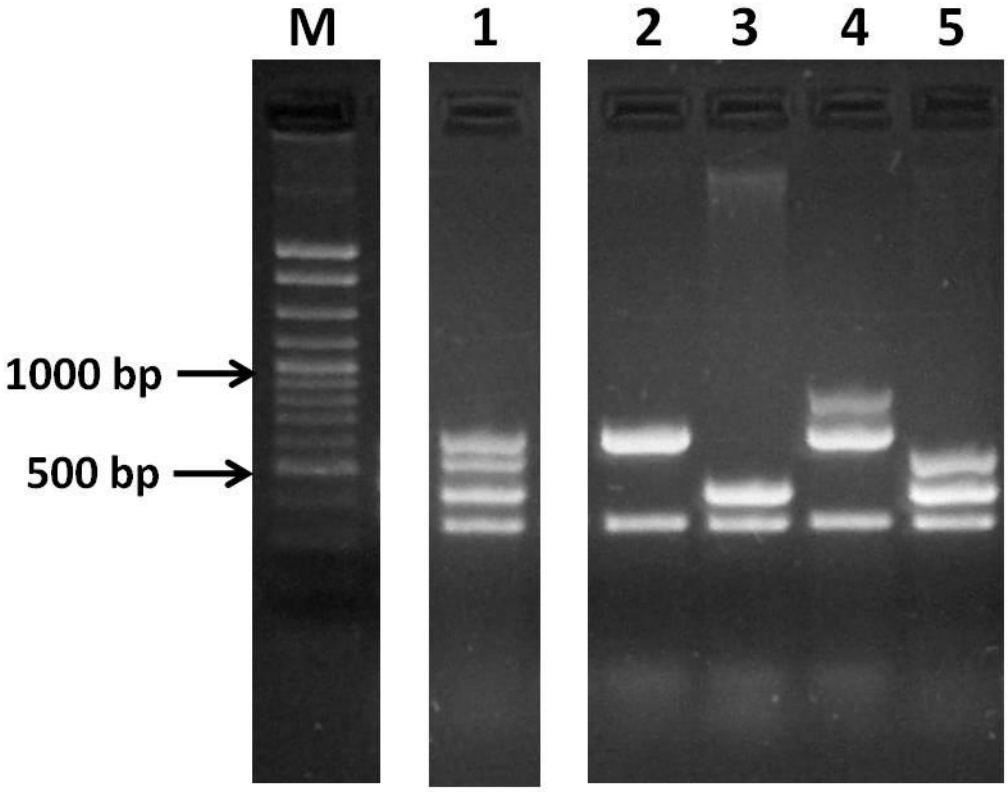



**Figure 2.1 Multiplex serotyping results of an atypical 4b isolate.** Lane 1, 2007-618 (atypical serotype 4b isolate); lane 2, LW-A113 (1/2a); lane 3, G3986 (3b); lane 4, G3969 (1/2c); and lane 5, G3992 (4b). M, exACTGene cloning DNA ladder (Fisher Scientific, Fair Lawn, NJ). From top to bottom, the bands correspond to *lmo1118* (906 bp), *lmo0737* (691 bp), ORF2110 (597 bp), ORF2819 (471 bp) and *prs* (370 bp). Similar results were obtained for other atypical serotype 4b isolates.

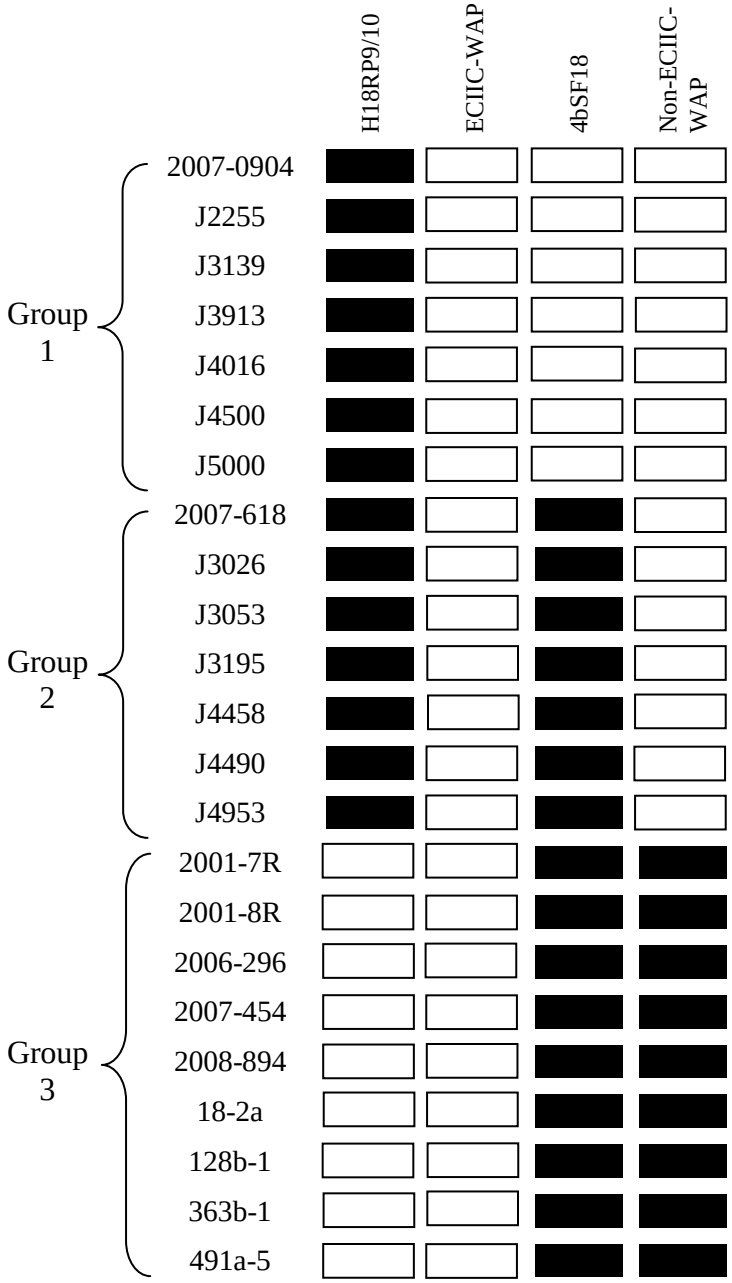



**Figure 2.2 DNA-DNA hybridization results.** Positive and negative results were represented with black and white squares.

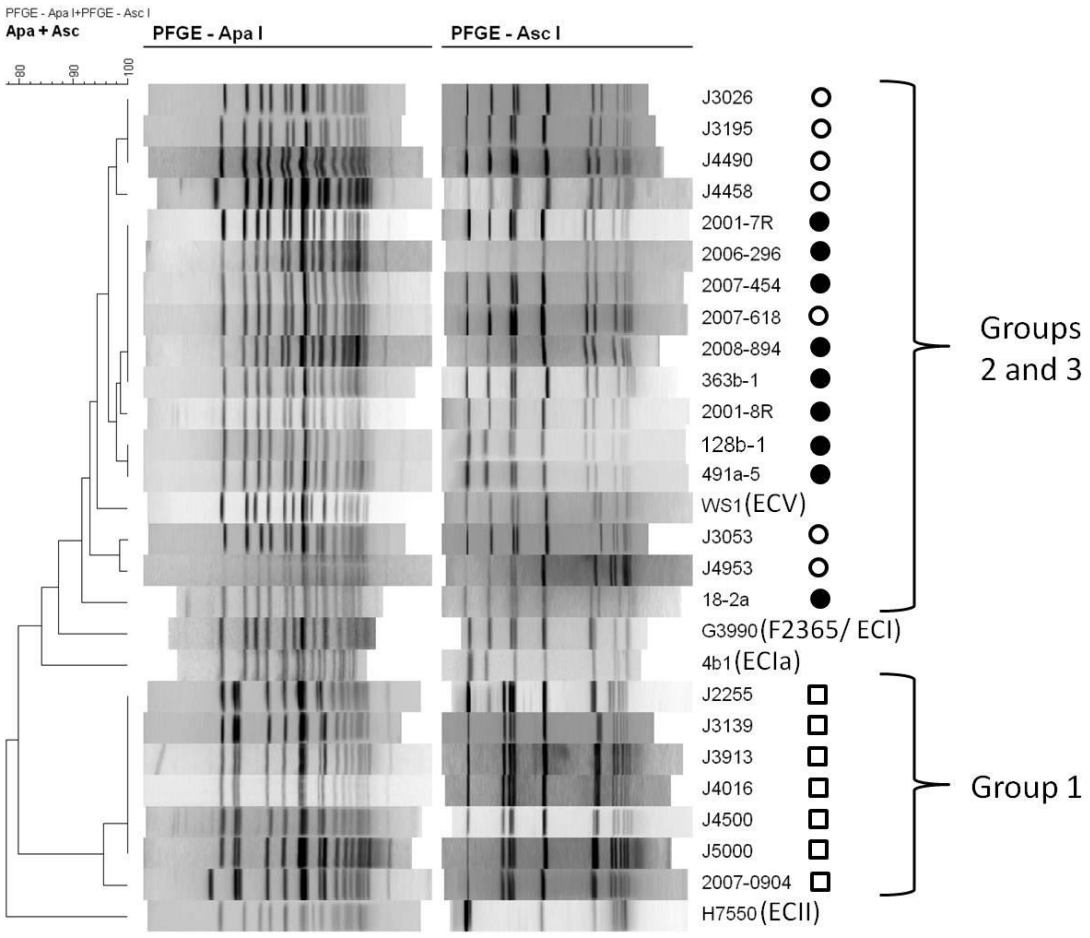
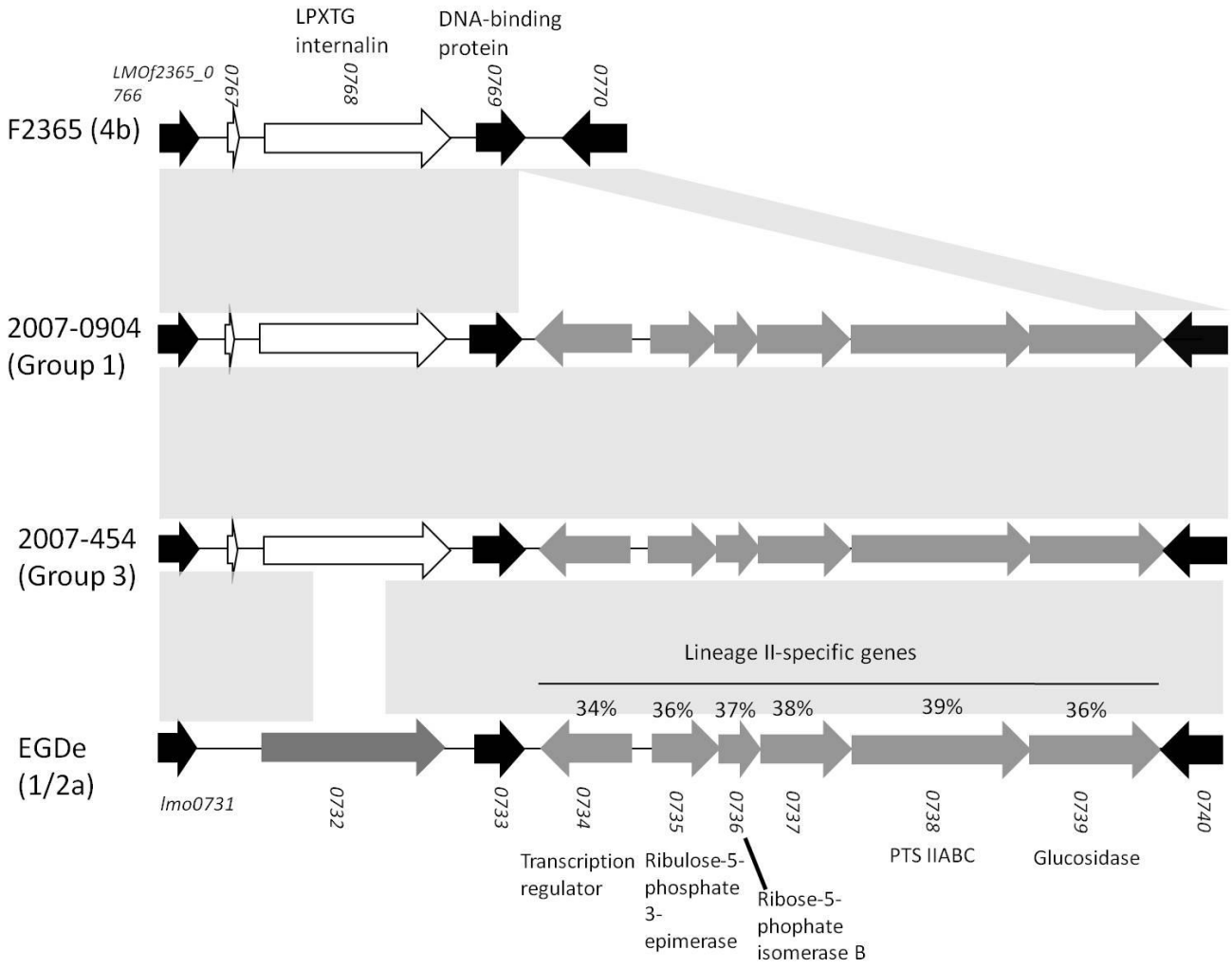


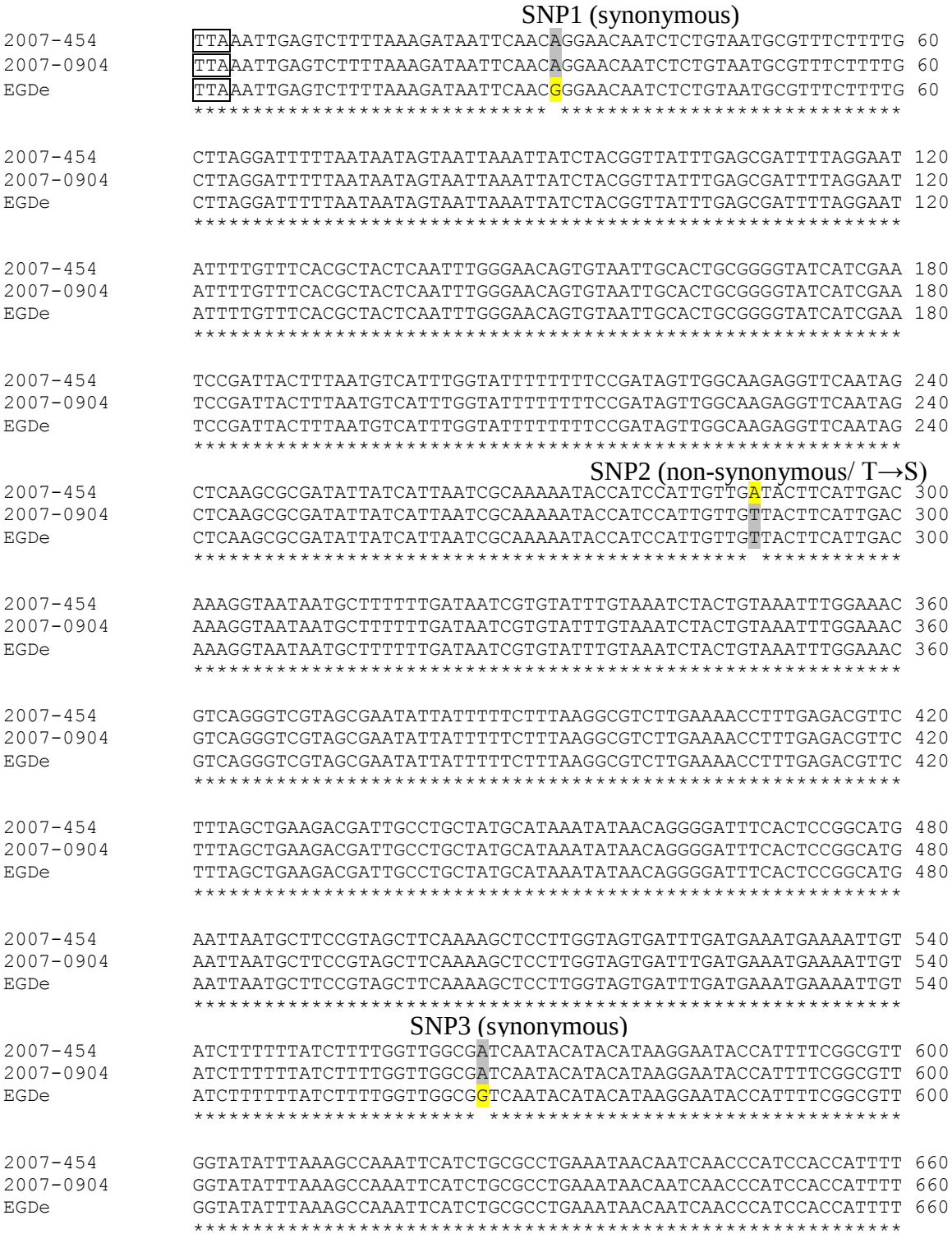

**Figure 2.3 PFGE dendrogram of atypical serotype 4b.** Groups 1-3 were marked with open rectangle, open circle and black circle, respectively. Serotype 4b strains F2365 (ECI), 4b1 (ECIa), H7550 (ECII) and WS1 (ECV) were included for reference.

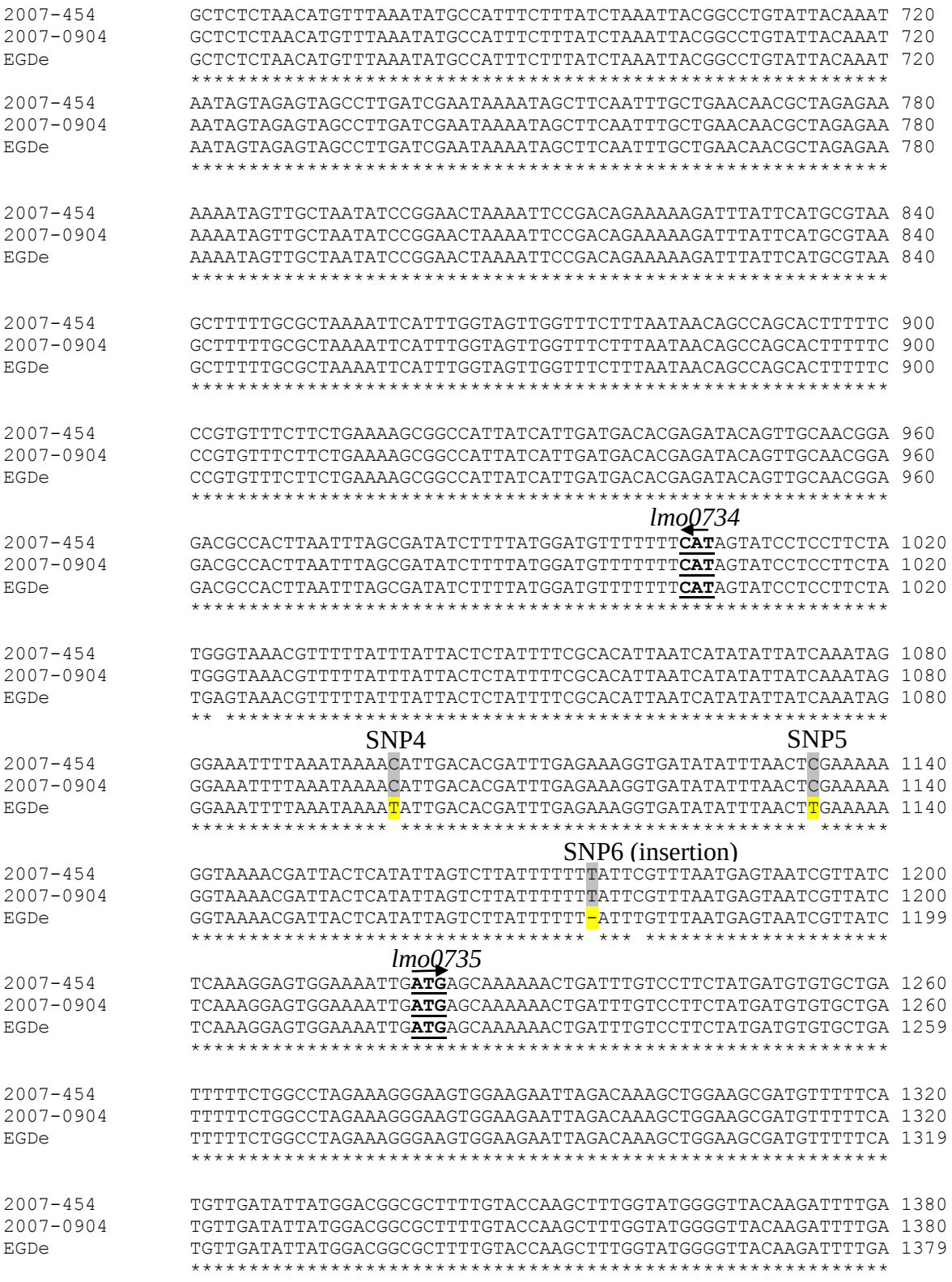
**Figure 2.4 DNA sequence comparison.** (A) Comparison of the genetic region harboring the *lmo0734-lmo0739* in EGDe (1/2a), F2365 (4b) and atypical serotype 4b isolates, 2007-0904 (Group 1) and 2007-454 (Group 3). The genes that are commonly present were represented in black; those conserved among F2365 and atypical 4b isolates were marked in white; the lineage II-specific genes including *lmo0737* were colored in grey; and the upstream divergent LPXTG internalin gene in EGDe was signified in darker grey. Homologous DNA sequences were marked with light grey box. The percentage above the lineage II-specific genes represents GC content of each gene. (B and C) Multiple sequence alignment of the *lmo0734-lmo0739* gene cassette of EGDe, 2007-454 and 2007-0904 (B) and of the *lmo0737* gene of 2007-454, 2007-0904, Australian strains along with EGDe and two additional lineage II strains (10M131 and 09M3788) (C). The start codons were boldfaced and underlined and the stop codons were marked with boxes. For each SNP, the dominant alleles were colored in grey and the other allele was marked in yellow. Next to the non-synonymous SNPs, the amino acid sequence change in comparison with the EGDe was shown. In C, the synonymous and non-synonymous SNPs were determined assuming that the DNA sequences from the Australian strains and the additional lineage II strains would be translated from the same reading frame as EGDe.

**A.**


LPXTG internalin
DNA-binding protein
LMOf2365_0766
0767
0768
0769
0770
F2365 (4b)
2007-0904 (Group 1)
2007-454 (Group 3)
Lineage II-specific genes
34%
36%
37%
38%
39%
36%
EGDe (1/2a)
lmo0731
0732
0733
0734
0735
0736
0737
0738
0739
0740
Transcription regulator
Ribulose-5-phosphate 3-epimerase
Ribose-5-phophate isomerase B
PTS IIABC
Glucosidase

**B.**


SNP1 (synonymous)
2007-454 TTAAATTGAGTCTTTTAAAGATAATTCAACAGGAACAATCTCTGTAATGCGTTTCTTTTG 60
2007-0904 TTAAATTGAGTCTTTTAAAGATAATTCAACAGGAACAATCTCTGTAATGCGTTTCTTTTG 60
EGDe TTAAATTGAGTCTTTTAAAGATAATTCAACGGGAACAATCTCTGTAATGCGTTTCTTTTG 60
2007-454 CTTAGGATTTTTAATAATAGTAATTAAATTATCTACGGTTATTTGAGCGATTTTAGGAAT 120
2007-0904 CTTAGGATTTTTAATAATAGTAATTAAATTATCTACGGTTATTTGAGCGATTTTAGGAAT 120
EGDe CTTAGGATTTTTAATAATAGTAATTAAATTATCTACGGTTATTTGAGCGATTTTAGGAAT 120
2007-454 ATTTTGTTTCACGCTACTCAATTTGGGAACAGTGTAATTGCACTGCGGGGTATCATCGAA 180
2007-0904 ATTTTGTTTCACGCTACTCAATTTGGGAACAGTGTAATTGCACTGCGGGGTATCATCGAA 180
EGDe ATTTTGTTTCACGCTACTCAATTTGGGAACAGTGTAATTGCACTGCGGGGTATCATCGAA 180
2007-454 TCCGATTACTTTAATGTCATTTGGTATTTTTTTTCCGATAGTTGGCAAGAGGTTCAATAG 240
2007-0904 TCCGATTACTTTAATGTCATTTGGTATTTTTTTTCCGATAGTTGGCAAGAGGTTCAATAG 240
EGDe TCCGATTACTTTAATGTCATTTGGTATTTTTTTTCCGATAGTTGGCAAGAGGTTCAATAG 240
SNP2 (non-synonymous/ T→S)
2007-454 CTCAAGCGCGATATTATCATTAATCGCAAAAATACCATCCATTGTTGATACTTCATTGAC 300
2007-0904 CTCAAGCGCGATATTATCATTAATCGCAAAAATACCATCCATTGTTGTTACTTCATTGAC 300
EGDe CTCAAGCGCGATATTATCATTAATCGCAAAAATACCATCCATTGTTGTTACTTCATTGAC 300
2007-454 AAAGGTAATAATGCTTTTTTTGATAATCGTGTATTTGTAAATCTACTGTAAATTTGGAAAC 360
2007-0904 AAAGGTAATAATGCTTTTTTTGATAATCGTGTATTTGTAAATCTACTGTAAATTTGGAAAC 360
EGDe AAAGGTAATAATGCTTTTTTTGATAATCGTGTATTTGTAAATCTACTGTAAATTTGGAAAC 360
2007-454 GTCAGGGTCGTAGCGAATATTATTTTTCTTTAAGGCGTCTTGAAAACCTTTGAGACGTTC 420
2007-0904 GTCAGGGTCGTAGCGAATATTATTTTTCTTTAAGGCGTCTTGAAAACCTTTGAGACGTTC 420
EGDe GTCAGGGTCGTAGCGAATATTATTTTTCTTTAAGGCGTCTTGAAAACCTTTGAGACGTTC 420
2007-454 TTTAGCTGAAGACGATTGCCTGCTATGCATAAATATAACAGGGGATTTCACTCCGGCATG 480
2007-0904 TTTAGCTGAAGACGATTGCCTGCTATGCATAAATATAACAGGGGATTTCACTCCGGCATG 480
EGDe TTTAGCTGAAGACGATTGCCTGCTATGCATAAATATAACAGGGGATTTCACTCCGGCATG 480
2007-454 AATTAATGCTTCCGTAGCTTCAAAAGCTCCTTGGTAGTGATTTGATGAAATGAAAATTGT 540
2007-0904 AATTAATGCTTCCGTAGCTTCAAAAGCTCCTTGGTAGTGATTTGATGAAATGAAAATTGT 540
EGDe AATTAATGCTTCCGTAGCTTCAAAAGCTCCTTGGTAGTGATTTGATGAAATGAAAATTGT 540
SNP3 (synonymous)
2007-454 ATCTTTTTTATCTTTTGGTTGGCGATCAATACATACATAAGGAATACCATTTTCGGCGTT 600
2007-0904 ATCTTTTTTATCTTTTGGTTGGCGATCAATACATACATAAGGAATACCATTTTCGGCGTT 600
EGDe ATCTTTTTTATCTTTTGGTTGGCGGTCAATACATACATAAGGAATACCATTTTCGGCGTT 600
2007-454 GGTATATTTAAAGCCAAATTCATCTGCGCCTGAAATAACAATCAACCCATCCACCATTTT 660
2007-0904 GGTATATTTAAAGCCAAATTCATCTGCGCCTGAAATAACAATCAACCCATCCACCATTTT 660
EGDe GGTATATTTAAAGCCAAATTCATCTGCGCCTGAAATAACAATCAACCCATCCACCATTTT 660


2007-454 GCTCTCTAACATGTTTAAATATGCCATTTCTTTATCTAAATTACGGCCTGTATTACAAAT 720
2007-0904 GCTCTCTAACATGTTTAAATATGCCATTTCTTTATCTAAATTACGGCCTGTATTACAAAT 720
EGDe GCTCTCTAACATGTTTAAATATGCCATTTCTTTATCTAAATTACGGCCTGTATTACAAAT 720
2007-454 AATAGTAGAGTAGCCTTGATCGAATAAAATAGCTTCAATTTGCTGAACAACGCTAGAGAA 780
2007-0904 AATAGTAGAGTAGCCTTGATCGAATAAAATAGCTTCAATTTGCTGAACAACGCTAGAGAA 780
EGDe AATAGTAGAGTAGCCTTGATCGAATAAAATAGCTTCAATTTGCTGAACAACGCTAGAGAA 780
2007-454 AAAATAGTTGCTAATATCCGGAACTAAAATTCCGACAGAAAAAGATTTATTCATGCGTAA 840
2007-0904 AAAATAGTTGCTAATATCCGGAACTAAAATTCCGACAGAAAAAGATTTATTCATGCGTAA 840
EGDe AAAATAGTTGCTAATATCCGGAACTAAAATTCCGACAGAAAAAGATTTATTCATGCGTAA 840
2007-454 GCTTTTTGCGCTAAAATTCATTTGGTAGTTGGTTTCTTTAATAACAGCCAGCACTTTTTC 900
2007-0904 GCTTTTTGCGCTAAAATTCATTTGGTAGTTGGTTTCTTTAATAACAGCCAGCACTTTTTC 900
EGDe GCTTTTTGCGCTAAAATTCATTTGGTAGTTGGTTTCTTTAATAACAGCCAGCACTTTTTC 900
2007-454 CCGTGTTTCTTCTGAAAAGCGGCCATTATCATTGATGACACGAGATACAGTTGCAACGGA 960
2007-0904 CCGTGTTTCTTCTGAAAAGCGGCCATTATCATTGATGACACGAGATACAGTTGCAACGGA 960
EGDe CCGTGTTTCTTCTGAAAAGCGGCCATTATCATTGATGACACGAGATACAGTTGCAACGGA 960
lmo0734
2007-454 GACGCCACTTAATTTAGCGATATCTTTTATGGATGTTTTTTTCATAGTATCCTCCTTCTA 1020
2007-0904 GACGCCACTTAATTTAGCGATATCTTTTATGGATGTTTTTTTCATAGTATCCTCCTTCTA 1020
EGDe GACGCCACTTAATTTAGCGATATCTTTTATGGATGTTTTTTTCATAGTATCCTCCTTCTA 1020
2007-454 TGGGTAAACGTTTTTATTTATTACTCTATTTTCGCACATTAATCATATATTATCAAATAG 1080
2007-0904 TGGGTAAACGTTTTTATTTATTACTCTATTTTCGCACATTAATCATATATTATCAAATAG 1080
EGDe TGAGTAAACGTTTTTATTTATTACTCTATTTTCGCACATTAATCATATATTATCAAATAG 1080
SNP4
SNP5
2007-454 GGAAATTTTAAATAAAACATTGACACGATTTGAGAAAGGTGATATATTTAACTCGAAAAA 1140
2007-0904 GGAAATTTTAAATAAAACATTGACACGATTTGAGAAAGGTGATATATTTAACTCGAAAAA 1140
EGDe GGAAATTTTAAATAAAATATTGACACGATTTGAGAAAGGTGATATATTTAACTTGAAAAA 1140
SNP6 (insertion)
2007-454 GGTAAAACGATTACTCATATTAGTCTTATTTTTTTATTCGTTTAATGAGTAATCGTTATC 1200
2007-0904 GGTAAAACGATTACTCATATTAGTCTTATTTTTTTATTCGTTTAATGAGTAATCGTTATC 1200
EGDe GGTAAAACGATTACTCATATTAGTCTTATTTTTT-ATTTGTTTAATGAGTAATCGTTATC 1199
lmo0735
2007-454 TCAAAGGAGTGGAAAATTGATGAGCAAAAAACTGATTTGTCCTTCTATGATGTGTGCTGA 1260
2007-0904 TCAAAGGAGTGGAAAATTGATGAGCAAAAAACTGATTTGTCCTTCTATGATGTGTGCTGA 1260
EGDe TCAAAGGAGTGGAAAATTGATGAGCAAAAAACTGATTTGTCCTTCTATGATGTGTGCTGA 1259
2007-454 TTTTTCTGGCCTAGAAAGGGAAGTGGAAGAATTAGACAAAGCTGGAAGCGATGTTTTTCA 1320
2007-0904 TTTTTCTGGCCTAGAAAGGGAAGTGGAAGAATTAGACAAAGCTGGAAGCGATGTTTTTCA 1320
EGDe TTTTTCTGGCCTAGAAAGGGAAGTGGAAGAATTAGACAAAGCTGGAAGCGATGTTTTTCA 1319
2007-454 TGTTGATATTATGGACGGCGCTTTTGTACCAAGCTTTGGTATGGGGTTACAAGATTTTGA 1380
2007-0904 TGTTGATATTATGGACGGCGCTTTTGTACCAAGCTTTGGTATGGGGTTACAAGATTTTGA 1380
EGDe TGTTGATATTATGGACGGCGCTTTTGTACCAAGCTTTGGTATGGGGTTACAAGATTTTGA 1379

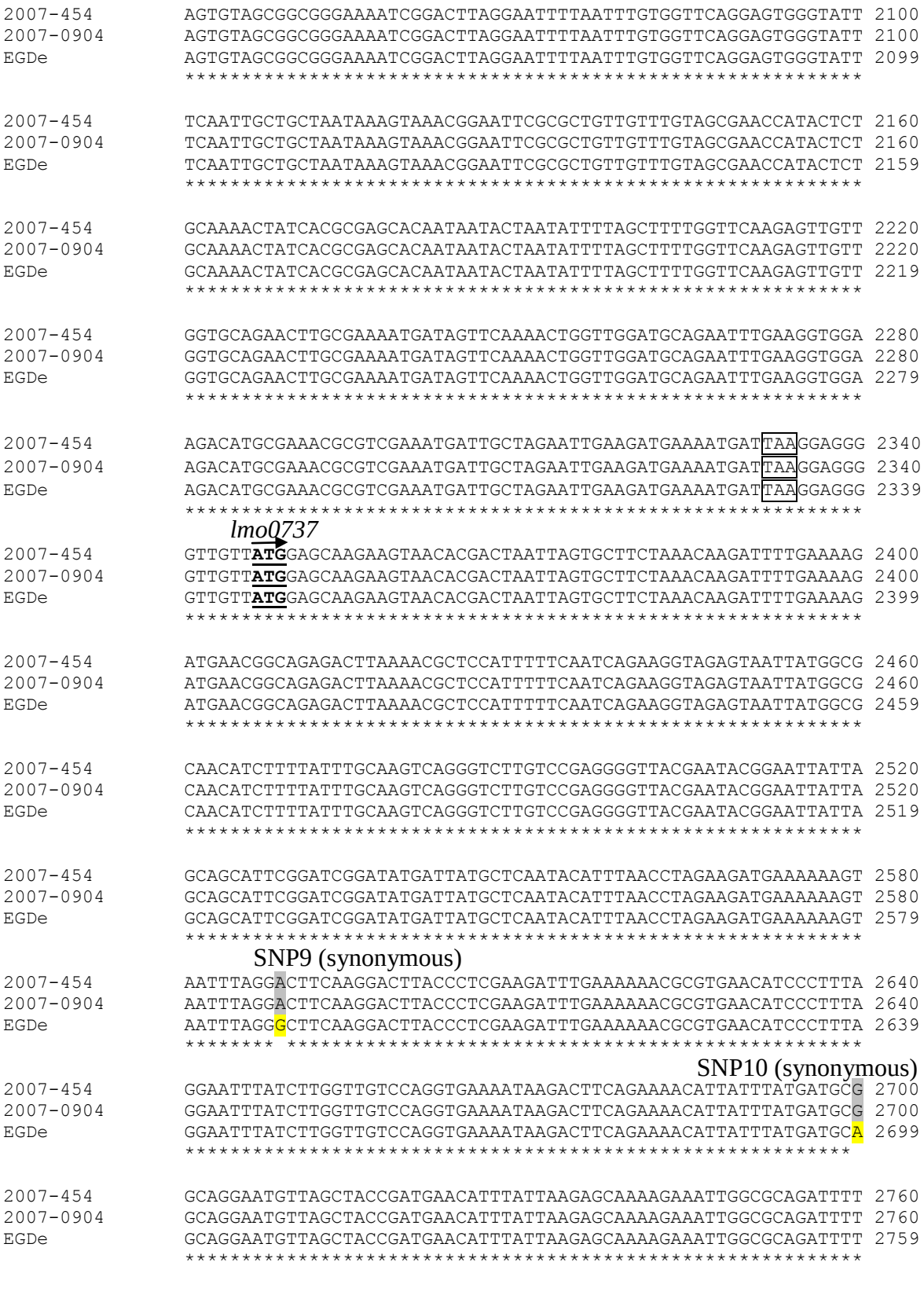
SNP7 (synonymous)

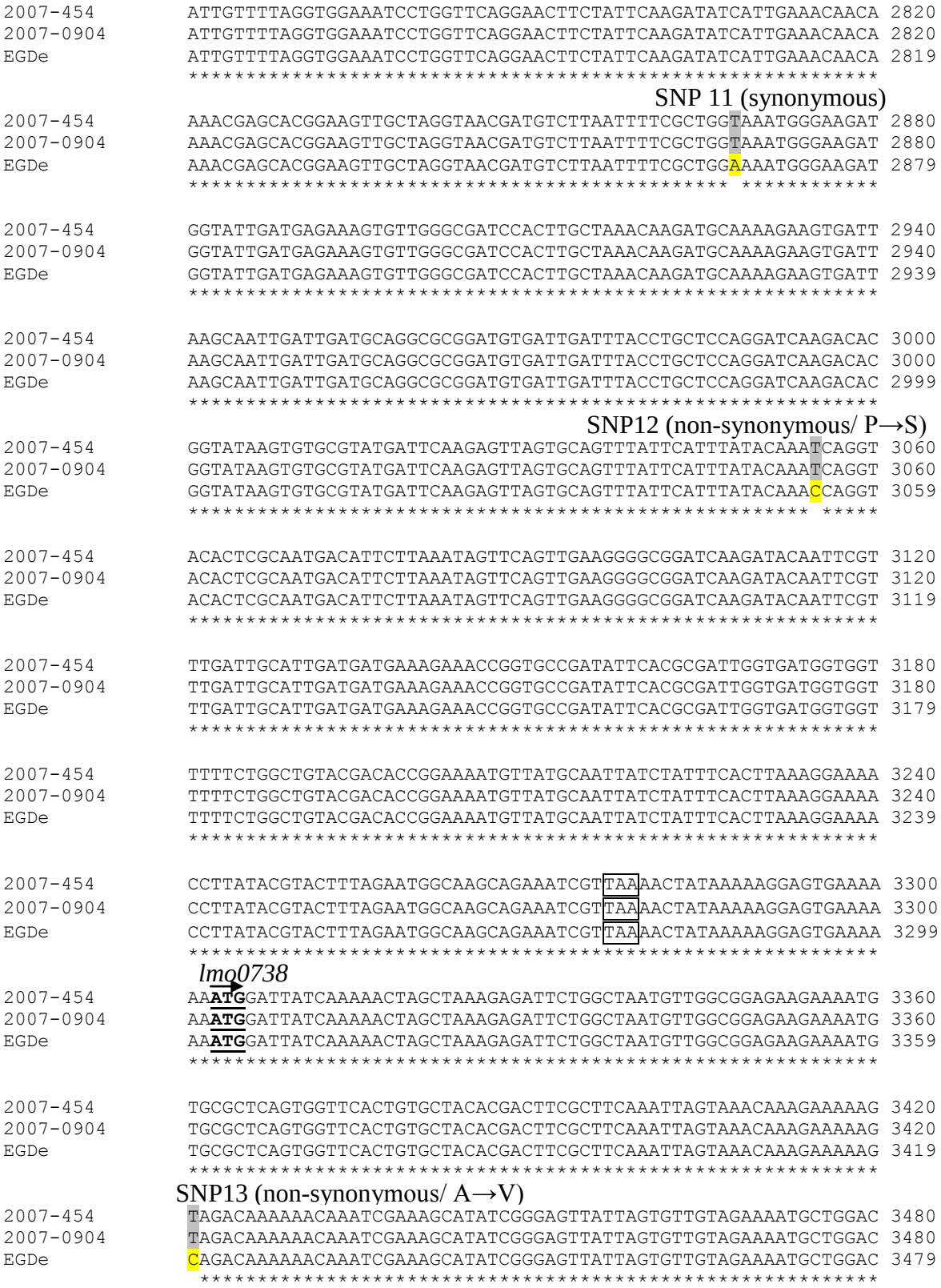
```
2007-454        AGAAATTCGTCAGAAAACTCAAAAGTTAGTTGACGTCCATTTGATGATCATGAATCCAGG 1440
2007-0904       AGAAATTCGTCAGAAAACTCAAAAGTTAGTTGACGTCCATTTGATGATCATGAATCCAGG 1440
EGDe            AGAAATTCGTCAAAAAACTCAAAAGTTAGTTGACGTCCATTTGATGATCATGAATCCAGG 1439
                ************ ************************************************

2007-454        TGATTATGTGGAAATGTTTGCTGATAAAGGTGCAGATATTATTTACATTCATCCAGAAGC 1500
2007-0904       TGATTATGTGGAAATGTTTGCTGATAAAGGTGCAGATATTATTTACATTCATCCAGAAGC 1500
EGDe            TGATTATGTGGAAATGTTTGCTGATAAAGGTGCAGATATTATTTACATTCATCCAGAAGC 1499
                ************************************************************

2007-454        TGATTTACATCCAGCTAGAACATTAGACAAAATCCGAAATAAGGGGAAGAAAGTTGGTAT 1560
2007-0904       TGATTTACATCCAGCTAGAACATTAGACAAAATCCGAAATAAGGGGAAGAAAGTTGGTAT 1560
EGDe            TGATTTACATCCAGCTAGAACATTAGACAAAATCCGAAATAAGGGGAAGAAAGTTGGTAT 1559
                ************************************************************

2007-454        TGCGATTAATCCGGGAACGGCGATTGCAACAATTAAAGAACTTTTACCATTAGTTGATAA 1620
2007-0904       TGCGATTAATCCGGGAACGGCGATTGCAACAATTAAAGAACTTTTACCATTAGTTGATAA 1620
EGDe            TGCGATTAATCCGGGAACGGCGATTGCAACAATTAAAGAACTTTTACCATTAGTTGATAA 1619
                ************************************************************

2007-454        TGTGATGGTGATGACAGTTAATCCAGGTTTTGCAGGACAAAAATATTTGGACTATGTAGA 1680
2007-0904       TGTGATGGTGATGACAGTTAATCCAGGTTTTGCAGGACAAAAATATTTGGACTATGTAGA 1680
EGDe            TGTGATGGTGATGACAGTTAATCCAGGTTTTGCAGGACAAAAATATTTGGACTATGTAGA 1679
                ************************************************************

2007-454        TGTGAAAATCGCAGAACTTGTCGAATTGAAAAAGGATTATACTTTTGAAATTATGGTGGA 1740
2007-0904       TGTGAAAATCGCAGAACTTGTCGAATTGAAAAAGGATTATACTTTTGAAATTATGGTGGA 1740
EGDe            TGTGAAAATCGCAGAACTTGTCGAATTGAAAAAGGATTATACTTTTGAAATTATGGTGGA 1739
                ************************************************************

2007-454        TGGGGCAATATCTCCAGAAAAAATCAGCAAACTTTCTCAGTTAGGTGTTAAGGGGTTTGT 1800
2007-0904       TGGGGCAATATCTCCAGAAAAAATCAGCAAACTTTCTCAGTTAGGTGTTAAGGGGTTTGT 1800
EGDe            TGGGGCAATATCTCCAGAAAAAATCAGCAAACTTTCTCAGTTAGGTGTTAAGGGGTTTGT 1799
                ************************************************************

2007-454        GCTAGGAACATCCGCATTATTCGGAAAAAAAGAAAGTTATGCTGAAATAATCAAAAAGTT 1860
2007-0904       GCTAGGAACATCCGCATTATTCGGAAAAAAAGAAAGTTATGCTGAAATAATCAAAAAGTT 1860
EGDe            GCTAGGAACATCCGCATTATTCGGAAAAAAAGAAAGTTATGCTGAAATAATCAAAAAGTT 1859
                ************************************************************
```

SNP8 (non-synonymous/ K→R)

*lmo0736* →

```
2007-454        AAGACAAGAAAAGATGGAGGAATTACGATGAAAATTGCTATTGGAAATGATCATGTTGGA 1920
2007-0904       AAGACAAGAAAAGATGGAGGAATTACGATGAAAATTGCTATTGGAAATGATCATGTTGGA 1920
EGDe            AAAACAAGAAAAGATGGAGGAATTACGATGAAAATTGCTATTGGAAATGATCATGTTGGA 1919
                ** ************************************************************

2007-454        ATCGAGCTTAAGCCAGTTATTGTTGCGTATTTACAGGATTTAGGTCATGAAGTAGATGAT 1980
2007-0904       ATCGAGCTTAAGCCAGTTATTGTTGCGTATTTACAGGATTTAGGTCATGAAGTAGATGAT 1980
EGDe            ATCGAGCTTAAGCCAGTTATTGTTGCGTATTTACAGGATTTAGGTCATGAAGTAGATGAT 1979
                ************************************************************

2007-454        TTTGGTGCTTTCTCAAATGAACGAACTGATTATCCTGAATATGGTAAAAAAGTTGCAGAA 2040
2007-0904       TTTGGTGCTTTCTCAAATGAACGAACTGATTATCCTGAATATGGTAAAAAAGTTGCAGAA 2040
EGDe            TTTGGTGCTTTCTCAAATGAACGAACTGATTATCCTGAATATGGTAAAAAAGTTGCAGAA 2039
                ************************************************************
```


2007-454 AGTGTAGCGGCGGGAAAATCGGACTTAGGAATTTTAATTTGTGGTTCAGGAGTGGGTATT 2100
2007-0904 AGTGTAGCGGCGGGAAAATCGGACTTAGGAATTTTAATTTGTGGTTCAGGAGTGGGTATT 2100
EGDe AGTGTAGCGGCGGGAAAATCGGACTTAGGAATTTTAATTTGTGGTTCAGGAGTGGGTATT 2099
2007-454 TCAATTGCTGCTAATAAAGTAAACGGAATTCGCGCTGTTGTTTGTAGCGAACCATACTCT 2160
2007-0904 TCAATTGCTGCTAATAAAGTAAACGGAATTCGCGCTGTTGTTTGTAGCGAACCATACTCT 2160
EGDe TCAATTGCTGCTAATAAAGTAAACGGAATTCGCGCTGTTGTTTGTAGCGAACCATACTCT 2159
2007-454 GCAAAACTATCACGCGAGCACAATAATACTAATATTTTAGCTTTTGGTTCAAGAGTTGTT 2220
2007-0904 GCAAAACTATCACGCGAGCACAATAATACTAATATTTTAGCTTTTGGTTCAAGAGTTGTT 2220
EGDe GCAAAACTATCACGCGAGCACAATAATACTAATATTTTAGCTTTTGGTTCAAGAGTTGTT 2219
2007-454 GGTGCAGAACTTGCGAAAATGATAGTTCAAAACTGGTTGGATGCAGAATTTGAAGGTGGA 2280
2007-0904 GGTGCAGAACTTGCGAAAATGATAGTTCAAAACTGGTTGGATGCAGAATTTGAAGGTGGA 2280
EGDe GGTGCAGAACTTGCGAAAATGATAGTTCAAAACTGGTTGGATGCAGAATTTGAAGGTGGA 2279
2007-454 AGACATGCGAAACGCGTCGAAATGATTGCTAGAATTGAAGATGAAAATGATTAAGGAGGG 2340
2007-0904 AGACATGCGAAACGCGTCGAAATGATTGCTAGAATTGAAGATGAAAATGATTAAGGAGGG 2340
EGDe AGACATGCGAAACGCGTCGAAATGATTGCTAGAATTGAAGATGAAAATGATTAAGGAGGG 2339
lmo0737
2007-454 GTTGTTATGGAGCAAGAAGTAACACGACTAATTAGTGCTTCTAAACAAGATTTTGAAAAG 2400
2007-0904 GTTGTTATGGAGCAAGAAGTAACACGACTAATTAGTGCTTCTAAACAAGATTTTGAAAAG 2400
EGDe GTTGTTATGGAGCAAGAAGTAACACGACTAATTAGTGCTTCTAAACAAGATTTTGAAAAG 2399
2007-454 ATGAACGGCAGAGACTTAAAACGCTCCATTTTTCAATCAGAAGGTAGAGTAATTATGGCG 2460
2007-0904 ATGAACGGCAGAGACTTAAAACGCTCCATTTTTCAATCAGAAGGTAGAGTAATTATGGCG 2460
EGDe ATGAACGGCAGAGACTTAAAACGCTCCATTTTTCAATCAGAAGGTAGAGTAATTATGGCG 2459
2007-454 CAACATCTTTTATTTGCAAGTCAGGGTCTTGTCCGAGGGGTTACGAATACGGAATTATTA 2520
2007-0904 CAACATCTTTTATTTGCAAGTCAGGGTCTTGTCCGAGGGGTTACGAATACGGAATTATTA 2520
EGDe CAACATCTTTTATTTGCAAGTCAGGGTCTTGTCCGAGGGGTTACGAATACGGAATTATTA 2519
2007-454 GCAGCATTCGGATCGGATATGATTATGCTCAATACATTTAACCTAGAAGATGAAAAAAGT 2580
2007-0904 GCAGCATTCGGATCGGATATGATTATGCTCAATACATTTAACCTAGAAGATGAAAAAAGT 2580
EGDe GCAGCATTCGGATCGGATATGATTATGCTCAATACATTTAACCTAGAAGATGAAAAAAGT 2579
SNP9 (synonymous)
2007-454 AATTTAGGACTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTA 2640
2007-0904 AATTTAGGACTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTA 2640
EGDe AATTTAGGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTA 2639
SNP10 (synonymous)
2007-454 GGAATTTATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCG 2700
2007-0904 GGAATTTATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCG 2700
EGDe GGAATTTATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCA 2699
2007-454 GCAGGAATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTT 2760
2007-0904 GCAGGAATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTT 2760
EGDe GCAGGAATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTT 2759


2007-454 ATTGTTTTAGGTGGAAATCCTGGTTCAGGAACTTCTATTCAAGATATCATTGAAACAACA 2820
2007-0904 ATTGTTTTAGGTGGAAATCCTGGTTCAGGAACTTCTATTCAAGATATCATTGAAACAACA 2820
EGDe ATTGTTTTAGGTGGAAATCCTGGTTCAGGAACTTCTATTCAAGATATCATTGAAACAACA 2819
SNP 11 (synonymous)
2007-454 AAACGAGCACGGAAGTTGCTAGGTAACGATGTCTTAATTTTCGCTGGTAAATGGGAAGAT 2880
2007-0904 AAACGAGCACGGAAGTTGCTAGGTAACGATGTCTTAATTTTCGCTGGTAAATGGGAAGAT 2880
EGDe AAACGAGCACGGAAGTTGCTAGGTAACGATGTCTTAATTTTCGCTGGAAAATGGGAAGAT 2879
2007-454 GGTATTGATGAGAAAGTGTTGGGCGATCCACTTGCTAAACAAGATGCAAAAGAAGTGATT 2940
2007-0904 GGTATTGATGAGAAAGTGTTGGGCGATCCACTTGCTAAACAAGATGCAAAAGAAGTGATT 2940
EGDe GGTATTGATGAGAAAGTGTTGGGCGATCCACTTGCTAAACAAGATGCAAAAGAAGTGATT 2939
2007-454 AAGCAATTGATTGATGCAGGCGCGGATGTGATTGATTTACCTGCTCCAGGATCAAGACAC 3000
2007-0904 AAGCAATTGATTGATGCAGGCGCGGATGTGATTGATTTACCTGCTCCAGGATCAAGACAC 3000
EGDe AAGCAATTGATTGATGCAGGCGCGGATGTGATTGATTTACCTGCTCCAGGATCAAGACAC 2999
SNP12 (non-synonymous/ P→S)
2007-454 GGTATAAGTGTGCGTATGATTCAAGAGTTAGTGCAGTTTATTCATTTATACAAATCAGGT 3060
2007-0904 GGTATAAGTGTGCGTATGATTCAAGAGTTAGTGCAGTTTATTCATTTATACAAATCAGGT 3060
EGDe GGTATAAGTGTGCGTATGATTCAAGAGTTAGTGCAGTTTATTCATTTATACAAACCAGGT 3059
2007-454 ACACTCGCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGT 3120
2007-0904 ACACTCGCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGT 3120
EGDe ACACTCGCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGT 3119
2007-454 TTGATTGCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGT 3180
2007-0904 TTGATTGCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGT 3180
EGDe TTGATTGCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGT 3179
2007-454 TTTTCTGGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAA 3240
2007-0904 TTTTCTGGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAA 3240
EGDe TTTTCTGGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAA 3239
2007-454 CCTTATACGTACTTTAGAATGGCAAGCAGAAATCGTTAAAACTATAAAAAGGAGTGAAAA 3300
2007-0904 CCTTATACGTACTTTAGAATGGCAAGCAGAAATCGTTAAAACTATAAAAAGGAGTGAAAA 3300
EGDe CCTTATACGTACTTTAGAATGGCAAGCAGAAATCGTTAAAACTATAAAAAGGAGTGAAAA 3299
lmo0738
2007-454 AAATGGATTATCAAAAACTAGCTAAAGAGATTCTGGCTAATGTTGGCGGAGAAGAAAATG 3360
2007-0904 AAATGGATTATCAAAAACTAGCTAAAGAGATTCTGGCTAATGTTGGCGGAGAAGAAAATG 3360
EGDe AAATGGATTATCAAAAACTAGCTAAAGAGATTCTGGCTAATGTTGGCGGAGAAGAAAATG 3359
2007-454 TGCGCTCAGTGGTTCACTGTGCTACACGACTTCGCTTCAAATTAGTAAACAAAGAAAAAG 3420
2007-0904 TGCGCTCAGTGGTTCACTGTGCTACACGACTTCGCTTCAAATTAGTAAACAAAGAAAAAG 3420
EGDe TGCGCTCAGTGGTTCACTGTGCTACACGACTTCGCTTCAAATTAGTAAACAAAGAAAAAG 3419
SNP13 (non-synonymous/ A→V)
2007-454 TAGACAAAAAACAAATCGAAAGCATATCGGGAGTTATTAGTGTTGTAGAAAATGCTGGAC 3480
2007-0904 TAGACAAAAAACAAATCGAAAGCATATCGGGAGTTATTAGTGTTGTAGAAAATGCTGGAC 3480
EGDe CAGACAAAAAACAAATCGAAAGCATATCGGGAGTTATTAGTGTTGTAGAAAATGCTGGAC 3479

```
2007-454        AATTACAAGTAATCATTGGCAATACCGTTGGAGATGTATATAAAGCACTGGGTTCTTTTA 3540
2007-0904       AATTACAAGTAATCATTGGCAATACCGTTGGAGATGTATATAAAGCACTGGGTTCTTTTA 3540
EGDe            AATTACAAGTAATCATTGGCAATACCGTTGGAGATGTATATAAAGCACTGGGTTCTTTTA 3539
                ************************************************************

2007-454        CTAAACTAACAGATGATGGCGACTCGGAGATAGCAAAAGGGACAAAAGATTCAGATGGGA 3600
2007-0904       CTAAACTAACAGATGATGGCGACTCGGAGATAGCAAAAGGGACAAAAGATTCAGATGGGA 3600
EGDe            CTAAACTAACAGATGATGGCGACTCGGAGATAGCAAAAGGGACAAAAGATTCAGATGGGA 3599
                ************************************************************
                                          SNP14 (synonymous)
2007-454        ACTTTTTGAGTAAGGCAATTGATGTTATTTCAGGCATATTCACACCTATTTTGGGAGCGC 3660
2007-0904       ACTTTTTGAGTAAGGCAATTGATGTTATTTCAGGCATATTCACACCTATTTTGGGAGCGC 3660
EGDe            ACTTTTTGAGTAAGGCAATTGATGTTATTTCAGGTATATTCACACCTATTTTGGGAGCGC 3659
                ********************************** ************************

2007-454        TTGCAGGTGGTGGTATGCTCAAAGGATTGCTGATGATTCTTACTACTTTTGGATGGTTGA 3720
2007-0904       TTGCAGGTGGTGGTATGCTCAAAGGATTGCTGATGATTCTTACTACTTTTGGATGGTTGA 3720
EGDe            TTGCAGGTGGTGGTATGCTCAAAGGATTGCTGATGATTCTTACTACTTTTGGATGGTTGA 3719
                ************************************************************

2007-454        CCGAATCAAGTGGTACGTATCAGATTTTATACGCGGCTGCTGATAGTGTGTTCTATTTCT 3780
2007-0904       CCGAATCAAGTGGTACGTATCAGATTTTATACGCGGCTGCTGATAGTGTGTTCTATTTCT 3780
EGDe            CCGAATCAAGTGGTACGTATCAGATTTTATACGCGGCTGCTGATAGTGTGTTCTATTTCT 3779
                ************************************************************

2007-454        TACCGCTTATTCTTGCTTATACTGCTGCGAGAAAATTTGGTGCTAATCCTCCGGTCGCGA 3840
2007-0904       TACCGCTTATTCTTGCTTATACTGCTGCGAGAAAATTTGGTGCTAATCCTCCGGTCGCGA 3840
EGDe            TACCGCTTATTCTTGCTTATACTGCTGCGAGAAAATTTGGTGCTAATCCTCCGGTCGCGA 3839
                ************************************************************

2007-454        TTGCAGCAGCAGGTGCACTAGTATATCCTACAATGATTAATCTATTTAATGAGGGAGCTC 3900
2007-0904       TTGCAGCAGCAGGTGCACTAGTATATCCTACAATGATTAATCTATTTAATGAGGGAGCTC 3900
EGDe            TTGCAGCAGCAGGTGCACTAGTATATCCTACAATGATTAATCTATTTAATGAGGGAGCTC 3899
                ************************************************************

2007-454        ATATTACTTTCTTACAAATACCAGTCGTATTAATGAGTTATTCTTTCTCCGTTATTCCAA 3960
2007-0904       ATATTACTTTCTTACAAATACCAGTCGTATTAATGAGTTATTCTTTCTCCGTTATTCCAA 3960
EGDe            ATATTACTTTCTTACAAATACCAGTCGTATTAATGAGTTATTCTTTCTCCGTTATTCCAA 3959
                ************************************************************

2007-454        TTATCTTGGCAGTTTGGTTTTTATCTATTTTAGAAAGATTTCTTAATAGTAAAATTCATG 4020
2007-0904       TTATCTTGGCAGTTTGGTTTTTATCTATTTTAGAAAGATTTCTTAATAGTAAAATTCATG 4020
EGDe            TTATCTTGGCAGTTTGGTTTTTATCTATTTTAGAAAGATTTCTTAATAGTAAAATTCATG 4019
                ************************************************************
                                                        SNP15 (synonymous)
2007-454        AGGCAGCGAAGACCTTTTTAACGCCGATGATTTGCCTGATGCTTATCGTACCACTTACAT 4080
2007-0904       AGGCAGCGAAGACCTTTTTAACGCCGATGATTTGCCTGATGCTTATCGTACCACTTACAT 4080
EGDe            AGGCAGCGAAGACCTTTTTAACGCCGATGATTTGCCTGATGCTTATCGTACCGCTTACAT 4079
                ************************************************** *******
SNP16 (synonymous)
2007-454        TCCTAGCGTTTGGACCTCTGGGTACGTTTATTAGTCAAGGTTTAGCTAGCGGGTATACAT 4140
2007-0904       TCCTAGCGTTTGGACCTCTGGGTACGTTTATTAGTCAAGGTTTAGCTAGCGGGTATACAT 4140
EGDe            TCTTAGCGTTTGGACCTCTGGGTACGTTTATTAGTCAAGGTTTAGCTAGCGGGTATACAT 4139
                ** ************************************************************
                                                    SNP17 (synonymous)
2007-454        TTATTTATAACTTGAGCCCAATTGTAGCGGGTGCTTTTATGGGTGCGTTCTGGCAAGTAT 4200
2007-0904       TTATTTATAACTTGAGCCCAATTGTAGCGGGTGCTTTTATGGGTGCATTCTGGCAAGTAT 4200
EGDe            TTATTTATAACTTGAGCCCAATTGTAGCGGGTGCTTTTATGGGTGCGTTCTGGCAAGTAT 4199
                ******************************************** **************
```

SNP18 (synonymous)

```
2007-454        TAGTTATTTTTGGGATTCATTGGGGCTTTGTACCAATCATGATTAACAACTTATCACGCT 4260
2007-0904       TAGTTATTTTTGGGATTCATTGGGGCTTTGTACCAATCATGATTAACAACTTATCACGCT 4260
EGDe            TAGTTATTTTTGGGATTCATTGGGGCTTTGTGCCAATCATGATTAACAACTTATCACGCT 4259
                ****************************** ****************************

2007-454        ATGGGCGTGACACCATGATTGCGATGGTTGGACCATCAAACTTTGCTCAAGCAGGTGCAT 4320
2007-0904       ATGGGCGTGACACCATGATTGCGATGGTTGGACCATCAAACTTTGCTCAAGCAGGTGCAT 4320
EGDe            ATGGGCGTGACACCATGATTGCGATGGTTGGACCATCAAACTTTGCTCAAGCAGGTGCAT 4319
                ************************************************************

2007-454        CACTTGGAGTATTTTTAAAAACAAAAAAACCGGAAGTCAAAGCTATTGCTGGTTCTGCCG 4380
2007-0904       CACTTGGAGTATTTTTAAAAACAAAAAAACCGGAAGTCAAAGCTATTGCTGGTTCTGCCG 4380
EGDe            CACTTGGAGTATTTTTAAAAACAAAAAAACCGGAAGTCAAAGCTATTGCTGGTTCTGCCG 4379
                ************************************************************

2007-454        CACTTACAGGCTTTTTTGGAATTACAGAACCGTCTATTTACGGGGTTACTTTAAAATACA 4440
2007-0904       CACTTACAGGCTTTTTTGGAATTACAGAACCGTCTATTTACGGGGTTACTTTAAAATACA 4440
EGDe            CACTTACAGGCTTTTTTGGAATTACAGAACCGTCTATTTACGGGGTTACTTTAAAATACA 4439
                ************************************************************

2007-454        AAAAACCATTTGTTATCGCGAGTATCGCCGGTGCAATTGGTGGTGCGATAGTTGGTGCAG 4500
2007-0904       AAAAACCATTTGTTATCGCGAGTATCGCCGGTGCAATTGGTGGTGCGATAGTTGGTGCAG 4500
EGDe            AAAAACCATTTGTTATCGCGAGTATCGCCGGTGCAATTGGTGGTGCGATAGTTGGTGCAG 4499
                ************************************************************

2007-454        CTGGGTCTTCTGGTGCAGCGAACGCTATCCCGGGAATTTTAACATTACCAATTTTTATCG 4560
2007-0904       CTGGGTCTTCTGGTGCAGCGAACGCTATCCCGGGAATTTTAACATTACCAATTTTTATCG 4560
EGDe            CTGGGTCTTCTGGTGCAGCGAACGCTATCCCGGGAATTTTAACATTACCAATTTTTATCG 4559
                ************************************************************

2007-454        GAAAAGGATTTGTTGGCTTTATTCTCGGAATTGCTGTAGCCTATATTTTATCAGCAATCG 4620
2007-0904       GAAAAGGATTTGTTGGCTTTATTCTCGGAATTGCTGTAGCCTATATTTTATCAGCAATCG 4620
EGDe            GAAAAGGATTTGTTGGCTTTATTCTCGGAATTGCTGTAGCCTATATTTTATCAGCAATCG 4619
                ************************************************************

2007-454        GCACATACTTCTTCGGTTATAAAGACGAGATGGCAGACGGAATTGCTCCAACTACAAAAG 4680
2007-0904       GCACATACTTCTTCGGTTATAAAGACGAGATGGCAGACGGAATTGCTCCAACTACAAAAG 4680
EGDe            GCACATACTTCTTCGGTTATAAAGACGAGATGGCAGACGGAATTGCTCCAACTACAAAAG 4679
                ************************************************************
```

SNP19 (non-synonymous/ V→I)

```
2007-454        AAGCGAAAGAAACTGGCGTCGAAGCAGAAATTATTGTAAGCCCTATTCGCGGCAATATTG 4740
2007-0904       AAGCGAAAGAAACTGGCGTCGAAGCAGAAATTATTGTAAGCCCTATTCGCGGCAATATTG 4740
EGDe            AAGCGAAAGAAACTGGCGTCGAAGCAGAAGTTATTGTAAGCCCTATTCGCGGCAATATTG 4739
                ****************************** ****************************
```

SNP20 (non-synonymous/ N→S)

```
2007-454        TTCCGTTAAGTGAGGTAAAAGATGAAGCTTTTTCAGCAGGGTTACTTGGAAAAGGTGTAG 4800
2007-0904       TTCCGTTAAGTGAGGTAAAAGATGAAGCTTTTTCAGCAGGGTTACTTGGAAAAGGTGTAG 4800
EGDe            TTCCGTTAAATGAGGTAAAAGATGAAGCTTTTTCAGCAGGGTTACTTGGAAAAGGTGTAG 4799
                ********* ************************************************
```

SNP21 (synonymous)

```
2007-454        CAATCGTTCCACAAGAAGGTAAATTAATCTCACCAGTAAACGGCACTATCGAAACAGCTT 4860
2007-0904       CAATCGTTCCACAAGAAGGTAAATTAATCTCACCAGTAAACGGCACTATCGAAACAGCTT 4860
EGDe            CAATCGTTCCGCAAGAAGGTAAATTAATCTCACCAGTAAACGGCACTATCGAAACAGCTT 4859
                ********** ************************************************
```

SNP22 (non-synonymous/ L→I)

```
2007-454        TTCCAACAGGACATGCGATTGGAATTCGTTCTGATAAAGGGGTAGAAATTTTGATTCACG 4920
2007-0904       TTCCAACAGGACATGCGATTGGAATTCGTTCTGATAAAGGGGTAGAAATTTTGATTCACG 4920
EGDe            TTCCAACAGGACATGCGATTGGAATTCGTTCTGATAAAGGGGTAGAAATTTTGCTTCACG 4919
                ************************************************** ******

2007-454        TTGGCTTTGATACTGTTCAATTAAATGGGAAATATTTTAAATTACTTGTGGCTCAAGGTG 4980
2007-0904       TTGGCTTTGATACTGTTCAATTAAATGGGAAATATTTTAAATTACTTGTGGCTCAAGGTG 4980
EGDe            TTGGCTTTGATACTGTTCAATTAAATGGGAAATATTTTAAATTACTTGTGGCTCAAGGTG 4979
                ************************************************************
```

SNP23 (synonymous)

```
2007-454        ATAGAGTTTTGGTAGGGCAAGCGTTACTTGAATTTGACCTAGAAGCAATTAAGGCAGATG 5040
2007-0904       ATAGAGTTTTGGTAGGGCAAGCGTTACTTGAATTTGACCTAGAAGCAATTAAGGCAGATG 5040
EGDe            ATAGAGTTTTGGTAGGGCAAGCATTACTTGAATTTGACCTAGAAGCAATTAAGGCAGATG 5039
                ********************* ************************************

2007-454        GTTATGATATTACGACTCCGATAGTTGTAACTAATACGGATGCTTACTTAGATGTACTCA 5100
2007-0904       GTTATGATATTACGACTCCGATAGTTGTAACTAATACGGATGCTTACTTAGATGTACTCA 5100
EGDe            GTTATGATATTACGACTCCGATAGTTGTAACTAATACGGATGCTTACTTAGATGTACTCA 5099
                ************************************************************
```

SNP24 (synonymous)

```
2007-454        TCTCGGATCAAAAAACTGTTAATTATGAAGATACGCTATTAACACCAGTACTTTAATATC 5160
2007-0904       TCTCGGATCAAAAAACTGTTAATTATGAAGATACGCTATTAACACCAGTACTTTAATATC 5160
EGDe            TCTCGGATCAAAAAACTGTTAATTATGAAGATACGCTTTTAACACCAGTACTTTAATATC 5159
                ************************************ ********************
```

*lmo0739* →

```
2007-454        GAAGGAGGATAAAAAATGTTTAATTTGCAAAAAGGTTTTCCAGAAAATTTTAAATGGGGT 5220
2007-0904       GAAGGAGGATAAAAAATGTTTAATTTGCAAAAAGGTTTTCCAGAAAATTTTAAATGGGGT 5220
EGDe            GAAGGAGGATAAAAAATGTTTAATTTGCAAAAAGGTTTTCCAGAAAATTTTAAATGGGGT 5219
                ************************************************************

2007-454        AGCTCCACTAATGCGCAACAATTTGAAGGTGGCTATAAAGAAGGTGGAAAAGGGCTTTCG 5280
2007-0904       AGCTCCACTAATGCGCAACAATTTGAAGGTGGCTATAAAGAAGGTGGAAAAGGGCTTTCG 5280
EGDe            AGCTCCACTAATGCGCAACAATTTGAAGGTGGCTATAAAGAAGGTGGAAAAGGGCTTTCG 5279
                ************************************************************

2007-454        ATTGCGGATGTGAGGGTTATTCCAGATATGCCAGACGAATCTGATTTTGAATCGTTCAAA 5340
2007-0904       ATTGCGGATGTGAGGGTTATTCCAGATATGCCAGACGAATCTGATTTTGAATCGTTCAAA 5340
EGDe            ATTGCGGATGTGAGGGTTATTCCAGATATGCCAGACGAATCTGATTTTGAATCGTTCAAA 5339
                ************************************************************

2007-454        ACGGCTTCAGACCATTATCATCATTATAAAGAGGATATTGCTTATTACGGGGAAATGGGC 5400
2007-0904       ACGGCTTCAGACCATTATCATCATTATAAAGAGGATATTGCTTATTACGGGGAAATGGGC 5400
EGDe            ACGGCTTCAGACCATTATCATCATTATAAAGAGGATATTGCTTATTACGGGGAAATGGGC 5399
                ************************************************************
```

SNP25 (synonymous)

```
2007-454        TTCCAAATTTATCGTTTTACAATGGCTTGGTCGAGAATTTTCCCGAACGGTGACGAAACA 5460
2007-0904       TTCCAAATTTATCGTTTTACAATGGCTTGGTCGAGAATTTTCCCGAACGGTGACGAAACA 5460
EGDe            TTCCAAATTTATCGTTTTACAATGGCTTGGTCGAGAATTTTCCCGAACGGAGACGAAACA 5459
                ************************************************ *********

2007-454        GAACCAAACGATGCTGGAGTTGAGTTTTATAGTAATATGTTGGCAGAGTTAGAAAAATAT 5520
2007-0904       GAACCAAACGATGCTGGAGTTGAGTTTTATAGTAATATGTTGGCAGAGTTAGAAAAATAT 5520
EGDe            GAACCAAACGATGCTGGAGTTGAGTTTTATAGTAATATGTTGGCAGAGTTAGAAAAATAT 5519
                ************************************************************
```

SNP26 (non-synonymous/ E→G)

```
2007-454        AATATCGAACCAGTTGTTACTTTATATGCTTATGATATGCCGTTACAACTGCTTGGAAAA 5580
2007-0904       AATATCGAACCAGTTGTTACTTTATATGCTTATGATATGCCGTTACAACTGCTTGGAAAA 5580
EGDe            AATATCGAACCAGTTGTTACTTTATATGCTTATGATATGCCGTTACAACTGCTTGAAAAA 5579
                ********************************************************** ****

2007-454        TATAATGGTTGGCTAGATAGAGCAATTATCAAAGATTACCTACATTATGTAGAAACGGTC 5640
2007-0904       TATAATGGTTGGCTAGATAGAGCAATTATCAAAGATTACCTACATTATGTAGAAACGGTC 5640
EGDe            TATAATGGTTGGCTAGATAGAGCAATTATCAAAGATTACCTACATTATGTAGAAACGGTC 5639
                ************************************************************

2007-454        GTGAAATTATTTAAAGGTCGAGTGAAATATTGGGTTCCTTTCAACGAGCAAAACTTTATT 5700
2007-0904       GTGAAATTATTTAAAGGTCGAGTGAAATATTGGGTTCCTTTCAACGAGCAAAACTTTATT 5700
EGDe            GTGAAATTATTTAAAGGTCGAGTGAAATATTGGGTTCCTTTCAACGAGCAAAACTTTATT 5699
                ************************************************************

2007-454        TCGATTGATTCTGAATACATGAGTGGTTATCGTGCTAAAAATAAAGCAGAAGTTTTCCAA 5760
2007-0904       TCGATTGATTCTGAATACATGAGTGGTTATCGTGCTAAAAATAAAGCAGAAGTTTTCCAA 5760
EGDe            TCGATTGATTCTGAATACATGAGTGGTTATCGTGCTAAAAATAAAGCAGAAGTTTTCCAA 5759
                ************************************************************

2007-454        ATTCAACATCATTTTAATTTGTGTTATGCAGAAGCAACGAAGCTTGTACACCAAATTGAT 5820
2007-0904       ATTCAACATCATTTTAATTTGTGTTATGCAGAAGCAACGAAGCTTGTACACCAAATTGAT 5820
EGDe            ATTCAACATCATTTTAATTTGTGTTATGCAGAAGCAACGAAGCTTGTACACCAAATTGAT 5819
                ************************************************************

2007-454        CCAGATGCAAAAGTTGGTGGAAATATTGGTAACATTTGCCCTTACCCAATGACTTGCAAA 5880
2007-0904       CCAGATGCAAAAGTTGGTGGAAATATTGGTAACATTTGCCCTTACCCAATGACTTGCAAA 5880
EGDe            CCAGATGCAAAAGTTGGTGGAAATATTGGTAACATTTGCCCTTACCCAATGACTTGCAAA 5879
                ************************************************************

2007-454        CCCGAAGATGTCGAAGCAAGTGACAAAGTAGCGCAACAGTTAGGTTATGCTTATGGAGAT 5940
2007-0904       CCCGAAGATGTCGAAGCAAGTGACAAAGTAGCGCAACAGTTAGGTTATGCTTATGGAGAT 5940
EGDe            CCCGAAGATGTCGAAGCAAGTGACAAAGTAGCGCAACAGTTAGGTTATGCTTATGGAGAT 5939
                ************************************************************

2007-454        ATTTATTTCAGAGGATACTATCCAAAATACTTCCTTAAAGAATATGAAGGCGTAGATTTT 6000
2007-0904       ATTTATTTCAGAGGATACTATCCAAAATACTTCCTTAAAGAATATGAAGGCGTAGATTTT 6000
EGDe            ATTTATTTCAGAGGATACTATCCAAAATACTTCCTTAAAGAATATGAAGGCGTAGATTTT 5999
                ************************************************************

2007-454        GAGCAAATCATTTTAGATGATGACTTAACTATTATCAAAAGTTCTGAGCCAGACTTCATG 6060
2007-0904       GAGCAAATCATTTTAGATGATGACTTAACTATTATCAAAAGTTCTGAGCCAGACTTCATG 6060
EGDe            GAGCAAATCATTTTAGATGATGACTTAACTATTATCAAAAGTTCTGAGCCAGACTTCATG 6059
                ************************************************************

2007-454        TCGTTAACATACTATATGTCGAGTGCGATTGAAGCAAAAGGTGAAGAGGAAGTTGTGGTA 6120
2007-0904       TCGTTAACATACTATATGTCGAGTGCGATTGAAGCAAAAGGTGAAGAGGAAGTTGTGGTA 6120
EGDe            TCGTTAACATACTATATGTCGAGTGCGATTGAAGCAAAAGGTGAAGAGGAAGTTGTGGTA 6119
                ************************************************************

2007-454        ATGAACGGGATCAAAGCGCCAAACCCATATTGTGAAACAACTGAATGGGGTTGGACAATT 6180
2007-0904       ATGAACGGGATCAAAGCGCCAAACCCATATTGTGAAACAACTGAATGGGGTTGGACAATT 6180
EGDe            ATGAACGGGATCAAAGCGCCAAACCCATATTGTGAAACAACTGAATGGGGTTGGACAATT 6179
                ************************************************************
```

```
2007-454        GATCCATATGGTTTTAAACATTATTTGCAAGAGTTCTATCATCGCTATCAATTACCAATT 6240
2007-0904       GATCCATATGGTTTTAAACATTATTTGCAAGAGTTCTATCATCGCTATCAATTACCAATT 6240
EGDe            GATCCATATGGTTTTAAACATTATTTGCAAGAGTTCTATCATCGCTATCAATTACCAATT 6239
                ************************************************************

2007-454        CTAATTCTGGAGAATGGGATGGGTGCACGCGACGAAAAAAATACCGATGATACAATTGAT 6300
2007-0904       CTAATTCTGGAGAATGGGATGGGTGCACGCGACGAAAAAAATACCGATGATACAATTGAT 6300
EGDe            CTAATTCTGGAGAATGGGATGGGTGCACGCGACGAAAAAAATACCGATGATACAATTGAT 6299
                ************************************************************
                          SNP27 (synonymous)                SNP28 (synonymous)
2007-454        GATACGTACAGAATTGATTATTTGGCTTCACATATTGCGCGTATGCAAGAAGCAGTTGAA 6360
2007-0904       GATACGTACAGAATTGATTATTTGGCTTCACATATTGCGCGTATGCAAGAAGCAGTTGAA 6360
EGDe            GATACGTACCGAATTGATTATTTGGCTTCACATATTGCGCGTATGCAAGAAGCTGTTGAA 6359
                ********* ****************************************** ******

2007-454        GAGGGGTGCGAAATTATTGGTTATCTTACATGGTCCGCAACAGATCTTTATTCAACGCGT 6420
2007-0904       GAGGGGTGCGAAATTATTGGTTATCTTACATGGTCCGCAACAGATCTTTATTCAACGCGT 6420
EGDe            GAGGGGTGCGAAATTATTGGTTATCTTACATGGTCCGCAACAGATCTTTATTCAACGCGT 6419
                ************************************************************
                                                  SNP29 (synonymous)
2007-454        GAAGGTTTCGAGAAAAGATATGGCTTTGTTTATGTAGATAAAGATAATAGCTATAAACGT 6480
2007-0904       GAAGGTTTCGAGAAAAGATATGGCTTTGTTTATGTAGATAAAGATAATAGCTATAAACGT 6480
EGDe            GAAGGTTTCGAGAAAAGATATGGCTTTGTTTATGTTGATAAAGATAATAGCTATAAACGT 6479
                ********************************** ************************

2007-454        TTGAAGAAGAAAAGTTTCTATTGGTATAAAAAAGTAATTGAAACTAATGGAAATGATTTA 6540
2007-0904       TTGAAGAAGAAAAGTTTCTATTGGTATAAAAAAGTAATTGAAACTAATGGAAATGATTTA 6540
EGDe            TTGAAGAAGAAAAGTTTCTATTGGTATAAAAAAGTAATTGAAACTAATGGAAATGATTTA 6539
                ************************************************************
                SNP30 (non-synonymous/ N→S)
2007-454        AGTTATTAA 6549
2007-0904       AGTTATTAA 6549
EGDe            AATTATTAA 6548
                * ******
```

**C.**

```
10M127          ------------------------------------------------------------
10M198          ------------------------------------------------------------
10M130          ------------------------------------------------------------
10M138A         ------------------------------------------------------------
2007-454        ATGGAGCAAGAAGTAACACGACTAATTAGTGCTTCTAAACAAGATTTTGAAAAGATGAAC 60
2007-0904       ATGGAGCAAGAAGTAACACGACTAATTAGTGCTTCTAAACAAGATTTTGAAAAGATGAAC 60
10M131 (1/2a)   ------------------------------------------------------------
09M3788 (1/2c)  ------------------------------------------------------------
EGDe (1/2a)     ATGGAGCAAGAAGTAACACGACTAATTAGTGCTTCTAAACAAGATTTTGAAAAGATGAAC 60


10M127          ------------------------------------------------------------
10M198          ------------------------------------------------------------
10M130          ------------------------------------------------------------
10M138A         ------------------------------------------------------------
2007-454        GGCAGAGACTTAAAACGCTCCATTTTTCAATCAGAAGGTAGAGTAATTATGGCGCAACAT 120
2007-0904       GGCAGAGACTTAAAACGCTCCATTTTTCAATCAGAAGGTAGAGTAATTATGGCGCAACAT 120
10M131 (1/2a)   ------------------------------------------------------------
09M3788 (1/2c)  ------------------------------------------------------------
EGDe (1/2a)     GGCAGAGACTTAAAACGCTCCATTTTTCAATCAGAAGGTAGAGTAATTATGGCGCAACAT 120
```

```
10M127          ------------------------------------------------------------
10M198          ------------------------------------------------------------
10M130          ------------------------------------------------------------
10M138A         ------------------------------------------------------------
2007-454        CTTTTATTTGCAAGTCAGGGTCTTGTCCGAGGGGTTACGAATACGGAATTATTAGCAGCA 180
2007-0904       CTTTTATTTGCAAGTCAGGGTCTTGTCCGAGGGGTTACGAATACGGAATTATTAGCAGCA 180
10M131 (1/2a)   ------------------------------------------------------------
09M3788 (1/2c)  ------------------------------------------------------------
EGDe (1/2a)     CTTTTATTTGCAAGTCAGGGTCTTGTCCGAGGGGTTACGAATACGGAATTATTAGCAGCA 180


10M127          ------------------------------------------------------------A 1
10M198          ------------------------------------------------------------A 1
10M130          ------------------------------------------------------------A 1
10M138A         ------------------------------------------------------------A 1
2007-454        TTCGGATCGGATATGATTATGCTCAATACATTTAACCTAGAAGATGAAAAAAGTAATTTA 240
2007-0904       TTCGGATCGGATATGATTATGCTCAATACATTTAACCTAGAAGATGAAAAAAGTAATTTA 240
10M131 (1/2a)   ------------------------------------------------------------A 1
09M3788 (1/2c)  ------------------------------------------------------------A 1
EGDe (1/2a)     TTCGGATCGGATATGATTATGCTCAATACATTTAACCTAGAAGATGAAAAAAGTAATTTA 240
                                                                            *
```

SNP 1 (synonymous)

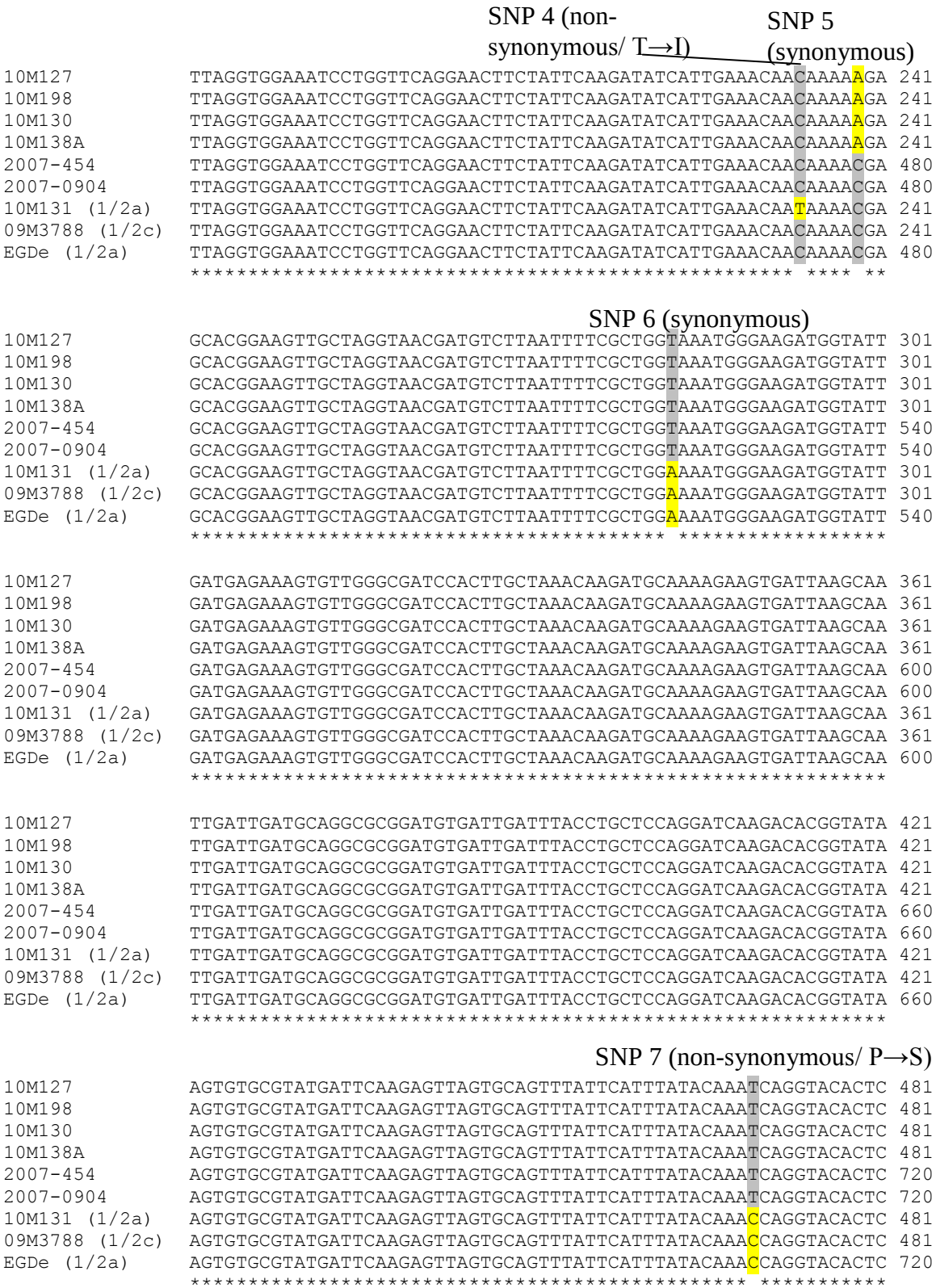
```
10M127          GGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 61
10M198          GGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 61
10M130          GGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 61
10M138A         GGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 61
2007-454        GGACTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 300
2007-0904       GGACTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 300
10M131 (1/2a)   GGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 61
09M3788 (1/2c)  GGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 61
EGDe (1/2a)     GGGCTTCAAGGACTTACCCTCGAAGATTTGAAAAAACGCGTGAACATCCCTTTAGGAATT 300
                ** ************************************************************
```

SNP 2 (synonymous)

```
10M127          TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCGGCAGGA 121
10M198          TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCGGCAGGA 121
10M130          TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCGGCAGGA 121
10M138A         TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCGGCAGGA 121
2007-454        TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCGGCAGGA 360
2007-0904       TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCGGCAGGA 360
10M131 (1/2a)   TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCAGCAGGA 121
09M3788 (1/2c)  TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCAGCAGGA 121
EGDe (1/2a)     TATCTTGGTTGTCCAGGTGAAAATAAGACTTCAGAAAACATTATTTATGATGCAGCAGGA 360
                ************************************************ ******
```

SNP 3 (non-synonymous/ T→A)

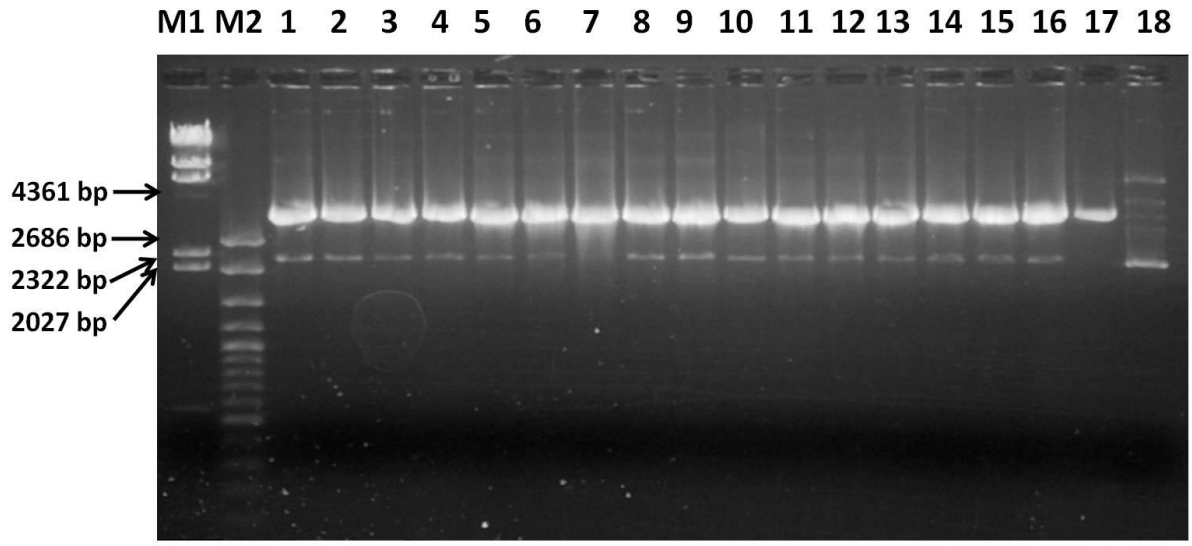
```
10M127          ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 181
10M198          ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 181
10M130          ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 181
10M138A         ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 181
2007-454        ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 420
2007-0904       ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 420
10M131 (1/2a)   ATGTTAGCTGCCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 181
09M3788 (1/2c)  ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 181
EGDe (1/2a)     ATGTTAGCTACCGATGAACATTTATTAAGAGCAAAAGAAATTGGCGCAGATTTTATTGTT 420
                ********* ************************************************
```


SNP 4 (non-synonymous/ T→I)
SNP 5 (synonymous)
SNP 6 (synonymous)
SNP 7 (non-synonymous/ P→S)

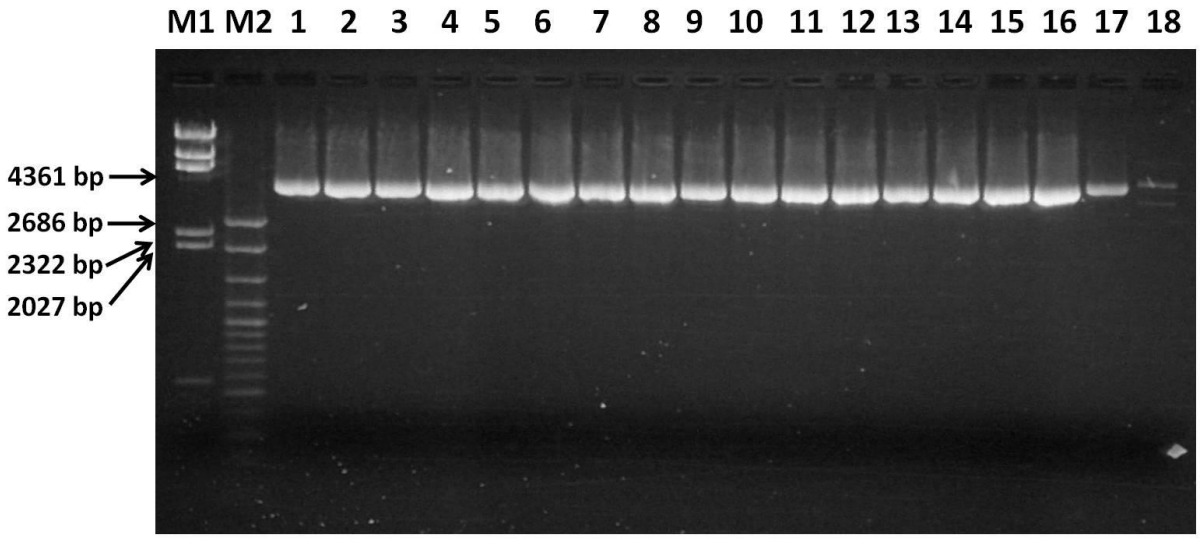
```
10M127          GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 541
10M198          GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 541
10M130          GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 541
10M138A         GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 541
2007-454        GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 780
2007-0904       GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 780
10M131 (1/2a)   GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 541
09M3788 (1/2c)  GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 541
EGDe (1/2a)     GCAATGACATTCTTAAATAGTTCAGTTGAAGGGGCGGATCAAGATACAATTCGTTTGATT 780
                ************************************************************

10M127          GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 601
10M198          GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 601
10M130          GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 601
10M138A         GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 601
2007-454        GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 840
2007-0904       GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 840
10M131 (1/2a)   GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 601
09M3788 (1/2c)  GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 601
EGDe (1/2a)     GCATTGATGATGAAAGAAACCGGTGCCGATATTCACGCGATTGGTGATGGTGGTTTTTCT 840
                ************************************************************

10M127          GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 661
10M198          GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 661
10M130          GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 661
10M138A         GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 661
2007-454        GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 900
2007-0904       GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 900
10M131 (1/2a)   GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 661
09M3788 (1/2c)  GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 661
EGDe (1/2a)     GGCTGTACGACACCGGAAAATGTTATGCAATTATCTATTTCACTTAAAGGAAAACCTTAT 900
                ************************************************************

10M127          ACGTACTTTAGAATGGCAAGCAGAAATCGT--- 691
10M198          ACGTACTTTAGAATGGCAAGCAGAAATCGT--- 691
10M130          ACGTACTTTAGAATGGCAAGCAGAAATCGT--- 691
10M138A         ACGTACTTTAGAATGGCAAGCAGAAATCGT--- 691
2007-454        ACGTACTTTAGAATGGCAAGCAGAAATCGTTAA 933
2007-0904       ACGTACTTTAGAATGGCAAGCAGAAATCGTTAA 933
10M131 (1/2a)   ACGTACTTTAGAATGGCAAGCAGAAATCGT--- 691
09M3788 (1/2c)  ACGTACTTTAGAATGGCAAGCAGAAATCGT--- 691
EGDe (1/2a)     ACGTACTTTAGAATGGCAAGCAGAAATCGTTAA 933
                ****************************
```

**Figure 2.5 PCR amplicons from primers annealing flanking genes and *lmo0737*.** (A) PCR products from primers lmo0733F and lmo0737R. Lane 1, 2007-618; lane 2, J3026; lanes 3, J3053; lanes 4, J3195; lane 5, J4458; lane 6, J4490; lane 7, J4953; lane 8, 18-2a; lane 9, 128b-1; lane 10, 2001-75R; lane 11, 2001-8R; lane 12, 2008-296; lane 13, 2007-454; lane 14, 2008-894; lane 15, 363b-1; lane 16, 491a-5; lane 17, EGDe (positive control); and lane 18, F2365 (negative control). (B) PCR products using primers lmo0737F and lmo0740R. Lane 1, J3026; lanes 2, J3053; lanes 3, J3195; lane 4, J4458; lane 5, J4490; lane 6, J4953; lane 7, 18-2a; lane 8, 128b-1; lane 9, 2001-75R; lane 10, 2001-8R; lane 11, 2008-296; lane 12, 2007-454; lane 13, 2008-894; lane 14, 363b-1; lane 15, 491a-5; lane 16, J5031; lane 17, EGDe (positive control); and lane 18, F2365 (negative control). M1, DNA Molecular Weight Marker II (Roche). M2, exACTCGene cloning DNA ladder (Fisher Scientific). Similar results were observed from other atypical serotype 4b isolates.

**A.**


M1 M2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
4361 bp
2686 bp
2322 bp
2027 bp

**B.**


M1 M2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
4361 bp
2686 bp
2322 bp
2027 bp

## REFERENCES


1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403-410.
2. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics. **21:**3422-3423.
3. **Cheng, Y., J. W. Kim, S. Lee, R. M. Siletzky, and S. Kathariou.** 2010. DNA probes for unambiguous identification of *Listeria monocytogenes* epidemic clone II strains. Appl. Environ. Microbiol. **76:**3061-3068.
4. **Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, and J. D. Thompson.** 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. **31:**3497-3500.
5. **den Bakker, H. C., X. Didelot, E. D. Fortes, K. K. Nightingale, and M. Wiedmann.** 2008. Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. BMC Evol. Biol. **8:**277.
6. **Doumith, M., C. Buchrieser, P. Glaser, C. Jacquet, and P. Martin.** 2004. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. J. Clin. Microbiol. **42:**3819-3822.
7. **Doumith, M., C. Cazalet, N. Simoes, L. Frangeul, C. Jacquet, F. Kunst, P. Martin, P. Cossart, P. Glaser, and C. Buchrieser.** 2004. New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. Infect. Immun. **72:**1072-1083.

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

9. **Eifert, J. D., P. A. Curtis, M. C. Bazaco, R. J. Meinersmann, M. E. Berrang, S. Kernodle, C. Stam, L. A. Jaykus, and S. Kathariou.** 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog. Dis. **2:**192-200.

10. **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:**2383-2390.

11. **Gasanov, U., D. Hughes, and P. M. Hansbro.** 2005. Methods for the isolation and identification of *Listeria* spp. and *Listeria monocytogenes*: a review. FEMS Microbiol. Rev. **29:**851-875.

12. **Graves, L. M., and B. Swaminathan.** 2001. PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. Int. J. Food Microbiol. **65:**55-62.

13. **Huang, B., N. Fang, K. Dimovski, X. Wang, G. Hogg, and J. Bates.** 2011. Observation of a new pattern in serogroup-related PCR typing of *Listeria monocytogenes* 4b isolates. J. Clin. Microbiol. **49:**426-429.

14. **Kathariou, S.** 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J. Food Prot. **65:**1811-1829.

15. **Kathariou, S., L. Graves, C. Buchrieser, P. Glaser, R. M. Siletzky, and B. Swaminathan.** 2006. Involvement of closely related strains of a new clonal group of *Listeria monocytogenes* in the 1998-99 and 2002 multistate outbreaks of foodborne listeriosis in the United States. Foodborne Pathog. Dis. **3:**292-302.

16. **Kibbe, W. A.** 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res. **35:**W43-6.

17. **MacDonald, P. D., R. E. Whitwam, J. D. Boggs, J. N. MacCormack, K. L. Anderson, J. W. Reardon, J. R. Saah, L. M. Graves, S. B. Hunter, and J. Sobel.** 2005. Outbreak of listeriosis among Mexican immigrants as a result of consumption of illicitly produced Mexican-style cheese. Clin. Infect. Dis. **40:**677-682.

18. **Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng, and S. H. Bryant.** 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. **39:**D225-9.

19. **McClelland, M., M. Nelson, and E. Raschke.** 1994. Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases. Nucleic Acids Res. **22:**3640-3659.

20. **Meinersmann, R. J., R. W. Phillips, M. Wiedmann, and M. E. Berrang.** 2004. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes

reveals clonal and recombination histories of three lineages. Appl. Environ. Microbiol. **70:**2193-2203.

21. **Milillo, S. R., J. M. Badamo, and M. Wiedmann.** 2009. Contributions to selected phenotypic characteristics of large species- and lineage-specific genomic regions in *Listeria monocytogenes*. Food Microbiol. **26:**212-223.
22. **Milillo, S. R., J. M. Badamo, and M. Wiedmann.** 2009. Contributions to selected phenotypic characteristics of large species- and lineage-specific genomic regions in *Listeria monocytogenes*. Food Microbiol. **26:**212-223.
23. **Orsi, R. H., Q. Sun, and M. Wiedmann.** 2008. Genome-wide analyses reveal lineage specific contributions of positive selection and recombination to the evolution of *Listeria monocytogenes*. BMC Evol. Biol. **8:**233.
24. **Ragon, M., T. Wirth, F. Hollandt, R. Lavenir, M. Lecuit, A. Le Monnier, and S. Brisse.** 2008. A new perspective on *Listeria monocytogenes* evolution. PLoS Pathog. **4:**e1000146.
25. **Rice, P., I. Longden, and A. Bleasby.** 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. **16:**276-277.
26. **Roberts, A., K. Nightingale, G. Jeffers, E. Fortes, J. M. Kongo, and M. Wiedmann.** 2006. Genetic and phenotypic characterization of *Listeria monocytogenes* lineage III. Microbiology. **152:**685-693.
27. **Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin.** 2011. Foodborne illness acquired in the United States-major pathogens. Emerg. Infect. Dis. **17:**7-15.

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

29. **Toledo-Arana, A., O. Dussurget, G. Nikitas, N. Sesto, H. Guet-Revillet, D. Balestrino, E. Loh, J. Gripenland, T. Tiensuu, K. Vaitkevicius, M. Barthelemy, M. Vergassola, M. A. Nahori, G. Soubigou, B. Regnault, J. Y. Coppee, M. Lecuit, J. Johansson, and P. Cossart.** 2009. The *Listeria* transcriptional landscape from saprophytism to virulence. Nature. **459:**950-956.

30. **Ward, T. J., T. F. Ducey, T. Usgaard, K. A. Dunn, and J. P. Bielawski.** 2008. Multilocus genotyping assays for single nucleotide polymorphism-based subtyping of *Listeria monocytogenes* isolates. Appl. Environ. Microbiol. **74:**7629-7642.

31. **Ward, T. J., L. Gorski, M. K. Borucki, R. E. Mandrell, J. Hutchins, and K. Pupedis.** 2004. Intraspecific phylogeny and lineage group identification based on the *prfA* virulence gene cluster of *Listeria monocytogenes*. J. Bacteriol. **186:**4994-5002.

32. **Ward, T. J., T. Usgaard, and P. Evans.** 2010. A targeted multilocus genotyping assay for lineage, serogroup, and epidemic clone typing of *Listeria monocytogenes*. Appl. Environ. Microbiol. **76:**6680-6684.

33. **Yildirim, S., D. Elhanafi, W. Lin, A. D. Hitchins, R. M. Siletzky, and S. Kathariou.** 2010. Conservation of genomic localization and sequence content of Sau3AI-like restriction-modification gene cassettes among *Listeria monocytogenes* epidemic clone I and selected strains of serotype 1/2a. Appl. Environ. Microbiol. **76:**5577-5584.

34. **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:**4158-4164.

## Chapter 3 Genetic Diversity of the Immigration Control Region of *Listeria monocytogenes*


### ABSTRACT

*Listeria monocytogenes* is the etiological agent of a severe foodborne disease (listeriosis) for humans and animals. An epidemic clonal group of serotype 4b *L. monocytogenes* (epidemic clone I [ECI]) possesses genomic DNA resistant to Sau3AI due to a Sau3AI-like restriction-modification (RM) system. This RM system was also identified in the genomically equivalent location (Region 85) in certain serotype 1/2a and 4a strains. In this study, we showed through *in silico* analysis and sequencing of Region 85 in selected strains that this region harbored multiple RM systems and a group of related lipoprotein genes. DNA probes were designed based on these genes and hybridized with *L. monocytogenes* strains of various serotypes and sources. Our findings suggest that a family of Type I RM systems with various specificities reside in Region 85, predominantly among serotype 4b isolates. Most of the isolates with Type I RM systems harbored the Type IV RM systems *mcrB* and *mrr*; however, the epidemic-associated clonal group ECIa was characteristically missing *mcrB* among serotype 4b, suggesting that this feature might be utilized as a genetic marker for this clonal group. Many serotype 1/2a, 1/2b and 1/2c isolates harbored *mcrB* and/or *mrr* in Region 85, but Type I RM systems were rarely found among isolates of these serotypes. The putative lipoprotein gene in this region (*LMOf2365_0329*) was found to be prevalent among *L. monocytogenes* isolates but was under-represented among animal isolates and lineage III and IV isolates. Low GC contents and the analysis of

the flanking regions indicate that the Type I and IV RM systems as well as the Sau3AI-like RM system might have been acquired as one fragment by an ancestral strain through horizontal gene transfer meditated by site-specific recombination. Diversity in Region 85 seems to result primarily from additions and serial deletions of the RM systems, and deletion or diversification of lipoprotein genes. The pronounced genetic diversity and abundance of RM systems suggest that Region 85 might be an Immigration Control Region of *L. monocytogenes* with potentially important roles in phage resistance, evolution and stress responses of this pathogen.

## INTRODUCTION

*Listeria monocytogenes* is a Gram-positive foodborne pathogen that is responsible for a potentially fatal disease (listeriosis) in humans and animals (44, 48). Human listeriosis occurs primarily among infants, the elderly, pregnant women and immunocompromised individuals, resulting in septicemia, meningitis, stillbirths and abortions (40, 44, 48). In the United States, listeriosis is estimated to cause ca. 1600 cases and 250 deaths in a year, exhibiting a mortality rate of approximately 16% (36). Since *L. monocytogenes* is ubiquitously present in the environment including food processing facilities, the elimination of this bacterium has been a considerable challenge for food manufacturers and food safety regulators (20, 44, 48). To facilitate the control of *L. monocytogenes*, a mixture of bacteriophages was approved in 2006 for use in foods as a biocontrol agent against *L. monocytogenes* (43).

The population of *L. monocytogenes* consists of three major lineages (lineages I, II, III) that include different serotypes (5, 20). Lineage I (serotypes 4b, 1/2b and 3b) exhibits pronounced clonality and a small number of serotype 4b clonal groups have often been associated with listeriosis outbreaks (5). On the other hand, lineage II (serotypes 1/2a, 3a, 1/2c and 3c) and lineage III (serotypes 4a, 4c and certain isolates of serotype 4b) are genetically heterogeneous and overrepresented among environmental and animal isolates, respectively (9, 31, 35). A subset of lineage III strains has been assigned into a separate lineage (lineage IV) (45).

Restriction-modification (RM) systems typically cleave DNA sequences lacking proper methylation while methylating the genomic DNA for protection against restriction

(42). These systems have been divided into four types (Types I to IV) that differ in their components, substrates and the relationship between recognition sites and cleavage sites (29, 42). Type I RM systems are composed of HsdM (N6-adenosine methylation), HsdS (specificity) and HsdR (restriction) subunits, and these proteins act as hetero-oligomers (29, 42). These RM systems are further divided into families that possess similar HsdM and HsdR but distinct HsdS (29). The cleavage sites of the Type I RM systems are located at some distance from the recognition sites (29, 42). Type II RM systems include a methyltransferase and a restriction endonuclease that function independently and typically at the same site (29, 42). Type III RM systems are composed of Mod and Res subunits (29, 39, 42). Similar to Type I RM systems, these subunits form oligomers and the DNA is cut at a site distant from the recognition site (29, 39, 42). Lastly, Type IV RM systems consist of restriction endonucleases that cleave only methylated DNA and are not accompanied by a cognate methyltransferase (42).

RM systems confer to bacterial cells protection against bacteriophages (19); hence, the RM systems in *L. monocytogenes* are expected to negatively influence the effectiveness of bacteriophages as a control measure. Also, RM systems function as a barrier against introduction of foreign DNA (41). This feature of RM systems may contribute to the clonality of lineage I of *L. monocytogenes* although the genetic homogeneity of lineage I may also result from a bottleneck that this lineage may have experienced in its history (9, 26, 46). Recently, the Type IV RM system Mrr of *Escherichia coli* was reported to trigger the SOS response after exposure to high pressure, suggesting that RM systems might be linked to the bacterial stress response (1).

One of the most extensively studied *L. monocytogenes* RM systems is the Sau3AI-like RM system in epidemic clone I (ECI), one of the four clonal groups (ECI, Ia, II, and V) of serotype 4b *L. monocytogenes* associated with multiple listeriosis outbreaks (49). This RM system renders the genomic DNA of the ECI strains resistant to Sau3AI digestion, and among serotype 4b strains it was harbored only by ECI. Thus, the Sau3AI-like RM system has been exploited as a genetic marker specific for ECI against serotype 4b strains (6, 50). A highly homologous RM system was also identified at the genomically equivalent location in some strains of serotypes 1/2a and 4a (49). However, only limited information is available regarding genetic content of other *L. monocytogenes* strains in the region in which ECI carries this RM system. Thus, this study was conducted to identify novel genes and examine genetic diversity in this genomic region.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** A total of 465 isolates were examined in this study a majority of which were provided by the Centers for Disease Control and Prevention (CDC), U.S. Food and Drug Administration (FDA), U.S. Department of Agriculture (USDA) and North Carolina Department of Health and Human Services (NCDHHS) (Table 1).The bacterial cells were routinely grown overnight on Brain Heart Infusion (BHI; Becton, Dickinson and Company, Sparks, MD) with 1.2% agar (Becton, Dickinson and Company) or in BHI broth at 37ºC. When indicated, cells were also incubated overnight at 25 and 42ºC, for one week at 8ºC, and for approximately 27 days at 4ºC.

**Genomic DNA extraction, restriction digestion and Polymerase Chain Reaction (PCR).** Genomic DNA was isolated from overnight cultures grown at 37ºC with the DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA) as described by the manufacturer. Genomic DNA was digested with Sau3AI and MboI (New England Biolabs, Ipswich, MA) and loaded on Tris/Borate/EDTA (TBE) gels containing 0.8% agarose (Apex BioResearch Products, Research Triangle Park, NC) in parallel with undigested genomic DNA.

PCR was conducted in a thermocycler (Biometra, Goettingen, Germany) under the conditions specified in Chapter 4. DNA polymerase exTaq (Takara, Otsu, Shiga, Japan) was employed in the PCR and all the primers used in this study (Table 4) were synthesized by Eurofins MWG Operon (Huntsville, AL). Multiplex PCR to designate serotypes was performed as described by Doumith et al. (11). After PCR, the amplicons were visualized on TBE gels and amplicon size was determined with visual comparison with molecular markers, using manual calculations or with the Kodak 1D image analysis program (EDAS290; Eastman Kodak, New Haven, CT).

**Sequencing, annotation and DNA analysis.** Template DNA for sequencing was prepared by excising the PCR amplicon from the gel and purifying it with the QIAquick gel extraction kit (QIAGEN). Sequencing was conducted by GENEWIZ (South Plainfield, NJ). The sequences were annotated with open reading frame (ORF) finder (www.ncbi.nlm.nih.gov/projects/gorf/), BLASTp (2) and conserved domain searches (25) in the National Center for Biotechnology Information (NCBI) homepage (http://www.ncbi.nlm.nih.gov/). In addition to the sequencing results, DNA sequences and protein sequences were retrieved from the NCBI database and included in the sequence

analysis. DNA and protein sequences were subjected to BLASTn and BLASTp, respectively (2), and aligned with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (7) or EMBOSS Align (http://www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html). Conserved sequences were visualized with WebLogo (http://weblogo.threeplusone.com/create.cgi) (8). DNA sequences were also compared with the online Artemis Comparison Tool (WebACT, http://www.webact.org/WebACT/home) (4) and the GC contents were calculated with Oligonucleotide Properties Calculator (http://www.basic.northwestern.edu/biotools/oligocalc.html) (22). The protein sequence translated from the DNA sequence was obtained with EMBOSS Transeq (http://www.ebi.ac.uk/Tools/emboss/transeq/) (34).

**DNA-DNA dot blots.** All the DNA probes are listed in Table 4. The genomic DNA was spotted on the nylon membrane (GE Water &Process Technologies, Trevose, PA) and bound to the DNA probes labeled with the non-radioactive digioxigenin (DIG) system (Roche, Florence, SC) as described in Chapter 4 The hybridization and stringency washes were routinely performed at 42ºC; however, in some hybridizations with probe 2007-640LP, the temperature was elevated to 47ºC for greater stringency.

**Phage infection assay.** Broad host range *Listeria* phage 20422-1 (23) was amplified in the reference strain DP-L862 and the LmoJ3 strain J3115 as described in Chapter 4. The 20422-1 derivative propagated in J3115 was termed 20422-1MJ3. Both 20422-1 and 20422-1MJ3 infected the reference strain DP-L862 and the strain J3976 grown in 4, 8, 25, 37 and 42ºC according the procedures shown in Chapter 4. The Efficiency of Plaquing (EOP) was defined as the ratio of pfu/ml of J3976 to pfu/ml of the reference strain, DP-L862. The

experiments were conducted at least three times and the results were analyzed by means of the mixed model of ANOVA with SAS (SAS Institute Inc., Cary, NC).

**Pulsed-Field Gel Electrophoresis (PFGE), multilocus genotyping (MLGT) and targeted MLGT (TMLGT).** PFGE with AscI (New England Biolabs) and ApaI (Roche) was conducted as described by Graves et al. (16) and analyzed with BioNumerics (Applied Maths, Austin, TX) with the optimization level and position tolerance at 1.5%. However, these settings were modified when necessary. MLGT and TMLGT were conducted as described (12, 45, 47).

**Nucleotide sequences accession numbers.** DNA sequences determined in this study have been deposited to GenBank under accession numbers JN235991 to JN236005.

## RESULTS

**The genomic region harboring the ECI-specific Sau3AI-like RM system exhibits diversity in RM systems and in a putative lipoprotein gene.** Analysis of the locus genomically equivalent to that harboring the Sau3AI-like RM system in the ECI strain F2365 revealed that in the sequenced genomes of *L. monocytogenes* and *Listeria* spp. this region was flanked by the conserved PTS IIA and threonine aldolase genes upstream and downstream, respectively, although *L. seeligeri* SLCC3954 lacked the PTS IIA gene (Fig. 1). In this study, the DNA sequence between these two conserved genes will be termed Region 85 after the designation of the DNA fragment that first revealed the Sau3AI-like RM system (18).

In addition to the Sau3AI-like RM system, the *in silico* analysis revealed that Region 85 harbored Type I RM systems (e.g. *L. monocytogenes* HPB2262 and *L. seeligeri* SLCC3954) and two Type IV RM system genes each putatively encoding Mrr and McrB (e.g. *L. monocytogenes* FSL J2-003 and H7858), suggesting that this locus might be particularly rich in RM systems in *Listeria* spp. (Fig. 1). Bioinformatics analysis also revealed that the lipoprotein gene in Region 85 of F2365 (*LMOf2365_0329*) was generally conserved among *L. monocytogenes* genomes; however, this gene was missing or diversified in certain lineage III *L. monocytogenes* strains or in non-pathogenic *Listeria* spp. such as *L. innocua* CLIP11262 and *L. welshimeri* SLCC5334 (Fig. 1). Thus, we hybridized the genomic DNA from a panel of isolates (n=114) with the DNA probe derived from this lipoprotein gene (LP) in an attempt to identify additional diversity in the genetic content in this region.

Of the 25 LP negative isolates that were identified, 22 were chosen for further characterization of alternative sequences that may be present in the region (Table 3). DNA from these isolates was used as template in PCR with primers annealing to the PTS IIA and threonine aldolase genes. Eleven PCR amplicons of different sizes (Table 3) were sequenced partially or in their entirety. Sequence analysis revealed novel sequence content in several strains in this region, including a novel *hsdS* gene, two novel Type II RM systems (LmoJ2 and LmoJ3) and new combinations of previously identified Region 85 genes (Fig. 1A and B).

The combined results from the analysis of sequenced genomes and from the sequencing of the PCR products suggested five categories of sequence content in this region (Fig. 1). The first category includes different Type II RM systems, e.g. the Sau3AI-like RM system in F2365 (ECI) and the serotype 4a strain FSL J1-208, LmoJ2 (strain J2479) and

LmoJ3 (strain J3115) (Fig. 1A). Further diversity in this category was conferred by different sequence content of the lipoprotein gene, as shown between ECI F2365 and serotype 4a FSL J1-208 (Fig. 1A).

The second category features three Type I RM systems identified in *L. monocytogenes* serotype 4b ECIa strain HPB2262 (10), *L. monocytogenes* serotype 4b strain J3976, and *L. seeligeri* SLCC3954 (Fig. 1B). These RM systems possessed highly homologous *hsdM* and *hsdR* genes but dissimilar *hsdS* genes (Fig. 1B), which suggest that these Type I RM systems belong to one family.

To compare the *hsdS* genes in Region 85 and identify the sources of these genes, the *hsdS* genes in *L. monocytogenes* J3976, HPB2262 and *L. seeligeri* SLCC3954 were aligned. In the multiple sequence alignment, the three *hsdS* genes revealed three conserved regions consisting of 84 bp at the 5' terminus, 186 bp in the central region and 121 bp at the 3' end (Fig. 2). Variable regions encoded two target recognition domains (TRDs) (Methylase_S, pfam01420), which here will be termed the amino-proximal recognition domain (ARD) and the carboxy-proximal recognition domain (CRD), respectively (27, 29).

In BLASTp searches, the two halves of the HsdS proteins frequently exhibited homology with different proteins, suggesting that the two variable regions might originate from different sources. In the specificity subunit from J3976 (J3976 HsdS), we could not find homologs for the N-terminal half whereas the C-terminal half shared a high similarity with the C-terminal half of other putative HsdS proteins such as LCK_00904 in *Leuconostoc citreum* KM 20 (YP_001728176.1), a putative HsdS in *Streptococcus thermophilus*

(CAE52399.1), and pN42p01 in *Lactobacillus delbrueckii* subsp. lactis NCC88 plasmid pN42 (NP_858129.1).

In the *L. seeligeri* HsdS encoded by *lse_0280,* the N-terminal half was homologous to the N-terminal half of the HsdS proteins including the abovementioned LCK_00904 and the HsdS EfmE1039_1939 in *Enterococcus faecium* E1039 (ZP_06675197.1). Meanwhile, the C-terminal half of Lse_0280 resembled different proteins such as the N-terminal half of the HsdS CUO_2559 in *E. faecium* PC4.1 (ZP_06625490.1), the C-terminal half of the HsdS HMPREF0793_0814 in *Staphylococcus epidermidis* M23864:W1 (ZP_04817946.1) and the N-terminal half of the HsdS HMPREF0496_0710 in *L. brevis* subsp. *gravesensis* ATCC 27305 (ZP_03938596.1).

The N-terminal half of the HsdS LMSG_01574 in *L. monocytogenes* HPB2262 was distinct in that it was homologous to the N-terminal half of proteins encoded by *hsdS* in other *L. monocytogenes* strains, in which *hsdS* was localized in another RM-rich location to be described in Chapter 6. Such proteins include LM5578_0546 in serotype 1/2a 08-5578 (YP_003412663.1) and Lm4b_00545 in serotype 4b CLIP81459 (YP_002757255.1). Outside *Listeria* spp., homologous protein sequences consisted of the N-terminal half of the HsdS LRH_01797 in *L. rhamnosus* HN001 (ZP_03213071.1) and the hypothetical protein EUBHAL_00244 in *Eubacterium hallii* DSM 3353 (ZP_03715198.1) (percent identity 64%). On the other hand, the C-terminal half of LMSG_01574 revealed homology with the C-terminal half of the HsdS pCI305_p3 in *Lactococcus lactis* subsp. lactis plasmid UC317 pCI305 (NP_062461.1), the entire region of the HsdS pL2_p3 in *L. lactis* subsp. lactis

plasmid pL2 (YP_879301.1) (71% identity) and the N-terminal half of a HsdS in *L. delbrueckii* subsp. lactis NCC88 (CAI79467.1) (51% identity).

Downstream of these Type I RM systems, *L. seeligeri* SLCC3954 and *L. monocytogenes* J3976 harbored *mrr* and *mcrB* whereas HPB2262 only harbored *mcrB* (Fig. 1B). Comparison of Region 85 among these genomes revealed other genomic changes such as the diversification of the lipoprotein genes and the replacement of *LMSG_03073* with the novel hypothetical protein gene *J3976NG* (Fig. 1B).

The third category refers to strains that harbor Type IV RM systems in Region 85 without Type I RM systems (Fig. 1C and D). The genome of the serotype 1/2a strain FSL J2-003 harbored both *mrr* and *mcrB* genes in this region, which was otherwise similar to that of HPB2262 except for the presence of *mcrB* and absence of the Type I RM system (Fig. 1C). Compared with FSL J2-003, many strains including H7858 (serotype 4b, ECII [Fig. 1C]), EGDe (serotype 1/2a) and F6854 (serotype 1/2a, ECIII) were missing *mrr* gene and a series of the downstream genes but possessed an otherwise similar Region 85 (Fig. 1C). In *L. monocytogenes* 2007-640 (serotype 4c), Region 85 shared similarity with that of H7858 except for the divergent lipoprotein gene (Fig. 1D).

The fourth category consisted of strains which harbored lipoprotein genes without RM systems in Region 85 (Fig. 1A, C and D). Although RM systems were absent, Region 85 gene content of some of these strains was closely related to other strains. For instance, *L. monocytogenes* strain 2003-151R (indeterminate serotype based on the multiplex PCR serotyping) possessed the Region 85 encoding one lipoprotein gene, which is homologous to the lipoprotein gene in Region 85 of FSL J1-208 (Fig. 1A). Also, Region 85 of the serotype

1/2b strain FSL J1-175 resembled that of H7858 and the only difference was the absence of *mcrB* (Fig. 1C). In Region 85, *L. innocua* CLIP11262 exhibited considerable homology to *L. monocytogenes* 2007-640 although a permease gene (*lin0332*) was located in the place of *mcrB* (Fig. 1D). Meanwhile, in *L. monocytogenes* HCC23 (serotype 4c) and 2001-75R (indeterminate serotype based on the multiplex PCR serotyping), Region 85 was unique and harbored multiple lipoprotein genes (Fig. 1D).

The last category represents strains in which this region harbors neither RM systems nor lipoprotein genes (Fig. 1E). In *L. monocytogenes* serotype 1/2b strain FSL J1-194, Region 85 consisted of three genes encoding a hypothetical protein, a DNA/RNA helicase and a 5'-nucleotidase, respectively (Fig. 1E). *L. welshimeri* SLCC5334 harbored homologs of these genes, which were preceded by four additional genes not encountered in other genomes in this region (Fig. 1E).

Similar, but not identical, lipoproteins in Region 85 created the necessity to systematically analyze these lipoproteins both at the DNA and protein levels. Meanwhile, in some strains, the DNA sequence homologous to other lipoprotein genes formed a pseudogene (FSL J2-003), a shorter gene lacking the 5'end of the other lipoprotein genes (*LmonocyFSL_020200006810* in FSL J1-208 and *lse_0287* in *L. seeligeri* SLCC3954) or two separate genes (*2001-75R_3* and the downstream gene in 2001-75R), suggesting that various genetic changes might have occurred within lipoprotein genes that influence the start and stop of open reading frames. These DNA sequences were included in the DNA sequence analysis along with other lipoprotein genes in Region 85 (Fig. 3A and B). Comparison at the DNA level revealed that all the sequences except for FSL J1-208 lipoprotein gene contained

ATG, which functions as the start codon of all the full-length lipoprotein genes, suggesting that the aforementioned short lipoprotein genes might have formed full-length lipoprotein genes before genetic divergence (Fig. 3B). The other common feature was that all the DNA sequences examined possessed the DNA sequence that encodes the signal peptide for lipoproteins (lipobox) $L_{-3}$-S/$A_{-2}$-A/$G_{-1}$-$C_{+1}$ (15) and the second amino acid sequence often deviated from the conserved sequence in some lipoproteins analyzed (Fig. 3B).

The DNA analysis assigned the lipoprotein genes and their homologous DNA sequences into five types (Fig. 3A). One type represented by the ECI F2365 lipoprotein gene (*LMOf2365_0329*) was the most frequently identified in Region 85 displaying different genetic content (Fig. 3A and B). The DNA sequence from FSL J2-003 revealed a unique additional adenosine located at nt 317, which resulted in its classification as a pseudogene (Fig. 3B).

The second most common lipoprotein type included the lipoproteins located in Region 85 of *L. innocua* CLIP11262 (*lin0331*), *L. monocytogenes* 2007-640 and J3976 (Fig. 3A and B). The third was comprised of the lipoprotein genes in *L. monocytogenes* HCC23 (*LMHCC_2328*) and 2001-75R (*2001-75R_1* and *2001-75R_2*), which bore a highly similar 36 bp additional sequence 2 bp downstream of the DNA sequence encoding the lipobox (Fig. 3A and B). In the middle of this additional sequence, *2001-75R_1* exhibited an extra 36 bp nucleotides (Fig. 3B).

The 36 bp common additional sequence was also harbored by another type of the lipoproteins that consists of the DNA sequences from 2003-151R and FSL J1-208 (Fig. 3B). In this group, FSL J1-208 revealed a sequence highly homologous to 2003-151R in the

upstream intergenic region of the lipoprotein gene; however, ATT in the place of the start codon ATG seemed to result in a shorter ORF (Fig. 3B).

The last type was exemplified by a lipoprotein in *L. monocytogenes* HCC23 (*LMHCC_2327*) and the DNA sequence encompassing the third lipoprotein and its downstream genes (shown 2001-75R_3* in Fig. 3A and B) in Region 85 of *L. monocytogenes* 2001-75R (Fig. 3A and B). Compared with *LMHCC_2327*, the DNA sequence from 2001-75R contained a 14 bp deletion between the end of the third lipoprotein gene (*2001-75R_3*) and the start of the downstream gene (Fig. 3B). The DNA fragment representing the upstream intergenic region and *lse_0287* in *L. seeligeri* SLCC3954 belonged to neither of the aforementioned types and revealed a 100 bp deletion upstream of *lse_0287* (Fig. 3A and B).

Alignment of the protein sequences was conducted with the full-length lipoproteins and the protein encoded by *LmonocyFSL_020200006810* in FSL J1-208 (Fig. 3C). This analysis revealed results consistent to the DNA sequence examination including the additional sequences, the conserved lipobox and the classification of similar lipoproteins (Fig. 3C). Also, this comparison identified six conserved regions (Fig. 3C). In the pair-wise comparisons between the representative lipoproteins, the percent identity at the protein level was lower than the percent identity at the DNA sequence level, which contrasts with the general tendency that similarity is higher at the protein level due to synonymous mutations (Table 5).

To investigate the prevalence of Region 85 genes and deepen our understanding on the diversity of Region 85, we designed 12 DNA probes derived from the RM system genes

(9 probes), lipoprotein genes (2 probes) and the helicase gene in Region 85 of *L. monocytogenes* FSL J1-194 (Table 4). Two additional probes (LMSG_01578 and J3976NG) were designed that represent the genetic differences between HPB2262 and J3976 downstream of the Region 85 Types I and IV RM systems (Fig. 1B and Table 4). Besides these Region 85 probes, we included five epidemic clone probes whose hybridization profiles along with probe 85M (originating from the methyltransferase gene of the Sau3AI-like RM system) unambiguously identify ECI and ECII (6) (Table 4). Also, these DNA probes provide clues for serotype 4b epidemic clones ECIa, ECV (6) and three clonal groups of atypical serotype 4b isolates with a gene cassette formerly considered to be lineage II-specific (Chapter 2). To facilitate the identification of ECV strains, we digested genomic DNA with MboI and Sau3AI since ECV strains are resistant to MboI but susceptible to Sau3AI (13). Probe lmo0737 was included to distinguish the atypical serotype 4b isolates from other serotype 4b strains (Chapter 2) (Table 4).

Hybridizations with these DNA probes were conducted with the total of 465 isolates that consisted of *L. monocytogenes* isolates except for two *L. innocua* isolates (Table 2). These *L. monocytogenes* isolates represent various serotypes: serotypes 1/2a (or 3a) [n=114], 1/2b (or 3b) [n=82], 1/2c (or 3c) [n=11], 4b [n=210] and the "other" serotype category including *L. monocytogenes* isolates of indeterminate serotype, serotypes 4a, 4c, 4d/e and *L. innocua* strains[n=48]. Isolates were from several sources (animal [n=58], environment/food [n=170] and clinical [n=215] origins) (Table 2). A majority of the isolates belonging to the "other" serotype category (n=43) were classified into lineages III or IV by MLGT (Table 1). Serotypes 1/2a (or 3a) and 4b were overrepresented among environmental/food and clinical

isolates, respectively, whereas the half of the animal isolates was classified into the "other" serotype category (Table 2). Most isolates were selected based on availability in our lab; however, isolates from certain other collections available in our laboratory were included in hybridization when they satisfied certain conditions mainly due to changes in the focus of the research. Such examples are all atypical serotype 4b isolates (Chapter 2), Sau3AI-resistant isolates among those collected by CDC in 2004, and isolates negative for probe LP among the Old *Listeria monocytogenes* (OLM) collection.

**Type II RM systems in Region 85 are present in strains of various serotypes.** Our bioinformatics and sequencing results showed that Region 85 harbors at least three Type II RM systems. The Sau3AI-like RM system was examined with probe 85M targeting the methyltransferase gene (*LMOf2365_0327*) (Fig. 1A and Table 4) (50). Hybridization with this probe revealed 75 positive isolates (16.1% of the total isolates) (Table 6). The highest prevalence of the positive isolates was observed among animal isolates (n=12, 20.7%) and clinical isolates (n=38, 17.7%) and among serotype 4b (n=56, 26.7%) (Table 6). We identified four and two positive isolates belonging to serotypes 1/2b (or 3b) and 4c, respectively (Table 6). Strains of these serotypes had not been reported to harbor the Region 85 Sau3AI-like RM system before (49). Three of the positive serotype 1/2b (or 3b) isolates displayed the same MLGT (1.40_1/2b) (Tables 1 and 7). Most of the serotype 4b isolates positive for 85M belonged to ECI; however, we found two non-ECI isolates positive for 85M (Tables 1 and 7).

Isolates positive for 85M differed in the hybridization with the lipoprotein gene probes (LP and 2007-640LP). Two lineage III isolates (CVM02 and FSL J1-208) hybridized

with neither of these probes (Tables 1 and 7). Eight isolates including the aforementioned non-ECI serotype 4b isolates hybridized with probe 2007-640LP representing the lipoprotein in *L. innocua* Region 85 (Tables 1 and 7). Except for one isolate (OLM 75), all these isolates were of animal origin and belonged to lineage III (Tables 1 and 7). Strain OLM 75 was the only serotype 1/2a (or 3a) isolate that was susceptible to Sau3AI digestion even though the hybridization and PCR results suggested the presence of the Sau3AI-like RM system (Tables 1 and 7). The other 85M-positive isolates except for one (OLM 65) bound only to probe LP targeting the frequently found *L. monocytogenes* Region 85 lipoprotein gene (*LMOf2365_0329*) (Tables 1 and 7). The serotype 4b isolates positive both for 85M and LP all belonged to ECI and conversely, all the ECI strains identified in our analysis exhibited this hybridization pattern except for one atypical ECI isolate (OLM 65) (Tables 1 and 7). This isolate generated signal for both LP and 2007-640LP and will be discussed in more detail in the section on the Type I RM systems (Tables 1 and 7).

For the novel Type II RM systems LmoJ2 and LmoJ3, we designed probes J2479Res and J3115Met, respectively (Fig. 1A and Table 4). Hybridization with these probes suggested that the novel RM systems might be present at low prevalence since we identified 14 LmoJ2 isolates (3.0% of the total isolates) and three LmoJ3 isolates (0.6% of the total isolates) (Table 6). Characteristically, neither of the other Region 85 probes hybridized with these isolates (Table 7). Other details on the LmoJ2 and LmoJ3 isolates will be discussed in Chapter 4.

**A family of Type I RM systems is present predominantly among serotype 4b isolates.** To identify Type I RM systems belonging to the same family as those identified

through the *in silico* analysis and sequencing, we designed probe LMSG_01573 targeting the *hsdM* gene (*LMSG_01573*) (Fig. 1B and Table 4) and identified 81 positive isolates (17.4% of the total isolates) (Table 6). Most (71/81) of these isolates belonged to serotype 4b; prevalence was markedly higher among serotype 4b isolates (33.8%) as compared among serotypes 1/2a (or 3a) (n=1, 0.9%), 1/2b (or 3b) (n=2, 2.4%) and the "other" serotype category (n=7, 14.6%) (Table 6). MLGT results revealed that the single serotype 1/2a (or 3a) and all the isolates belonging to the "other" serotype category belonged to lineage III (Tables 1 and 7).

Diversity in *hsdS* genes was examined with probe LMSG_01574C based on the central conserved region and two probes LMSG_01574 and J3976TypeIS_V representing the CRDs of the Region 85 *hsdS* genes of HPB2262 (ECIa) and J3976, respectively (Fig. 1B and Table 4). Except for one (J3913), all isolates positive for probe LMSG_01573 bound with probe LMSG_01574C (Tables 1 and 7). Thirty eight LMSG_01573-positive isolates hybridized with probe LMSG_01574 and another 13 isolates bound to J3976TypeIS_V (Table 7). These findings suggest that, in these isolates, at least one target recognition domain shares similarity to the carboxy-proximal recognition domains of HPB2262 and J3976 *hsdS* genes. Among LMSG_01573 negative isolates, probes LMSG_01574 and J3976TypeIS_V hybridized with 40 and one isolates, respectively (Table 7), and PCR results suggested that the homologous DNAs are located outside Region 85 (data not shown).The other 30 isolates positive for LMSG_01573 bound neither of LMSG_01574 nor J3976TypeIS_V, which is indicative of novel *hsdS* genes that cannot be identified with our DNA probes (Table 7).

Two LMSG_01573-positive isolates were negative for probes Mrr and McrB representing Region 85 Type IV RM systems; however, all others hybridized with one or both of these probes (Table 7). It is noteworthy that (1) all isolates negative for probe Mrr (n=6) belonged to lineages III or IV based on MLGT and (2) a majority of serotype 4b isolates negative for probe McrB (21 out of 22) displayed the same hybridization profile as ECIa strains 4b1 and B068 for all the probes used in this study (Tables 1 and 7).

On exceptional strain (OLM 65) was positive for LMSG_01573 and 85M (Tables 1 and 7). The PCR employing primers annealing to a flanking gene and a DNA probe target suggested that the Sau3AI-like RM system might be located between the upstream Type I RM system and *mrr* and the downstream *mcrB* (data not shown). This strain displayed the typical ECI hybridization profile for the epidemic clone probes (Table 1), was clustered with other ECI strains on the PFGE dendrogram (data not shown) and had the ECI MLGT (1.13_4b_Sw87_ECI) (Table 1).

Among the previously identified serotype 4b clonal groups, ECIa and atypical serotype 4b strains (Groups 1-3) were positive for MLSG_01573 (Tables 1 and 7). Hybridization results suggested that Region 85 of Group 1 strains might resemble that of J3976 whereas Groups 2 and 3 might harbor Region 85 similar to ECIa strains although they were positive for probe McrB (Tables 1 and 7). Hybridization profiles of Groups 2 and 3 were identical except for the *hsdS*-based probes (Tables 1 and7). Additionally, we identified 7 clusters of isolates based on the hybridization profiles for all DNA probes used (Tables 1 and 7). These clusters included one serotype 1/2b (or 3b) group (P4) and one serotype 4b

group (P5) that were closely related in PFGE dendrograms (data not shown). Strains within the serotype 4b group (P5) displayed the same MLGT (1.5_4b) (Table 1).

**Mrr and McrB are often identified among serotypes 1/2a, 1/2b and 1/2c without Type I RM systems in Region 85.** Hybridization with probe Mrr revealed 126 positive isolates (27.1% of the total isolates) (Table 6). These positive isolates were under-represented among animal isolates (n=3, 5.2%) in comparison with clinical (n=60, 27.9%) or environmental/food isolates (n=59, 34.7%), and a majority of them belonged to serotypes 1/2a (or 3a) (n=49, 43.0%) and 4b (n=70, 33.3%) (Table 6).

All serotype 4b isolates positive for probe Mrr also hybridized with probe LMSG_01573, suggesting a close linkage between Type I RM systems and Mrr in serotype 4b strains (Table 7). On the other hand, serotype 1/2a (or 3a) isolates positive for probe Mrr were negative for probe LMSG_01573 but positive for probe McrB (Table 7). These isolates exhibited the hybridization profile for Region 85 probes expected for serotype 1/2a *L. monocytogenes* FSL J2-003 (Fig. 1C and Table 7). Such a profile was the one most frequently identified among serotype 1/2a (or 3a) isolates and was not found among other serotypes (Table 7).

Positive isolates for the McrB probe were present at the highest percentage (n=268, 57.6% of the total isolates) among all Region 85 RM system probes (Table 6). The McrB probe hybridized more frequently with serotypes 1/2a (or 3a) (n=84, 73.7%), 1/2b (or 3b) (n=60, 73.2%) and 1/2c (or 3c) (n=11, 100%) than with serotype 4b (n=104, 49.5%) (Table 6). Like Mrr, isolates positive for the McrB probe were under-represented among animal

isolates (n=15, 25.9%) as compared with clinical (n=124, 57.7%) and environmental/food isolates (n=118, 69.4%) (Table 6).

Many McrB-positive isolates failed to hybridize with any of the Region 85 RM system probes (Table 7). Our hybridization and PCR results suggested that a majority of such isolates harbored Region 85 similar to serotype 4b *L. monocytogenes* H7858 (ECII) (n=155, 33.3% of the total isolates) (Fig. 1C and Table 7). This H7858-like Region 85 was the one most frequently identified among serotypes 1/2b (or 3b) (n=59, 72.0%) and 1/2c (or 3c) (n=11, 100%), and the second most prevalent among serotype 1/2b (or 3b) (n=33, 28.9%) after the Region 85 similar to FSL J2-003 (Table 7). The following two findings are noteworthy: (1) many serotype 1/2b (or 3b) isolates (n=40) harboring the H7858-like Region 85 hybridized with probe LMSG_01574 (based on the variable region of ECIa strain HPB2262 *hsdS* gene) although PCR results suggested that the target gene might be located outside Region 85; and (2) among serotype 4b isolates, only ECII and ECV isolates displayed the H7858 profile for Region 85 probes (Tables 1 and 7).

**A unique gene in *L. welshimeri* Region 85 is often found among animal isolates of *L. monocytogenes*.** The presence of the unique genes identified in Region 85 of *L. monocytogenes* FSL J1-194 and *L. welshimeri* SLCC5334 was examined with probe J1-194_1 based on the Region 85 DNA/RNA helicase gene (*LMBG_02168*) (Fig. 1E and Table 4). Hybridization results revealed 28 positive isolates (6% of the total isolates) belonging of serotypes 1/2a (or 3a) (n=9), 1/2b (or 3b) (n=8), 4b (n=2) and the "other" serotype category (n=9) (Table 6). Both serotype 4b isolates belonged to lineage III based on MLGT, indicating that no lineage I serotype 4b isolates are positive for FSL J1-194_1 (Tables 1 and 7).

Frequency of the isolates positive for probe J1-194_1 was markedly higher among animal isolates (n=14, 24.1%) compared with clinical (n=11, 5.1%) and environmental/food (n=2, 1.2%) isolates (Table 6). Most of J1-194_1-positive isolates did not hybridize with any of the other Region 85 probes; however, four bound to probe 2007-640LP representing the *L. innocua* Region 85 lipoprotein gene and another four isolates hybridized with probes 2007-640LP and McrB (Table 7). In the MLGT analysis, these eight isolates all belonged to lineages III or IV (Tables 1 and 7).

**Lipoprotein gene (*LMOf2365_0329*) and its *L. innocua* homolog are present in a mutually exclusive manner and the *L. innocua* homolog is over-represented among animal isolates of *L. monocytogenes*.** Hybridization with probe LP based on the Region 85 lipoprotein of ECI strain F2365 (*LMOf2365_0329*) revealed 351 positive isolates (75.7% of the total isolates), which is the highest prevalence of the Region 85 genes examined here (Table 6). The LP positive isolates were under-represented among animal isolates (n=14, 24.1%) and among the "other" serotype category (n=7, 14.6%) as compared among other sources or serotypes (75.8-100%) (Table 6). Hybridization with probe 2007-640LP targeting the Region 85 lipoprotein in *L. innocua* CLIP11262 (*lin0331*) suggested that isolates positive for this probe are negative for LP and vice versa except for one isolate (OLM 65) positive for both probes (Table 7). Thus, the isolates positive for probe 2007-640LP were over-represented among animal isolates (n=24, 41.4%) and the "other" serotype category (n=25, 52.1%) (Table 6).

This mutually exclusive feature was also identified among other Region 85 probes that hybridized with isolates positive for probes LP or 2007-640LP. Although a few exceptions

were observed, signal from probes LMSG_01574 or LMSG_01578 was identified only among the isolates positive for LP (Table 7) whereas the isolates positive for probes J3976TypeIS_V or J3976NG generated signal for probe 2007-640LP (Table 7).

Twenty eight isolates (6.0% of the total isolates) hybridized only with LP and one isolate hybridized only with probes LP and J3976TypeIS_V among Region 85 probes (Table 7). Our PCR results suggest that these isolates might harbor a Region 85 similar to *L. monocytogenes* FSL J1-175 (data not shown). Meanwhile, thirteen isolates (2.8% of the total isolates) were positive only for probe 2007-640LP and most of these isolates belonged to lineages III or IV (Tables 1 and 7). The PCR results with primers targeting probe 2007-640LP and those targeting flanking genes suggested that six of these isolates might harbor a Region 85 similar to 2003-151R (Table 1).

**Strains hybridized with no Region 85 probes all belonged to lineage III.** Ten isolates (2.2% of the total isolates) did not hybridize with any of the Region 85 probes and all belonged to lineage III (Tables 1 and 7). In the PCR amplifying the region between the flanking genes generated a single band of approximately 2, 2.8 (2001-75R, Fig. 1D), 4 or 7 kb (Tables 1 and 7).

**Region 85 genes are organized in a fixed order and the spacing is conserved among different strains.** The location of Region 85 probe targets was examined by PCR using primers annealing to the target gene and a flanking gene. This analysis suggested that Region 85 probe targets are located in the following order (upstream to downstream) : PTS IIA gene→Type I RM system→*mrr*→Sau3AI-like RM system→*LMSG_03070* or *J3976NG* homolog→lipoprotein gene homologous to *LMOf2365_0329* or *lin0331*→*mcrB*→threonine

aldolase gene (data not shown). Among the isolates positive for probe J1-194_1, the genes were located in the order of PTS IIA gene→target gene of probe J1-194_1→*lin0331* homolog→*mcrB*→threonine aldolase gene (data not shown). The size of the PCR amplicons indicated that the space between two genes is maintained even when genetic contents of Region 85 differ and that one missing gene leads to a reduced amplicon size, suggesting that no additional DNA might be present in the place of the missing gene (data not shown).

**Horizontal gene transfer might be involved in the Region 85 gene acquisition.** A majority of the genes in Region 85, except for *LMOh7858_0335* and its homologs (GC 38-40%), displayed GC contents ranging from 26 to 36% (Fig. 1), which is lower than the genome average of *L. monocytogenes* genome (38%) (30). This relatively low GC content contrasts to the flanking sequences encoding PTS IIA (GC 42-43%) and threonine aldolase (GC 41-45%) for which GC content was above the genome average (Fig. 1). Taken together, these lines of evidence suggest that the genes in and adjacent to Region 85 may have resulted from different horizontal gene transfer events.

In an attempt to investigate potential sources of these genes outside *Listeria* spp., the BLASTp was conducted and the hits from non-*Listeria* microorganisms were examined. Among the RM system proteins, HsdM and HsdR (encoded in *lse_0279* and *lse_0281* in *L. seeligeri* SLCC3954) shared considerable homology with the putative HsdM and HsdR proteins in non-*Listeria* organisms such as *Bacillus coagulans* 36D1 (ZP_04431903.1, 74% identity for HsdM; and ZP_04431905.1, 70% identity for HsdR), *Lactococcus lactis* subsp. lactis CV56 (ADZ63271.1, 67% identity for HsdM; and ADZ63273.1, 69% identity for HsdR); and *Alkaliphilus metalliredigens* QYMF (YP_001319443.1, 67% identity for HsdM;

and YP_001319445.1, 65% identity for HsdR). For the latter two microorganisms whose genomes were available, *hsdM* and *hsdR* were separated by *hsdS,* suggesting that these three genes might constitute one genetic cassette. Also, the components of the Sau3AI-like RM system (encoded in *LMOf2365_0325*, *LMOf2365_0326* and *LMOf2365_0327*) exhibited homology with various proteins from non-*Listeria* microorganisms; however, the hits of each protein did not form a genetic cassette. Among the Type IV RM systems, the top hits of the BLASTp with McrB (encoded in *LMOh7858_0337*) were restricted to *Listeria* spp. whereas Mrr (encoded by *lse_0282*) shared similarity with the Mrr proteins originating from non-*Listeria* microorganisms including various *Enterococcus faecalis* strains (up to 59% identity).

Among the Region 85 genes not involved in RM systems, we selected for BLASTp the lipoprotein (encoded in *LMOf2365_0329*) and two hypothetical proteins encoded in genes *LMSG_03073* and *J3976NG* that provided evidence for genetic diversity downstream of the Type I RM systems in Region 85. The lipoprotein and LMSG_03073 exhibited no significant hits outside *Listeria* spp., and the search with J3976NG revealed no significant homologs among and outside *Listeria* spp. Among the flanking genes of Region 85, the BLASTp search with the PTS IIA protein (encoded in *LMOf2365_0322*) suggested only a low degree of homology with the non-*Listeria* PTS IIA proteins. However, proteins similar to threonine aldolase (encoded in *LMOf2365_0330*) were identified in various non-*Listeria* microorganisms such as *Exiguobacterium sibiricum* 255-15 (YP_001813866.1, 60% identity) and *Paenibacillus* sp. Y412MC10 (YP_003243675.1, 60% identity).

To scrutinize the beginning of the genetic changes in Region 85 and to uncover DNA sequences involved in the horizontal gene transfer, we compared the DNA sequences encompassing a part of the flanking gene and the adjacent region within Region 85 from the strains that harbor various genetic contents in Region 85. In the upstream, partial PTS IIA gene sequences revealed similarity in all the strains although a few single nucleotide polymorphisms (SNPs) were identified mostly in one particular location (33th nucleotide upstream the start codon, -33) (Fig. 4A). Conserved sequence continued further downstream to the 92$^{nd}$ nucleotide into the intergenic region (+92) (Fig. 4A). One exception to this trend was *L. welshimeri* SLCC5334 that was considerably rich in SNPs and possessed a 261 bp additional sequence in the middle of the conserved intergenic region sequence (Fig. 4A). In this insertion, the start codon was located for the first gene in the Region 85 of *L. welshimeri* SLCC5334 (*lwe0277*) (Fig. 4A).

In the downstream, the partial threonine aldolase gene sequences contained more SNPs than the PTS IIA gene sequences although its DNA sequence was generally highly homologous among all the strains. In particular, the 8$^{th}$ and 9$^{th}$ nucleotides (-8 and -9) upstream the start codon particularly lacked homogeneity (Fig. 4B). Downstream the threonine aldolase gene, we identified a 28 bp DNA sequence (up to +28) that was highly conserved among all the strains examined (Fig. 4B). This general tendency was defied by the novel RM system isolates J2479 and J3115. The LmoJ2-harboring J2479 revealed abundant SNPs and three short insertions (9, 6, and 3 bp, respectively) whereas the LmoJ3-bearing J3115 showed an additional sequence of 39 bp (Fig. 4B).

Comparison of the upstream and downstream flanking regions identified one pair of the inverted repeats (CTCCTTTT…AAAAGGAG) in the upstream and downstream adjacent sequences of the flanking genes within Region 85 (Fig. 4A and B). These repeats resided near the start codon of the flanking genes, i.e., 8 bp apart from the PTS IIA gene (Fig. 4A) and 9 bp from the threonine aldolase gene (Fig. 4B). These inverted repeats suggest that a site-specific recombination might have been involved in the horizontal gene transfer of Region 85 genes. Strains that manifested marked genetic changes in the conserved sequence often harbored disrupted inverted repeats. For instance, G was inserted in the upstream inverted repeat of *L. welshimeri* SLCC5334, resulting in CTCCTTTGT (Fig. 4A) whereas LmoJ2-harboring J2479 harbored one SNP and a 9 bp insertion in the downstream inverted repeat (AATATTTAGCTAGGGAG) (Fig. 4B).

Intriguingly, between the Sau3AI-like RM system and the lipoprotein gene in Region 85 of ECI strain F2365, we identified a hypothetical protein gene (*LMOf2365_0328*) whose BLASTp results included a recombinase in *Lactobacillus ruminis* ATCC 25644 (ZP_08079957.1, percent identity 38%) (Fig. 1A). This suggests that the protein encoded in this gene might be involved in horizontal gene transfer as a recombinase. However, the BLASTp analysis also exhibited many McrA proteins in the results, raising a possibility that this protein might be a McrA protein.

**Mrr did not provide protection against phage infection.** Characterization of the novel RM systems LmoJ2 and LmoJ3 to be described in Chapter 4 showed that these RM systems might methylate the 5-carbon of cytosine in GCWGC (W=A or T) and GCNGC (N=A, T, C and C), respectively (Chapter 4). Since 5-methylcytosine was contained in the

substrate of the some Mrr and McrB proteins (51), and LmoJ2 and LmoJ3 isolates were all negative for probes Mrr and McrB (Table 6), we hypothesized that the Mrr and/or McrB in Region 85 might cleave the DNA methylated by LmoJ2 and LmoJ3, thus preventing a strain from possessing these Type II and Type IV RM systems at the same time.

This hypothesis was tested for Mrr by infecting the reference strain DP-L862 (positive for the McrB probe but negative for the Mrr probe) and the strain J3976 (positive for probes for McrB and Mrr) grown at various temperatures (4, 8, 25, 37 and 42ºC) with the *Listeria* phage 20422-1 and its derivative methylated at GCNGC through propagation in the LmoJ3 isolate J3115 (20422-1MJ3). The methylation by LmoJ3 increased infectivity to J3976 (by ca. 24 fold, $p<0.0001$) at low temperatures (4 ºC) while no significant difference was observed at other temperatures ($p>0.05$). This finding is opposite to our expectation that the methylated phages might infect J3976 less effectively due to the action of Mrr (Fig. 5). The effect of McrB could not be estimated in this experiment since both strains were positive for the McrB probe and the restriction by the Type I RM system in J3976 is not likely to be affected by the methylation by LmoJ3 given that the methylation by Type I RM systems occurs at adenosine (42).

## DISCUSSION

In this study, our analysis revealed that a certain genetic location (Region 85) harbored a high density of RM systems including three Type II RM systems, a family of Type I RM systems and two Type IV RM systems (Mrr and McrB). Also, Region 85 displayed a remarkable genetic diversity, which contrasts to the general synteny displayed by

*Listeria* spp. genomes (3). Previously, a highly variable, RM system-rich region was identified among *Escherichia coli* and *Salmonella* Typhimurium strains, and termed the Immigration Control Region (ICR) (29, 33). Our findings suggest that Region 85 is the first ICR identified among *Listeria* spp. Abundance in RM systems in this locus is expected to play a pivotal role in phage resistance, evolution and stress response of *L. monocytogenes*.

The Sau3AI-like RM system is the first and most extensively studied RM system in the ICR Region 85. Our study provided new information on this RM system. For instance, this is the first report that the Sau3AI-like RM system is distributed among serotypes 1/2b or 4c in addition to the previously identified serotypes 1/2a, 4b and 4a (49). The Sau3AI-like RM system has been used as a genetic marker to distinguish ECI strains from other serotype 4b strains (6, 50). In this study, we discovered that a few non-ECI serotype 4b isolates belonging to lineage III harbored the Sau3AI-like RM system, challenging the strict association of this RM system with ECI strains among serotype 4b isolates. Lastly, one strain was positive for the Sau3AI-like RM system probe but revealed susceptibility to Sau3AI when its genomic DNA was digested with this enzyme. This finding suggests that, in some strains, the Sau3AI-like RM system might not be functional at least under the growth conditions that were used.

One of the main findings revealed by this study is that Type I RM systems belonging to the same family are harbored most frequently among serotype 4b as compared with other serotypes. Such Type I RM systems might contribute to clonality observed for lineage I although it still remains unknown why many serotype 4b isolates harbor Type I RM systems in Region 85. *In silico* analysis and hybridization with the *hsdS*-based probes indicated that

the variable regions in *hsdS* might have been acquired through horizontal gene transfer both at intra- and inter-species levels. In addition, many Type I RM systems in Region 85 seem to possess novel *hsdS* genes. Investigating the sequences of such novel *hsdS* genes could enhance our understanding of the genetic diversity of this *Listeria* ICR, in particular among serotype 4b strains.

All serotype 4b isolates examined in this study were assigned into previously identified serotype 4b clonal groups, i.e., ECI, ECIa, ECII, ECV and Groups 1-3 (described in Chapter 2). Our results suggested that each serotype 4b clonal group exhibited the same hybridization profile for Region 85 probes with the exception of one ECI isolate and one Group 1 isolate, suggesting that Region 85 has rarely been diversified among the strains belonging to the same clonal group. The Region 85 probe profile for an epidemic clone was not found among serotype 4b isolates not belonging to the particular epidemic clone. Hence, unique features of Region 85 hybridization profiles could be employed as new genetic markers for epidemic clones against serotype 4b strains. For instance, ECIa isolates are distinct from other serotype 4b strains in that they harbor the Type I RM system but lack *mcrB*, which accompanies the Type I RM system in non-ECIa serotype 4b isolates. Only one non-ECIa isolate exhibited the same characteristic. This finding could be instrumental as a genetic marker for ECIa given that reliable DNA probes specific to this clonal group have not been identified yet. On the other hand, the novel serotype 4b clonal groups Groups 1-3 shared the hybridization profiles of Region 85 probes with other serotype 4b isolates. Our analyses in Chapter 2 suggested that Group 1 is distinct from the others whereas Groups 2

and 3might be closely related to each other. Region 85 hybridization profiles were consistent to these results.

Our hybridization results suggested that serotype 4b isolates exhibited more complexity in the ICR Region 85 than serotypes 1/2a, 1/2b, and 1/2c that predominantly harbored Type IV systems without Type I RM systems in Region 85. This observation contrasts to the well-established finding that serotypes 1/2a and 1/2c (lineage II) tend to be more diverse than lineage I including serotype 4b (9, 31, 32).

Some Region 85 genes were over-represented among animal isolates and among lineages III or IV strains. Also, certain hybridization profiles were predominantly or exclusively identified among lineages III or IV strains. All these findings indicate that Region 85 might differ considerably between lineages III or IV strains and other lineages. Furthermore, the observation that isolates hybridized with neither of the Region 85 probes used here belonged to lineage III raises the possibility that novel genomic contents, e.g. potentially new RM systems, might be identified in Region 85 of lineages III or IV strains. These characteristics of lineages III or IV might contribute to the unique epidemiological characteristics of these lineages such as their frequent involvement with animal listeriosis cases.

In a study on the evolution of the *E. coli* ICR, Sibley et al. proposed that the Type I and IV RM systems in the *E. coli* ICR might have been acquired through multiple site-specific recombination events (37). In our study, we found two lines of evidence supporting the horizontal gene transfer through site-specific recombination: (1) all the RM systems and lipoprotein genes in this location exhibited GC contents lower than the genome average; and

(2) a pair of inverted repeats were located in the conserved flanking regions of Region 85. However, unlike the *E. coli* ICR whose genetic changes started abruptly after the inverted repeats, the conserved regions around Region 85 continued after the inverted repeats. Also, the order and distance between Region 85 genes seem to be maintained from one strain to another. These features led to the speculation that the genes in Region 85 might have been transferred as a single fragment into an ancestral strain of *Listeria* spp. and current diversity might have been established mainly due to serial deletions from the original DNA fragment. Such deletions have been reported in *L. welshimeri* SLCC3954 genomes in comparison with the *L. monocytogenes* and *L. innocua* genomes (17). Although little information is available on which genes were transferred in the putative ancient horizontal gene transfer, the identification of an atypical ECI isolate (OLM 65) possessing the Type I RM system, the Type II Sau3AI-like RM system, and the Type IV *mrr* and *mcrB* suggests that at least these genes would have composed the originally transferred DNA fragment. A candidate recombinase involved in the ancient event might be encoded in a putative recombinase gene downstream of the Sau3Ai-like RM system. Also, our analyses suggested that diversity identified in Region 85 might be increased by genetic changes in lipoprotein genes and some additional horizontal gene transfer potentially involving LmoJ2 and LmoJ3.

In conclusion, we have shown that the genetic diversity in the *Listeria* ICR Region 85 is associated with isolation sources, serotypes, and population lineages of *L. monocytogenes*. Such diversity seems to originate mainly from the serial deletions from a large DNA fragment horizontally acquired by an ancestral *Listeria* strain to the form optimized for the environmental niches. However, except for the Sau3AI-like RM system, we lack information

on whether individual genes are actually expressed and play a biologically important role in the cells. Further studies on this subject will increase our understanding of the RM systems affecting the evolution and ecology of *L. monocytogenes*.

## Table 3.1 Isolates examined in this study

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| CVM02 | 2001 | NC | A | 1/2b, 3b | - | + | ■ | □ | ■ | □ | □ | ■ | □ | 1 | | Lm4.1 (III[g]) | |
| FSL J1-208 | N/A[a] | N/A | A | 4a | - | + | ■ | □ | ■ | □ | □ | □ | □ | 1 | | | |
| OLM 75 | 1952 | Canada | C | 1/2a, 3a | + | + | ■ | □ | ■ | □ | □ | □ | □ | 2 | | 2.12_1/2a | |
| NRRL B-33190 | N/A | N/A | A | 4b | - | + | ■ | □ | ■ | □ | □ | □ | □ | 2 | | Lm3.12 | (45) |
| OLM 39 | 1939 | IA | A | 4b | - | + | ■ | □ | ■ | □ | □ | □ | □ | 2 | | Lm3.1 | |
| NRRL B-33184 | N/A | N/A | A | 4c | - | + | ■ | □ | ■ | □ | □ | □ | □ | 2 | | Lm3.8 | (45) |
| NRRL B-33382 | 2004 | N/A | A | Indeterminate | - | + | ■ | □ | ■ | □ | □ | □ | □ | 2 | | Lm3.35 | (45) |
| NRRL B-33383 | 2004 | N/A | A | Indeterminate | - | + | ■ | □ | ■ | □ | □ | □ | □ | 2 | | Lm3.36 | (45) |
| NRRL B-33177 | N/A | N/A | A | 4c | - | + | ■ | □ | ■ | □ | □ | ■ | □ | 2 | | Lm3.4 | (45) |
| NRRL B-33365 | 2002 | N/A | A | Indeterminate | - | + | ■ | □ | □ | □ | □ | □ | □ | 2 | | Lm3.20 | (45) |
| LW-A11 | 2001 | N/A | E/F | 1/2a, 3a | - | + | ■ | N[f] | N | N | N | N | ■ | 3 | | | |
| LW-A29 | 2002 | N/A | E/F | 1/2a, 3a | - | + | ■ | N | N | N | N | N | ■ | 3 | | | |
| LW-A33 | 2003 | N/A | E/F | 1/2a, 3a | - | + | ■ | N | N | N | N | N | ■ | 3 | | | |
| LW-A36 | 2003 | N/A | E/F | 1/2a, 3a | - | + | ■ | N | N | N | N | N | ■ | 3 | | | |
| LW-A8 | 2000 | N/A | E/F | 1/2a | - | + | ■ | N | N | N | N | N | ■ | 3 | | | |
| LW-A95 | 2004 | N/A | E/F | 1/2a, 3a | - | + | ■ | N | N | N | N | N | ■ | 3 | | | |
| SE-076 | N/A | N/A | E/F | 1/2a, 3a | - | + | ■ | N | N | N | N | N | ■ | 3 | | | |
| #22 | 2003 | NC | E/F | 1/2b, 3b | - | + | ■ | N | N | N | N | N | □ | 3 | | 1,40_1/2b | |
| 82a-6 | 2003 | NC | E/F | 1/2b, 3b | - | + | ■ | N | N | N | N | N | □ | 3 | | 1.40_1/2b | |
| 82b-1 | 2003 | NC | E/F | 1/2b, 3b | - | + | ■ | N | N | N | N | N | □ | 3 | | 1.40_1/2b | |
| 2008-114 | 2008 | NC | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| 2008-451 | 2008 | NC | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2010-0055 | 2010 | NC | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| 2010-0072A | 2010 | NC | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| AT-16 | 1988 | N/A | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | 1.13_4b_Sw87_EC1 | |
| F6812 | 1988 | N/A | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| FDA 5 | 1986 | N/A | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| FDA 93 | 1986 | N/A | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| FDA 97 | 1986 | N/A | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| G3990 (F2365)[h] | 1985 | CA | E/F | 4b | - | N | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | Lm1.11 | (12, 30) |
| J2213 | 2003 | AZ | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2269 | 2003 | GA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2275 | 2003 | PA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2282 | 2003 | MD | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2288 | 2003 | TX | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2302 | 2003 | CA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2313 | 2003 | TX | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2327 | 2003 | MI | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2353 | 2003 | IL | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2391 | 2003 | TX | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2584 | 2003 | VT | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (38) |
| J2854 | 2004 | AZ | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J2985 | 2004 | IL | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3106 | 2004 | NY | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3133 | 2004 | TX | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3232 | 2004 | OK | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3410 | 2005 | SC | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3535 | 2005 | RI | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3559 | 2005 | GA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3592 | 2005 | ME | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3709 | 2005 | NH | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3799 | 2005 | CT | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J3916 | 2006 | NM | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4001 | 2006 | TX | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4099 | 2006 | VA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4116 | 2006 | ME | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4187 | 2006 | WI | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4253 | 2006 | TN | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4274 | 2006 | NH | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4297 | 2006 | PA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4316 | 2006 | IA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J4685 | 2007 | MO | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| J5043 | 2008 | MA | C | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| L0228 | 2003 | VA | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (13) |
| L0328 | 2003 | VA | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (13) |
| L0616 | 2004 | VA | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (13) |
| L0617 | 2004 | VA | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | (13) |
| LW-A4 | 2001 | N/A | E/F | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| NC0064408a-3 | N/A | NC | N/A | 4b | N | N | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| NC0064408a-5 | N/A | NC | N/A | 4b | N | N | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NC0065706 48h H+ | N/A | NC | N/A | 4b | N | N | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| OLM 126* | 1959 | Canada | A | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| OLM 20 | 1937 | USA | A | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 3 | ECI | | |
| LW-A7 | 2000 | N/A | E/F | Indeterminate (1/2a) | - | + | ■ | N | N | N | N | N | □ | 3 | | | |
| #191A (90)[h] | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 4 | | 2.30_1/2a_T189 | (28) |
| L1522a (1747)[h] | 2005 | VA | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 4 | | 2.30_1/2a_T189 | (28) |
| LW-A75 | 2003 | N/A | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 4 | | 2.20_1/2a_T492 | |
| LW-A97 | 2004 | N/A | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 4 | | 2.20_1/2a_T492 | |
| (HI) 267 | 1999 | N/A | C | 4b | - | N | □ | □ | ■ | □ | □ | □ | □ | 4 | | Lm3.49 | |
| J2479 | 2003 | MI | C | 4b | + | - | □ | □ | ■ | □ | □ | □ | □ | 4 | | Lm3.42 | (38) |
| NRRL B-33191 | 2001 | N/A | A | 4b | + | + | □ | □ | ■ | □ | □ | □ | □ | 4 | | Lm3.13 | (45) |
| OLM 115 | 1955 | Canada | C | 4b | + | + | □ | □ | ■ | □ | ■ | ■ | □ | 4 | | 1.70_4b | |
| OLM 116 | 1956 | Canada | C | 4b | + | + | □ | □ | ■ | □ | ■ | ■ | □ | 4 | | 1.70_4b | |
| OLM 80 | 1954 | Canada | C | 4b | + | + | □ | □ | ■ | □ | ■ | ■ | □ | 4 | | 1.70_4b | |
| 2008-392 | 2008 | NC | C | 4c | + | + | □ | □ | ■ | □ | □ | □ | □ | 4 | | Lm3.46 | |
| NRRL B-33330 | 2002 | NH | E/F | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 4 | | Lm3.17 | (45) |
| NRRL B-33372 | 2004 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 4 | | Lm3.26 | (45) |
| OLM 62 | 1953 | Canada | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 4 | | Lm3.42 | |
| 2010-0072B | N/A | NC | C | 1/2a, 3a | + | + | □ | □ | □ | □ | □ | □ | ■ | 5 | | 2.92_1/2a | |
| 2007-584 | 2007 | NC | C | 4b | + | + | □ | □ | □ | □ | □ | ■ | □ | 5 | | 1.74_4b | |
| J3115 | 2004 | VA | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 5 | | 1.73_4b | |
| NRRL B-33374 | 2004 | N/A | A | 1/2a, 3a | + | + | □ | □ | □ | □ | □ | □ | ■ | 6 | | Lm3.28 | (45) |
| J3720 | 2005 | PA | C | 4b | + | + | □ | □ | ■ | □ | □ | □ | □ | 7 | | Lm3.42 | |
| NRRL B-33408 | 1997 | N/A | A | 4b | + | + | □ | □ | ■ | □ | □ | □ | □ | 8 | | Lm4.3 | (45) |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | Imo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3913 | 2006 | IN | C | 4b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | ■ | 9 | Group 1 | 1.17_4b | |
| 2007-0904 | 2007 | NC | C | 4b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | ■ | 10 | Group 1 | 1.17_4b | |
| J2255 | 2003 | GA | C | 4b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | ■ | 10 | Group 1 | 1.17_4b | (38) |
| J3139 | 2004 | WI | C | 4b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | ■ | 10 | Group 1 | 1.17_4b | |
| J4016 | 2006 | SC | C | 4b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | ■ | 10 | Group 1 | 1.17_4b | |
| J4500 | 2007 | NE | C | 4b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | ■ | 10 | Group 1 | 1.17_4b | |
| J5000 | 2008 | VA | C | 4b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | ■ | 10 | Group 1 | 1.17_4b | |
| J3182 | 2004 | OR | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 10 | P1 | | |
| J3751 | 2005 | OH | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 10 | P1 | | |
| J3976 | 2006 | MA | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 10 | P1 | 1.6_4b | |
| J4179 | 2006 | RI | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 10 | P1 | | |
| J4726 | 2007 | FL | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 10 | P1 | | |
| J5032 | 2008 | PA | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 10 | P1 | | |
| J3155 | 2004 | IL | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | 1.59_4b | |
| J3422 | 2005 | LA | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | | |
| J4432 | 2007 | PA | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | | |
| J4696 | 2007 | SD | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | | |
| J5421 | 2008 | FL | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | | |
| OLM 15 | 1934 | USA | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | | |
| OLM 18 | 1934 | USA | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | | |
| OLM 93 | 1954 | Canada | C | 4b | + | + | ☐ | ■ | ■ | ☐ | ☐ | ☐ | ☐ | 11 | P2 | | |
| 2007-629 | 2007 | NC | C | Indeterminate (4c) | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 12 | P3 | Lm3.45 | |
| NRRL B-33360 | N/A | IL | C | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 12 | P3 | Lm3.18 | (45) |
| NRRL B-33368 | 2003 | N/A | A | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 12 | P3 | Lm3.23 | (45) |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L0207 | 2003 | VA | E/F | 1/2b, 3b | + | + | □ | □ | □ | □ | □ | □ | □ | 13 | P4 | 1.34_1/2b | |
| L1210b | 2005 | VA | E/F | 1/2b, 3b | + | + | □ | □ | □ | □ | □ | □ | □ | 13 | P4 | | |
| 2007-582 | 2007 | NC | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| 4b1 | 1962 | N/A | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| 82a-2 | 2003 | NC | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| B068 (ScottA) | 1983 | MA | C | 4b | N | N | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | Lm1.1 | (12) |
| BUG912 | N/A | N/A | N/A | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| BUG914 | N/A | N/A | N/A | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| BUG922 | N/A | N/A | N/A | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| F7194 | 1989 | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| F7635 | 1989 | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| F7666 | 1989 | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| F7870 | 1989 | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| FDA 11 | 1987 | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| FDA 35 | 1988 | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| GA 311 | N/A | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| J2967 | 2004 | CA | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| J3290 | 2004 | ME | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| J3419 | 2005 | CA | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| J3768 | 2005 | CO | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| J3921 | 2006 | CT | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| LW-A104 | 2005 | N/A | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| OLM 120 | 1957 | Canada | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 14 | ECIa | | |
| FDA 107 | 1988 | N/A | E/F | 4b | + | N | □ | □ | □ | □ | ■ | ■ | □ | 14 | | 1.63_4b | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 128b-1 | 2004 | NC | E/F | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | (13) |
| 18-2a | 2003 | NC | E/F | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| 2001-7R | 2001 | NC | C | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| 2001-8R | 2001 | NC | C | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| 2006-296 | 2006 | NC | C | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| 2007-454 | 2007 | NC | C | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| 2008-894 | 2008 | NC | C | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| 363b-1 | 2005 | NC | E/F | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| 491a-5 | 2006 | NC | E/F | 4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | 15 | Group 3 | 1.60_4b | |
| J2621 | 2003 | OR | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | □ | 15 | P5 | 1.5_4b | (38) |
| J3917 | 2006 | CA | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | □ | 15 | P5 | 1.5_4b | |
| J4010 | 2006 | NE | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | □ | 15 | P5 | 1.5_4b | |
| J4109 | 2006 | NJ | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | □ | 15 | P5 | 1.5_4b | |
| 2007-618 | 2007 | NC | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | 16 | Group 2 | 1.46_4b | |
| J3026 | 2004 | CT | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | 16 | Group 2 | 1.46_4b | |
| J3053 | 2004 | MI | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | 16 | Group 2 | 1.46_4b | |
| J3195 | 2004 | OH | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | 16 | Group 2 | 1.46_4b | |
| J4458 | 2007 | MN | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | 16 | Group 2 | 1.46_4b | |
| J4490 | 2007 | SC | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | 16 | Group 2 | 1.46_4b | |
| J4953 | 2008 | OH | C | 4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | 16 | Group 2 | 1.46_4b | |
| #5 MI 6-1 | 2005 | MI | E/F | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 16 | P6 | | |
| J3085 | 2004 | MD | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 16 | P6 | | |
| J3728 | 2005 | NY | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 16 | P6 | 1.45_4b | |
| J3793 | 2005 | WV | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 16 | P6 | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J4291 | 2006 | NY | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 16 | P6 | | |
| NRRL B-33377 | 2004 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 17 | P7 | Lm3.31 | (45) |
| NRRL B-33379 | 2004 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 17 | P7 | Lm3.33 | (45) |
| NRRL B-33370 | 2003 | N/A | A | Indeterminate | + | + | □ | □ | □ | □ | □ | □ | □ | 18 | | Lm3.24 | (45) |
| NRRL B-33376 | 2004 | N/A | A | Indeterminate | + | + | □ | □ | □ | □ | □ | □ | □ | 19 | | Lm3.30 | (45) |
| OLM 65 | 1952 | Japan | A | 4b | - | + | ■ | ■ | ■ | □ | □ | □ | □ | 20 | ECI | 1.13_4b_Sw87_EC1 | |
| NRRL B-33182 | N/A | N/A | A | 1/2b, 3b (4c) | + | + | □ | □ | □ | □ | □ | □ | □ | 21 | | Lm3.6 | (45) |
| NRRL B-33405 | N/A | N/A | N/A | 1/2b, 3b (4c) | - | + | □ | □ | □ | □ | □ | □ | □ | 22 | | Lm4.1 | (45) |
| #171C | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #176A | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #18 | 2003 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #192A | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #201A | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #206A | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #209C | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #210A | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #24 | 2003 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| #265E | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 02-001 | 2002 | FL | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 101b-4 | 2003 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 143a-2 | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 158a-3 | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 171-b1 | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 176b-1 | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 18-4a | 2003 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 2001-82R | 2001 | NC | C | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 2002-119R | 2002 | NC | C | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 2002-119R BHA | 2002 | NC | C | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 2005-462 | 2005 | NC | C | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 2005-513 | 2005 | NC | C | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 2009-0037 | N/A | N/A | E/F | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 23 | | | |
| 2009-0038 | N/A | N/A | E/F | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 23 | | | |
| 2009-0039 | N/A | N/A | E/F | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 23 | | | |
| 2009-0040 | N/A | N/A | E/F | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 23 | | | |
| 2009-0041 | N/A | N/A | E/F | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 23 | | | |
| 201a-4 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 206a-5 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 210b-1 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 223b-4 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 231b-1 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 236a-1 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 239a-2 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 24-1b | 2003 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 26-1a | 2003 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 268b-2 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 279a-2 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 339b-5 | 2004 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |
| 367b-1 | 2005 | NC | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 23 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 369a-1 | 2005 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 494b-1 | 2006 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 506a-1 | 2006 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| 517a-1 | 2006 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| L1217 | 2005 | VA | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| L1624a | 2005 | VA | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| NC01-236R | 2001 | NC | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 23 | | | |
| NRRL B-33027 | 2000 | N/A | E/F | 1/2a, 3a | + | + | □ | □ | □ | □ | □ | □ | ■ | 23 | | Lm2.5 (T700) | (45) |
| NRRL B-33043 | N/A | N/A | E/F | 1/2a, 3a | + | + | □ | □ | □ | □ | □ | □ | ■ | 23 | | Lm2.11 | (45) |
| NRRL B-33371 | 2003 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 24 | | Lm3.25 | (45) |
| NRRL B-33185 | N/A | N/A | A | 4b | + | + | □ | □ | ■ | □ | ■ | □ | □ | 25 | | Lm3.9 | (45) |
| OLM 8 | 1933 | USA | A | 4b | + | + | □ | □ | ■ | □ | □ | □ | □ | 25 | | Lm3.42 | |
| 2007-640 | 2007 | NC | C | Indeterminate (4c) | + | + | □ | □ | ■ | □ | □ | □ | □ | 25 | | Lm3.50 | |
| NRRL B-33227 | N/A | N/A | N/A | 4c | + | + | □ | □ | □ | □ | □ | □ | □ | 25 | | Lm3.14 | (45) |
| NRRL B-33230 | N/A | N/A | E/F | 4c | + | + | □ | □ | □ | □ | □ | □ | □ | 25 | | Lm3.15 | (45) |
| NRRL B-33407 | N/A | N/A | N/A | 4c | + | + | □ | □ | □ | □ | □ | ■ | □ | 25 | | Lm3.39 | (45) |
| 04-454 | 2004 | NC | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 26 | | | |
| 2001-21R | 2001 | NC | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 26 | | | |
| 2002-327R | 2002 | NC | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 26 | | | |
| 2004-363 | 2004 | NC | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 26 | | | |
| 2005-397 | 2005 | NC | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 26 | | | |
| 2005-585 | 2005 | NC | C | 1/2a | + | P | □ | N | N | N | N | N | ■ | 26 | | | |
| 2006-644 | 2006 | NC | C | 1/2a | + | + | □ | N | N | N | N | N | ■ | 26 | | | |
| 2009-0084 | 2009 | NC | C | 1/2a | + | + | □ | □ | □ | □ | □ | □ | ■ | 26 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AT-01 | 1986 | N/A | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| BUG923 | N/A | N/A | N/A | 1/2a, 3a | - | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |
| BUG924 | N/A | N/A | N/A | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |
| DP-L862 | N/A | N/A | N/A | 1/2a | N | N | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |
| EGDe | 1926 | N/A | A | 1/2a | N | N | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |
| G2228 | N/A | N/A | C | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | |
| G3965 | 1988 | USA | E/F | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | (38) |
| G3975 | 1988 | USA | C | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | (38) |
| G3979 | 1988-1990 | USA | E/F | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | (38) |
| G3984 | 1981 | England | C | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | (38) |
| G3994 | 1981 | England | C | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | (38) |
| G3997 | 1988-1990 | USA | E/F | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | (38) |
| G4013 | 1988 | USA | E/F | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | | (38) |
| H7764 | N/A | N/A | C | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| H7788 | 1998 | N/A | E/F | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| J0161 | 2000 | USA | C | 1/2a | N | N | ☐ | N | N | N | N | N | ■ | 26 | | Lm2.1 | (38, 45) |
| L1009a | 2004 | VA | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| L1014a | 2004 | VA | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| L1508a | 2005 | VA | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| L1712a | 2005 | VA | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| LM 02-002 | 2002 | NC | C | 1/2a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| LW-A38 | 2003 | N/A | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| LW-A67 | 2001 | N/A | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |
| LW-A74 | 2003 | N/A | E/F | 1/2a, 3a | + | + | ☐ | N | N | N | N | N | ■ | 26 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| OLM 29 | 1938 | England | A | 1/2a, 3a | + | +/P | ☐ | N | N | N | N | N | ■ | 26 | | | |
| #1 MI 33 | 2004 | MI | E/F | 1/2b, 3b | + | - | ☐ | N | N | N | N | N | ☐ | 26 | | | |
| 414a-1 | 2005 | ND | E/F | 1/2b, 3b | + | - | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 26 | | | |
| BUG910 | N/A | N/A | N/A | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 26 | | | |
| BUG911 | N/A | N/A | N/A | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 26 | | | |
| FSL N1-017 | N/A | N/A | E/F | 1/2b, 3b (4b) | + | - | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 26 | | | |
| G3964 | N/A | Canada | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| G3978 | 1988-1990 | N/A | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| G4008 | N/A | Canada | E/F | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| G6003 | 1994 | IL | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | Lm1.18 | (12, 38) |
| G6004 | 1994 | IL | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| G6005 | 1994 | IL | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| G6006 | 1994 | IL | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| G6054 | 1994 | IL | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | Lm1.18 | (12, 38) |
| G6055 | 1994 | IL | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| G6060 | 1994 | IL | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| H3281 | N/A | N/A | C | 1/2b | + | - | ☐ | N | N | N | N | N | ☐ | 26 | | | (38) |
| L0607 | 2004 | VA | E/F | 1/2b, 3b | + | - | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 26 | | | |
| L0708 | 2004 | VA | E/F | 1/2b, 3b | + | - | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 26 | | | |
| LW-A96 | 2004 | N/A | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 26 | | | |
| #2 MI 40 | 2004 | MI | E/F | 1/2c, 3c | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |
| 2008-293 | 2008 | NC | C | 1/2c | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |
| 2008-911 | 2008 | NC | C | 1/2c, 3c | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |
| 283a-3 | 2004 | NC | E/F | 1/2c, 3c | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 26 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 516b-1 | 2006 | NC | E/F | 1/2c, 3c | + | + | □ | □ | □ | □ | □ | □ | ■ | 26 | | | |
| BUG803 | N/A | N/A | N/A | 1/2c, 3c | + | - | □ | □ | □ | □ | □ | □ | ■ | 26 | | | |
| BUG916 | N/A | N/A | N/A | 1/2c, 3c | + | + | □ | □ | □ | □ | □ | □ | ■ | 26 | | | |
| BUG920 | N/A | N/A | N/A | 1/2c, 3c | + | - | □ | □ | □ | □ | □ | □ | ■ | 26 | | | |
| NRRL B-33259 | 2001 | NJ | E/F | 1/2c, 3c | + | + | □ | □ | □ | □ | □ | □ | ■ | 26 | | Lm2.10 | (45) |
| NRRL B-33285 | 2001 | HI | E/F | 1/2c, 3c | + | + | □ | □ | □ | □ | □ | □ | ■ | 26 | | Lm2.10 | (45) |
| OLM 137 | 1960 | Canada | A | 1/2c, 3c | + | + | □ | □ | □ | □ | □ | □ | ■ | 26 | | | |
| #8 MI 15 | 2005 | MI | E/F | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | |
| 2009-0079 | 2009 | NC | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | |
| 2009-0117 | 2009 | NC | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | |
| 34-2b | 2003 | NC | E/F | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | |
| 34-6a | 2003 | NC | E/F | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (13) |
| Beef (franks) | 1992 | N/A | E/F | 4b | + | N | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | |
| H3396 | 1997 | CA | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| H6383 | 1996 | OH | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| H7550 | 1998 | USA | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38, 46) |
| H7738 | 1998 | USA | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | |
| H7762 | 1998 | USA | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| J1815 | 2002 | N/A | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| J1925 | 2002 | N/A | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| J2206 | 2003 | NJ | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| J2230 | 2003 | MA | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| J2446 | 2003 | OH | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |
| J2685 | 2004 | NY | C | 4b | + | + | □ | □ | □ | ■ | ■ | ■ | □ | 26 | ECII | | (38) |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3006 | 2004 | TX | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (38) |
| J3033 | 2004 | IL | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (38) |
| J3075 | 2004 | CT | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3170 | 2004 | MI | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3200 | 2004 | CT | C | 4b | P | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (38) |
| J3333 | 2004 | MN | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3527 | 2005 | WI | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3565 | 2005 | IL | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3606 | 2005 | MI | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3785 | 2005 | MA | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3840 | 2005 | MD | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J3881 | 2006 | CO | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J4101 | 2006 | WV | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J4105 | 2006 | AL | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J4120 | 2006 | OH | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J4210 | 2006 | GA | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| J4485 | 2007 | MA | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| L0226 | 2003 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0315 | 2003 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0327 | 2003 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0603 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0623 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0704 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0719 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L0720 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0724 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | (13) |
| L0928 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| L0929 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| L0930 | 2004 | VA | E/F | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| NRRL B-33589 | 2003 | MD | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| NRRL B-33591 | 2003 | GA | C | 4b | + | + | ☐ | ☐ | ☐ | ■ | ■ | ■ | ☐ | 26 | ECII | | |
| 2007-583 | 2007 | NC | C | 4b | + | - | ☐ | ☐ | ☐ | ☐ | ■ | ■ | ☐ | 26 | ECV | | |
| J3215 | 2004 | NJ | C | 4b | + | - | ☐ | ☐ | ☐ | ☐ | ■ | ■ | ☐ | 26 | ECV | | |
| J4045 | 2006 | MO | C | 4b | + | - | ☐ | ☐ | ☐ | ☐ | ■ | ■ | ☐ | 26 | ECV | | |
| WS1 | 2000 | NC | C | 4b | + | - | ☐ | ☐ | ☐ | ☐ | ■ | ■ | ☐ | 26 | ECV | | (24) |
| #158C | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| #171D | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| #192C | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| #201B | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| #210B | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| #4 MI 16 | 2004 | MI | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| 128b-2 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| 158b-2 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| 179b-3 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 192b-1 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 2005-625 | 2005 | NC | C | 1/2b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| 2007-269 | 2007 | NC | C | 1/2b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | 1.29_1/2b | |
| 2007-549 | 2007 | NC | C | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2008-018 | 2007 | NC | C | 1/2b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 201a-1 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 223a-1 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 231a-1 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 232b-1 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 258a-1 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 326b-5 | 2004 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 34-2a | 2003 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| 360b-1 | 2005 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 369a-4 | 2005 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 506a-3 | 2006 | NC | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| 91a-1 | 2003 | NC | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| AT-05 | 1987 | N/A | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| BUG915 | N/A | N/A | N/A | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| G4007 | N/A | United Kingdom | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 27 | | | (38) |
| G4598 | N/A | Italy | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 27 | | | (38) |
| G4599 | N/A | Italy | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 27 | | | (38) |
| G4600 | N/A | Italy | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 27 | | | (38) |
| G4601 | N/A | Italy | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 27 | | | (38) |
| G4602 | N/A | Italy | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 27 | | | (38) |
| L1214a | 2005 | VA | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| L1619a-1 | 2005 | VA | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | | |
| LW-A22 | 2001 | N/A | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| LW-A34 | 2003 | N/A | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LW-A42 | 2001 | N/A | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 27 | | | |
| NRRL B-33005 | 1985 | N/A | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | Lm1.30 | (12) |
| NRRL B-33028 | 2000 | N/A | E/F | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 27 | | Lm1.25 | (12) |
| 2003-184R | 2003 | NC | C | 1/2a | - | + | ☐ | N | N | N | N | N | ■ | 28 | | 2.69_1/2a | |
| 2009-0045 | 2009 | NC | C | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 28 | | 2.43_1/2a | |
| OLM 26 | 1938 | England | A | 1/2a, 3a | P | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 28 | | 2.60_1/2a | |
| OLM 67 | 1953 | Canada | A | 1/2a, 3a | - | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 28 | | 2.60_1/2a | |
| OLM 72 | 1937 | Canada | A | 1/2a, 3a | - | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 28 | | 2.60_1/2a | |
| OLM 73 | 1953 | Canada | A | 1/2a, 3a | - | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ■ | 28 | | 2.60_1/2a | |
| OLM 5 | 1933 | USA | A | 1/2a, 3a | + | + | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | Lm3.16 | |
| OLM 6 | 1933 | USA | A | 1/2a, 3a | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | | |
| FSL J1-194 | 1997 | N/A | C | 1/2b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | | |
| G3985 | N/A | USA | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 28 | | 1.20_1/2b | (38) |
| G4009 | N/A | USA | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 28 | | | (38) |
| G4025 | N/A | USA | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 28 | | | (38) |
| G4027 | N/A | USA | C | 1/2b | N | N | ☐ | N | N | N | N | N | ☐ | 28 | | | (38) |
| LW-A94 | 2004 | N/A | E/F | 1/2b, 3b | + | + | ☐ | N | N | N | N | N | ☐ | 28 | | 1/20_1/2b | |
| OLM 136* | 1960 | LA | C | 1/2b, 3b | + | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | 1.72_1/2b | |
| J2571 | 2003 | KY | C | 4b | - | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 28 | | Lm3.42 | (38) |
| FSL J2-071 | N/A | N/A | A | 4c | + | + | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | | |
| NRRL B-33115 | N/A | N/A | A | 4c | + | + | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | Lm3.16 | (45) |
| NRRL B-33231 | N/A | N/A | E/F | 4c | + | + | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | Lm3.16 | (45) |
| OLM 84 | 1949 | MD | C | Indeterminate | + | + | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | ☐ | 28 | | Lm3.16 | |
| NRRL B-33406 | N/A | N/A | N/A | 1/2b, 3b (4a) | - | + | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | 29 | | Lm4.2 | (45) |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NRRL B-33425 | N/A | N/A | A | 4e/d | + | + | □ | □ | ■ | □ | □ | □ | □ | 29 | | Lm3.40 | (45) |
| NRRL B-33364 | 2002 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 29 | | Lm3.21 | (45) |
| NRRL B-33366 | 2002 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 29 | | Lm3.21 | (45) |
| NRRL B-33367 | 2002 | N/A | A | 1/2a, 3a | + | + | □ | □ | □ | □ | □ | □ | ■ | 30 | | Lm3.22 | (45) |
| NRRL B-33077 | N/A | N/A | A | 4b | + | - | □ | □ | ■ | □ | □ | □ | □ | 30 | | Lm3.1 | (45) |
| J3017 | 2004 | IN | C | Indeterminate (4b) | + | + | □ | □ | □ | □ | □ | □ | □ | 30 | | Lm3.42 (IV[g]) | |
| NRRL B-33188 | N/A | N/A | A | 4c | + | + | □ | ■ | □ | □ | □ | □ | □ | 30 | | Lm3.11 | (45) |
| #1 MI 18 | 2004 | MI | E/F | 1/2a, 3a | - | + | □ | N | N | N | N | N | ■ | 31 | | | |
| #266A | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 31 | | | |
| 179b-1 | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 31 | | | |
| 1a (255) | N/A | N/A | N/A | 1/2a, 3a | - | + | □ | N | N | N | N | N | ■ | 31 | | | |
| 2001-25R | 2001 | NC | C | 1/2a | - | + | □ | N | N | N | N | N | ■ | 31 | | | |
| 266a-1 | 2004 | NC | E/F | 1/2a, 3a | + | + | □ | N | N | N | N | N | ■ | 31 | | | |
| BUG909 | N/A | N/A | N/A | 1/2a, 3a | + | + | □ | □ | □ | □ | □ | □ | ■ | 31 | | | |
| LW-A35 | 2003 | N/A | E/F | 1/2a | - | + | □ | N | N | N | N | N | ■ | 31 | | | |
| 2003-123R | 2003 | NC | C | 1/2b | + | + | □ | N | N | N | N | N | □ | 31 | | | |
| 2009-0042 | 2009 | NC | C | 1/2b | + | + | □ | □ | □ | □ | □ | □ | □ | 31 | | | |
| 2009-0050 | 2009 | NC | C | 1/2b | + | + | □ | □ | □ | □ | □ | □ | □ | 31 | | | |
| LW-A39 | 2003 | N/A | E/F | 1/2b, 3b | + | + | □ | N | N | N | N | N | □ | 31 | | | |
| NRRL B-33381 | 2004 | N/A | A | 1/2b, 3b | + | + | □ | □ | ■ | □ | □ | □ | □ | 31 | | Lm3.34 | (45) |
| J2422 | 2003 | RI | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | 1.7_4b | (38) |
| J3070 | 2004 | TN | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | 1.62_4b | |
| J3120 | 2004 | WA | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |
| J3309 | 2005 | IA | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |

## Table 3.1 Continued

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3554 | 2005 | MS | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |
| J3618 | 2005 | NJ | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |
| J3684 | 2005 | OR | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |
| J3762 | 2005 | IN | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |
| J3795 | 2005 | AZ | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |
| J3909 | 2006 | OR | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | | |
| J4428 | 2007 | OH | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 31 | | 1.62_4b | |
| J4221 | 2006 | MD | C | 4b | + | + | □ | □ | □ | □ | ■ | ■ | □ | 31 | | 1.75_4b | |
| NRRL B-33183 | N/A | N/A | A | 4b | + | + | □ | □ | □ | □ | □ | □ | □ | 31 | | Lm3.7 | (45) |
| NRRL B-33363 | 2002 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 31 | | Lm3.19 | (45) |
| NRRL B-33378 | 2004 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 31 | | Lm3.32 | (45) |
| J3415 | 2005 | TX | C | 4b | + | + | □ | ■ | ■ | □ | □ | □ | □ | 32 | | 1.7_4b | |
| CLIP 11262 | 1988 | Morocco | E/F | *L. innocua* | N | N | □ | □ | □ | □ | □ | □ | □ | 33 | | | |
| NRRL B-33314 | 2001 | CA | E/F | *L. innocua* | + | + | □ | □ | □ | □ | □ | □ | □ | 33 | | | |
| J5031[i] | 2008 | GA | C | 1/2a, 3a (4b) | + | + | □ | ■ | ■ | □ | □ | □ | ■ | 33 | | Lm3.39 | |
| L0703[i] | 2004 | VA | E/F | 1/2b, 3b | + | - | □ | □ | □ | □ | □ | □ | □ | 33 | | 1.37_1/2b | |
| NRRL B-33160[i] | N/A | N/A | E/F | 1/2b, 3b | + | + | □ | □ | □ | □ | □ | □ | □ | 33 | | Lm1.37 | (46) |
| 2003-151R[i] | 2003 | NC | C | Indeterminate (4b) | - | + | □ | □ | ■ | □ | □ | □ | □ | 33 | | Lm3.42 | |
| NRRL B-33229[i] | N/A | N/A | C | 4a /4c | - | + | □ | □ | ■ | □ | □ | □ | □ | 33 | | Lm3.38 | (45) |
| NRRL B-33403[i] | N/A | N/A | C | 4a | - | + | □ | □ | ■ | □ | □ | □ | □ | 33 | | Lm3.38 | (45) |
| OLM 59[j] | 1953 | Canada | C | Indeterminate | - | + | □ | □ | □ | □ | □ | □ | □ | 33 | | Lm4.2 | |
| NRRL B-33181[k] | N/A | N/A | A | 4b | - | + | □ | □ | ■ | □ | □ | □ | □ | 33 | | Lm3.5 | (45) |
| NRRL B-33187[k] | N/A | N/A | A | 4b | + | + | □ | □ | ■ | □ | □ | □ | □ | 33 | | Lm3.10 | (45) |
| NRRL B-33362[k] | 2002 | N/A | A | Indeterminate | + | + | □ | □ | ■ | □ | □ | □ | □ | 33 | | Lm3.27 | (45) |

**Table 3.1 Continued**

| Name | Isolation Year | Country/State | Source[b] | Serotype[c] | Sau3AI[d] | MboI[d] | 85M[e] | Non-ECIIC-WAP[e] | 4bSF18[e] | Hsp001219[e] | ECIIC-WAP[e] | H18RP9/10[e] | lmo0737[e] | Region 85 Hybridization Pattern | Clonal Group | MLGT | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| NRRL B-33373[k] | 2004 | N/A | A | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 33 | | Lm3.27 | (45) |
| NRRL B-33105[l] | N/A | N/A | A | 4b | + | - | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.3 | (45) |
| OLM 37[l] | 1937 | USA | A | 4b | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.52 | |
| OLM 21[l] | 1937 | USA | A | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.42 | |
| OLM 22[l] | 1937 | New Zealand | A | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.42 | |
| OLM 82[l] | 1953 | Germany | N/A | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.42 | |
| NRRL B-33375[l] | 2004 | N/A | A | Indeterminate | + | + | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.29 | (45) |
| 2001-75R[m] | 2001 | NC | C | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.2 | |
| NRRL B-33092[n] | N/A | N/A | A | 4b | + | + | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.2 | (45) |
| NRRL B-33426[o] | N/A | N/A | E/F | 4b | + | - | ☐ | ☐ | ■ | ☐ | ☐ | ☐ | ☐ | 34 | | Lm3.41 | (45) |
| NRRL B-33384[o] | 2004 | N/A | A | Indeterminate | + | + | ☐ | ☐ | ■ | ☐ | ■ | ☐ | ☐ | 34 | | Lm3.37 | (45) |

[a] N/A, information unknown.

[b] A, animal isolates; C, clinical isolates; and E/F, isolates from the environment or foods.

[c] The serotypes were determined with the serological method performed in the institutes that the isolates were collected and/or multiplex PCR (11) performed in our lab. When serotypes determined by both methods differed, the serotype obtained with the multiplex PCR was used for analysis and the serotype obtained with the serological method was shown within parentheses. When the serotype could not be determined by this multiplex PCR, the isolate was designated to the indeterminate serotype. The plus and

**Table 3.1 Continued**

minus signs represent susceptibility and resistance of the genomic DNA to a restriction enzyme, respectively. The letter P signifies partial digestion of the genomic DNA.

[e] White and black rectangles mark the absence and presence, respectively, of the signal in the DNA-DNA hybridization using a DNA probe.

[f] N, not tested.

[g] Lineage determined by TMLGT.

[h] Alternative strain name.

[i] Isolates that generated 0.6 kb amplicon from primers 2007-640LPF/2007-640LPR targeting probe 2007-640LP and 1.5 kb amplicon from primers H7858_0334R/H7858_0338F targeting the flanking genes. Since strain 2003-151R exhibited the same results, these isolates seem to harbor Region 85 similar to 2003-151R.

[j] Isolate that generated 0.5 kb amplicon from primers 2007-640LPF/2007-640LPR and 2.2 kb amplicon from primers H7858_0334R/H7858_0338F.

[k] Isolates that generated 0.6 and 1.2kb amplicons from primers 2007-640LPF/2007-640LPR and 7 kb amplicon from primers H7858_0334R/H7858_0338F.

**Table 3.1 Continued**

[l] Isolates that generated no amplicon from primers 2007-640LPF/2007-640LPR and 2 kb amplicon from primers H7858_0334R/H7858_0338F.

[m] Isolate that generated no amplicon from primers 2007-640LPF/2007-640LPR and 2.8 kb amplicon from primers H7858_0334R/H7858_0338F.

[n] Isolate that generated no amplicon from primers 2007-640LPF/2007-640LPR and 4 kb amplicon from primers H7858_0334R/H7858_0338F.

[o] Isolates that generated no amplicon from primers 2007-640LPF/2007-640LPR and 7 kb amplicon from primers H7858_0334R/H7858_0338F.

**Table 3.2 Number of isolates belonging to different serotypes and sources**

| | 1/2a, 3a | 1/2b, 3b | 1/2c, 3c | 4b | Others[a] | Total |
|---|---|---|---|---|---|---|
| A | 10 | 3 | 1 | 16 | 28 | 58 |
| C | 28 | 29 | 2 | 145 | 11 | 215 |
| E/F | 71 | 45 | 5 | 43 | 6 | 170 |
| N/A | 5 | 5 | 3 | 6 | 3 | 22 |
| Total | 114 | 82 | 11 | 210 | 48 | 465 |

[a] The "other" serotype category includes two *L. innocua* isolates, *L. monocytogenes* isolates belonging to lineage III and IV serotypes (4a and 4c) and *L. monocytogenes* isolates whose serotypes could not be determined by the multiplex PCR devised by Doumith et al. (11).

**Table 3.3 List of isolates whose Region 85 was sequenced**

| Name | Estimated amplicon size (sequenced DNA size) (bp) | GenBank accession number | Serotype | Region 85 Hybridization Pattern |
|---|---|---|---|---|
| 2003-151R | 1596 (1471)[a] | JN235996 | Indeterminate | 33 |
| 2007-640 | 3140 (2986)[a] | JN235991 | Indeterminate (4c) | 25 |
| 2001-75R | 3328 (2826)[a] | JN235995 | Indeterminate | 34 |
| 2007-584 | 3944 (865, 858)[b] | JN235999, JN236000 | 4b | 5 |
| J3115 | 4573 (3697)[a] | JN235992 | 4b | 5 |
| J2479 | 4850 (4532)[a] | JN235993 | 4b | 4 |
| J2571 | 5967 | | 4b | 28 |
| J3017 | 7080 | | Indeterminate (4b) | 30 |
| FSL J1-208 | 7408 (6631)[a] | JN235994 | 4a | 1 |
| 2007-0904 | 11333 | | 4b | 10 |
| J2255 | 11603 | | 4b | 10 |
| J3139 | 11828 | | 4b | 10 |
| J3182 | 12088 (789)[c] | JN236001 | 4b | 10 |
| J3155 | 12354 | | 4b | 11 |
| J3422 | 12626 (820, 868)[b] | JN236002, JN236003 | 4b | 11 |
| J3720 | 12626 | | 4b | 7 |
| 2007-629 | 13660 | | Indeterminate (4c) | 12 |
| J4016 | 13814 | | 4b | 10 |
| J4179 | 14010 (869, 886)[b] | JN236004, JN236005 | 4b | 10 |
| J3913 | 14078 | | 4b | 9 |
| J3751 | 14279 | | 4b | 10 |
| J3976 | 14279 (2247, 9302)[d] | JN235997, JN235998 | 4b | 10 |

[a] Region 85 amplicon was completely sequenced.

[b] Only the terminal parts of the Region 85 amplicon were sequenced. Within the parentheses, the first and second numbers represent the size of the 5' and 3' terminal fragments, respectively.

[c] Only the 5' terminal parts of the Region 85 amplicon was sequenced.

[d] The sequence is available for the internal part (2247 bp) including the *hsdS* gene and its flanking regions and the 3' terminal fragment of 9302 bp.

**Table 3.4 Primers and DNA probes used in this study**

| Probe | Primer | Sequence (5' to 3') | Target | Reference |
|---|---|---|---|---|
| | H7858_0334R | GTTCCCGAATCATTTCCAC | *LMOh7858_0334* | This study |
| | H7858_0338F | CTCGTGAATCTCCAAATGCG | *LMOh7858_0338* | This study |
| | LMSG_01575F | AGTTCAGTGGAGAACCGTAG | *LMSG_03070* | This study |
| | LMSG_03069R | CTAGATTCTCCACCACGAAC | *LMSG_03069* | This study |
| **<Region 85-based probes>** | | | | |
| 85M | F2365_85MF | AATATATTTTCAATGTTTGATGGT | *LMOf2365_0327* | (50) |
| | F2365_85MR | GCTAATTCAATCCCTATTCT | *LMOf2365_0327* | (50) |
| J2479Res | J2479Res_F | GACCTAGTAAAGCAGGTGCT | *lmoJ2R* | This study |
| | J2479_3R | AGGTACCCATTCGATAGTCG | *lmoJ2R* | This study |
| J3115Met | J3115_2F | TATCAGGCTTTCCCTGTCAA | *lmoJ3M* | This study |
| | J3115Met_R | GGTACTACAACCGAATTCCC | *lmoJ3M* | This study |
| LMSG_01573 | LMSG_01573F | TACAATTGGTCGGACACGTG | *LMSG_01573* | This study |
| | LMSG_01573R | AGAATCCGCTCATAAACAGC | *LMSG_01573* | This Study |
| LMSG_01574C | LMSG_01574_CF | CTAGACAACACTATCGCTC | *LMSG_01574* | This Study |
| | LMSG_01574_CR | ACGCTGTTCCCAATCATCAG | *LMSG_01574* | This Study |
| LMSG_01574 | LMSG_01574F | CCTCTAATTCGAGTATCC | *LMSG_01574* | This Study |
| | LMSG_01574R | GTCCCCTACAGCAACACCTTT | *LMSG_01574* | This Study |
| J3976TypeIS_V | J3976_TypeIS_VF | GCAAGTCTAAATCTACCAG | *LmoJ3976S* | This Study |
| | J3976_TypeIS_VR | ATGACTCTGCGGCAAGTAC | *LmoJ3976S* | This Study |
| Mrr | LMSG_01576F | CAGAATGGGATATAGCGGAG | *LMSG_01576* | This Study |
| | LMSG_01576R | CAGCTTCCACTCCAACTTTG | *LMSG_01576* | This Study |
| McrB | H7858_0337F | ATATCCATGCCCATCACCAC | *LMOh7858_0337* | This study |
| | H7858_0337R | CGGGAGGAATCTCGTTATAC | *LMOh7858_0337* | This study |
| LMSG_01578 | LMSG_01578F | GCTTGCAGGAAAGAAGGAAC | *LMSG_03073* | This Study |
| | LMSG_01578R | GGTTTTCTTTCGCAAGTGC | *LMSG_03073* | This Study |

**Table 3.4 Continued**

| Probe | Primer | Sequence (5' to 3') | Target | Reference |
|---|---|---|---|---|
| LP | LPF | GTTGTTGATAGTAAGTATA | *LMOf2365_0329* | This Study |
| | LPR | TCTGCTTTGACATCTATGTG | *LMOf2365_0329* | This Study |
| J3976NG | J3976NG_F | GCAGCATTAGTAGTTTGTAGC | *J3976NG* | This Study |
| | J3976_01575LP_3R | TGAAGTCTATAGTCTCCGGC | *J3976NG* | This Study |
| 2007-640LP | 2007-640_LPF | GTTAGTAGGATGTGGTAAAG | *lin0331* | This Study |
| | 2007-640_LPR | CTGTTCTCTTTGTCATCCAC | *lin0331* | This Study |
| J1-194_1 | J1-194_1 F | CGAGGTCTAGGACATTGTTC | Intergenic region between *LMBG_02168* and *LMBG_02169* | This Study |
| | J1-194_1 R | CATATGGTTACCGTCACTAG | *LMBG_02168* | This Study |
| **<EC probes>** | | | | |
| Non-ECIIC-WAP | ECIC-WAPF | ATGGAAATTGGGCATGGC | *LMOf2365_0450* | (6) |
| | ECIC-WAPR | GTAGTTCCAGTGGACATG | *LMOf2365_0450* | (6) |
| 4bSF18 | 4bSF18_F | ACGGGCGTTTTATATTAAATGGG | *LMOf2365_0466* | (14) |
| | 4bSF18_R | AATATCTCGAAAACTCCGAGT | *LMOf2365_0466* | (14) |
| Hsp001219 | Hspec01219F | GAGGCTATCGAAATTGCTCG | *LMOh7858_1168* | (6) |
| | Hspec01219R | AGGATTCGGAATTTCATCCA | *LMOh7858_1168* | (6) |
| ECIIC-WAP | ECIIC-WAPF | GGGAACTTTCCATTAGCC | *LMOh7858_0479* | (6) |
| | ECIIC-WAPR | TTAAATGGGATATGATGT | *LMOh7858_0479* | (6) |
| H18RP9/10 | H18RP9/10F | CTGGATTTGCAGCTTATGAT | *LMOh7858_0487* | (21) |
| | H18RP9/10R | CCTATTCTTTCCATAAGTAAT | *LMOh7858_0487* | (21) |
| **<DNA probes for atypical serotype 4b isolates>** | | | | |
| lmo0737 | lmo0737F | AGGGCTTCAAGGACTTACCC | *lmo0737* | (11) |
| | lmo0737R | ACGATTTCTGCTTGCCATTC | *lmo0737* | (11) |

**Table 3.5 Pair-wise comparisons of lipoproteins**

| *2003-151R* | *lin0331* | *LMHCC_2327* | *LMOf2365_0329* | *LMHCC_2328* | DNA/ Protein |
|---|---|---|---|---|---|
| 100/100[a] | 78.6/ 73.8 | 76.8/ 61.5 | 69.5/ 58.2 | 67.7/ 50.8 | *2003-151R* |
| 78.6/ 73.8 | 100/100 | 83.2/ 73.8 | 77.3/ 67.4 | 59.3/ 40.0 | *lin0331* |
| 76.8/ 61.5 | 83.2/ 73.8 | 100/100 | 74.0/ 63.0 | 59.6/ 42.6 | *LMHCC_2327* |
| 69.5/ 58.2 | 77.3/ 67.4 | 74.0/ 63.0 | 100/100 | 58.6/ 41.3 | *LMOf2365_0329* |
| 67.7/ 50.8 | 59.3/ 40.0 | 59.6/ 42.6 | 58.6/ 41.3 | 100/100 | *LMHCC_2328* |

[a] The numbers in front of and after the slash mark represent the percent identities at the DNA and protein levels, respectively, which were obtained with EMBOSS Align.

**Table 3.6 Distribution of the isolates positive for the Region 85 gene probes**

| | A (%) | C (%) | E/F (%) | N/A (%) | 1/2a, 3a (%) | 1/2b, 3b (%) | 1/2c, 3c (%) | 4b (%) | Others (%) | Total (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| 85M | 12 (20.7) | 38 (17.7) | 22 (12.9) | 3 (13.6) | 8 (7.0) | 4 (4.9) | 0 | 56 (26.7) | 7 (14.6) | 75 (16.1) |
| J2479Res | 3 (5.2) | 6 (2.8) | 5 (2.9) | 0 | 4 (3.5) | 0 | 0 | 6 (2.9) | 4 (8.3) | 14 (3.0) |
| J3115Met | 0 | 3 (1.4) | 0 | 0 | 1 (0.9) | 0 | 0 | 2 (1.0) | 0 | 3 (0.6) |
| LMSG_01573 | 8 (13.8) | 53 (24.7) | 17 (10.0) | 3 (13.6) | 1 (0.9) | 2 (2.4) | 0 | 71 (33.8) | 7 (14.6) | 81 (17.4) |
| LMSG_01574C | 8 (13.8) | 52 (24.2) | 17 (10.0) | 3 913.6) | 1 (0.9) | 2 (2.4) | 0 | 70 (33.8) | 7 (14.6) | 80 (17.2) |
| LMSG_01574 | 2 (3.4) | 29 (13.5) | 43 (25.3) | 4 (18.2) | 1 (0.9) | 40 (48.4) | 0 | 36 (17.1) | 1 (2.1) | 78 (16.8) |
| J3976TypeIS_V | 0 | 14 (6.5) | 0 | 0 | 0 | 0 | 0 | 14 (6.7) | 0 | 14 (3.0) |
| Mrr | 3 (5.2) | 60 (27.9) | 59 (34.7) | 4 (18.2) | 49 (43.0) | 4 (4.9) | 0 | 70 (33.3) | 3 (6.3) | 126 (27.1) |
| McrB | 15 (25.9) | 124 (57.7) | 118 (69.4) | 11 (50.0) | 84 (73.7) | 60 (73.2) | 11 (100) | 104 (49.5) | 9 (18.8) | 268 (57.6) |
| LMSG_01578 | 0 | 36 (16.7) | 57 (33.5) | 3 (13.6) | 49 (43.0) | 0 | 0 | 47 (22.4) | 0 | 96 (20.6) |
| LP | 14 (24.1) | 163 (75.8) | 157 (92.4) | 17 (77.3) | 97 (85.1) | 69 (84.1) | 11 (100) | 167 (79.5) | 7 (14.6) | 351 (75.5) |
| J3976NG | 3 (5.2) | 24 (11.2) | 0 | 1 (4.5) | 0 | 2 (2.4) | 0 | 23 (11.0) | 3 (6.3) | 28 (6.0) |
| 2007-640LP | 24 (41.4) | 32 (14.9) | 5 (2.9) | 4 (18.2) | 4 (3.5) | 5 (6.1) | 0 | 31 (14.8) | 25 (52.1) | 61 (13.1) |
| J1-194_1 | 14 (24.1) | 11 (5.1) | 2 (1.2) | 1 (4.5) | 9 (7.9) | 8 (9.8) | 0 | 2 (1.0) | 9 (18.8) | 28 (6.0) |
| Total | 58 | 215 | 170 | 22 | 114 | 82 | 11 | 210 | 48 | 465 |

**Table 3.7 Distribution of the different hybridization profiles of Region 85 probes**

| Region 85 Hybridization pattern | 85M | J2479Res | J3115Met | LMSG_01573 | LMSG_01574C | LMSG_01574 | J3976TypeIS_V | Mrr | McrB | LMSG_01578 | LP | J3976NG | 2007-640LP | J1-194_1 | 1/2a, 3a (%) | 1/2b, 3b (%) | 1/2c, 3c (%) | 4b (%) | Others (%) | Total (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | 0 | 1 (1.2)[a] | 0 | 0 | 1 (2.1)[a] | 2 (0.4) |
| 2 | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | 1 (0.9) | 0 | 0 | 2 (1.0)[a] | 5 (10.4)[a] | 8 (1.7) |
| 3 | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | 7 (6.1) | 3 (3.7) | 0 | 53 (25.2)[b] | 1 (2.1) | 64 (13.8) |
| 4 | □ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | 4 (3.5) | 0 | 0 | 6 (2.9) | 4 (8.3)[a] | 14 (3.0) |
| 5 | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | 1 (0.9) | 0 | 0 | 2 (1.0) | 0 | 3 (0.6) |
| 6 | □ | □ | □ | ■ | ■ | ■ | □ | □ | ■ | □ | □ | □ | ■ | □ | 1 (0.9)[a] | 0 | 0 | 0 | 0 | 1 (0.2)[a] |
| 7 | □ | □ | □ | ■ | ■ | ■ | □ | ■ | ■ | □ | □ | ■ | ■ | □ | 0 | 0 | 0 | 1 (0.5)[a] | 0 | 1 (0.2)[a] |
| 8 | □ | □ | □ | ■ | ■ | □ | □ | □ | ■ | □ | □ | □ | ■ | □ | 0 | 0 | 0 | 1 (0.5)[a] | 0 | 1 (0.2)[a] |
| 9 | □ | □ | □ | ■ | □ | □ | ■ | ■ | ■ | □ | □ | ■ | ■ | □ | 0 | 0 | 0 | 1 (0.5)[c] | 0 | 1 (0.2) |
| 10 | □ | □ | □ | ■ | ■ | □ | ■ | ■ | ■ | □ | □ | ■ | ■ | □ | 0 | 0 | 0 | 12 (5.7)[c] | 0 | 12 (2.6) |
| 11 | □ | □ | □ | ■ | ■ | □ | □ | ■ | ■ | □ | □ | ■ | ■ | □ | 0 | 0 | 0 | 8 (3.8) | 0 | 8 (1.7) |
| 12 | □ | □ | □ | ■ | ■ | □ | □ | ■ | □ | □ | □ | ■ | ■ | □ | 0 | 0 | 0 | 0 | 3 (6.3)[a] | 3 (0.6)[a] |
| 13 | □ | □ | □ | ■ | ■ | □ | □ | ■ | □ | □ | ■ | □ | □ | □ | 0 | 2 (2.4) | 0 | 0 | 0 | 2 (0.4) |
| 14 | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | ■ | ■ | □ | □ | □ | 0 | 0 | 0 | 22 (10.5)[d] | 0 | 22 (4.7) |
| 15 | □ | □ | □ | ■ | ■ | ■ | □ | ■ | ■ | ■ | ■ | □ | □ | □ | 0 | 0 | 0 | 13 (6.2)[e] | 0 | 13 (2.8) |
| 16 | □ | □ | □ | ■ | ■ | □ | □ | ■ | ■ | ■ | ■ | □ | □ | □ | 0 | 0 | 0 | 12 (5.7)[f] | 0 | 12 (2.6) |
| 17 | □ | □ | □ | ■ | ■ | □ | □ | □ | ■ | □ | ■ | □ | □ | □ | 0 | 0 | 0 | 0 | 2 (4.2)[a] | 2 (0.4)[a] |

**Table 3.7 Continued**

| Region 85 Hybridization pattern | 85M | J2479Res | J3115Met | LMSG_01573 | LMSG_01574C | LMSG_01574 | J3976TypeIS_V | Mrr | McrB | LMSG_01578 | LP | J3976NG | 2007-640LP | J1-194_1 | 1/2a, 3a (%) | 1/2b, 3b (%) | 1/2c, 3c (%) | 4b (%) | Others (%) | Total (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 18 | □ | □ | □ | ■ | ■ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | 0 | 0 | 0 | 0 | 1 (2.1)[a] | 1 (0.2)[a] |
| 19 | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | ■ | □ | □ | □ | 0 | 0 | 0 | 0 | 1 (2.1)[a] | 1 (0.2)[a] |
| 20 | ■ | □ | □ | ■ | ■ | □ | □ | ■ | ■ | □ | ■ | ■ | ■ | □ | 0 | 0 | 0 | 1 (0.5) | 0 | 1 (0.2) |
| 21 | □ | □ | □ | □ | □ | □ | □ | ■ | ■ | □ | □ | ■ | ■ | □ | 0 | 1 (1.2)[a] | 0 | 0 | 0 | 1 (0.2)[a] |
| 22 | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | ■ | ■ | □ | 0 | 2 (1.2)[a] | 0 | 0 | 0 | 1 (0.2)[a] |
| 23[g] | □ | □ | □ | □ | □ | □ | □ | ■ | ■ | ■ | ■ | □ | □ | □ | 49 (43.0) | 0 | 0 | 0 | 0 | 49 (10.5) |
| 24 | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | 0 | 0 | 0 | 0 | 1 (2.1)[a] | 1 (0.2)[a] |
| 25 | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | ■ | □ | 0 | 0 | 0 | 2 (1.0)[a] | 4 (8.3)[a] | 6 (1.3)[a] |
| 26[h] | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | ■ | □ | □ | □ | 33 (28.9) | 19 (23.2) | 11 (100.0) | 52 (24.8)[i] | 0 | 115 (24.7) |
| 27[h] | □ | □ | □ | □ | □ | ■ | □ | □ | ■ | □ | ■ | □ | □ | □ | 0 | 40 (48.8) | 0 | 0 | 0 | 40 (8.6) |
| 28 | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | 8 (7.0) | 7 (8.5) | 0 | 1 (0.5)[a] | 4 (8.3)[a] | 20 (4.3) |
| 29 | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | ■ | 0 | 1 (1.2)[a] | 0 | 0 | 3 (6.3)[a] | 4 (0.9)[a] |
| 30 | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | ■ | ■ | 1 (0.9)[a] | 0 | 0 | 1 (0.5)[a] | 2 (4.2)[a] | 4 (0.9)[a] |
| 31[j] | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | 8 (7.0) | 5 (6.1) | 0 | 13 (6.2) | 2 (4.2)[a] | 28 (6.0) |
| 32[j] | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | ■ | □ | □ | □ | 0 | 0 | 0 | 1 (0.5) | 0 | 1 (0.2) |
| 33 | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | 1 (0.9)[a] | 2 (2.4) | 0 | 2 (1.0)[a] | 8 (16.7) | 13 (2.8) |
| 34 | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | 0 | 0 | 0 | 4 (1.9)[a] | 6 (12.5)[a] | 10 (2.2)[a] |
| Total | | | | | | | | | | | | | | | 114 | 82 | 11 | 210 | 48 | 465 |

**Table 3.7 Continued**

[a] All isolates were classified into lineages III or IV by MLGT.

[b] All isolates belonged to ECI.

[c] Seven isolates belonged to Group 1. Strain J3976 displayed hybridization pattern 10.

[d] Twenty one isolates belonged to ECIa.

[e] Seven isolates belonged to Group 2.

[f] Nine isolates belonged to Group 3.

[g] Hybridization pattern 23 represents Region 85 similar to FSL J2-003 (Fig. 1C).

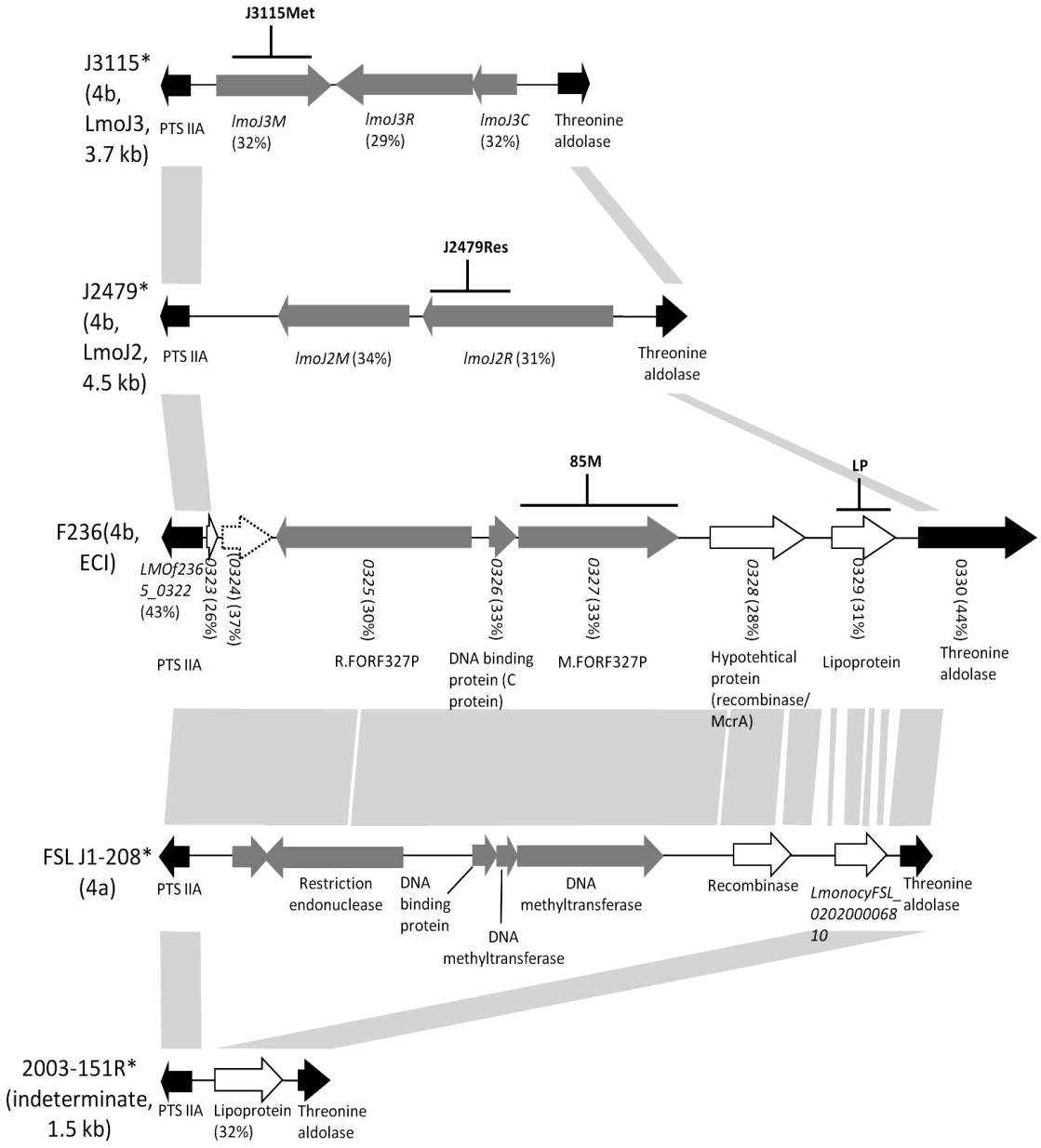
[h] Isolates displaying these hybridization patterns seem to harbor Region 85 similar to H7858 (Fig. 1C).

[i] All isolates belonged to ECII or ECV.

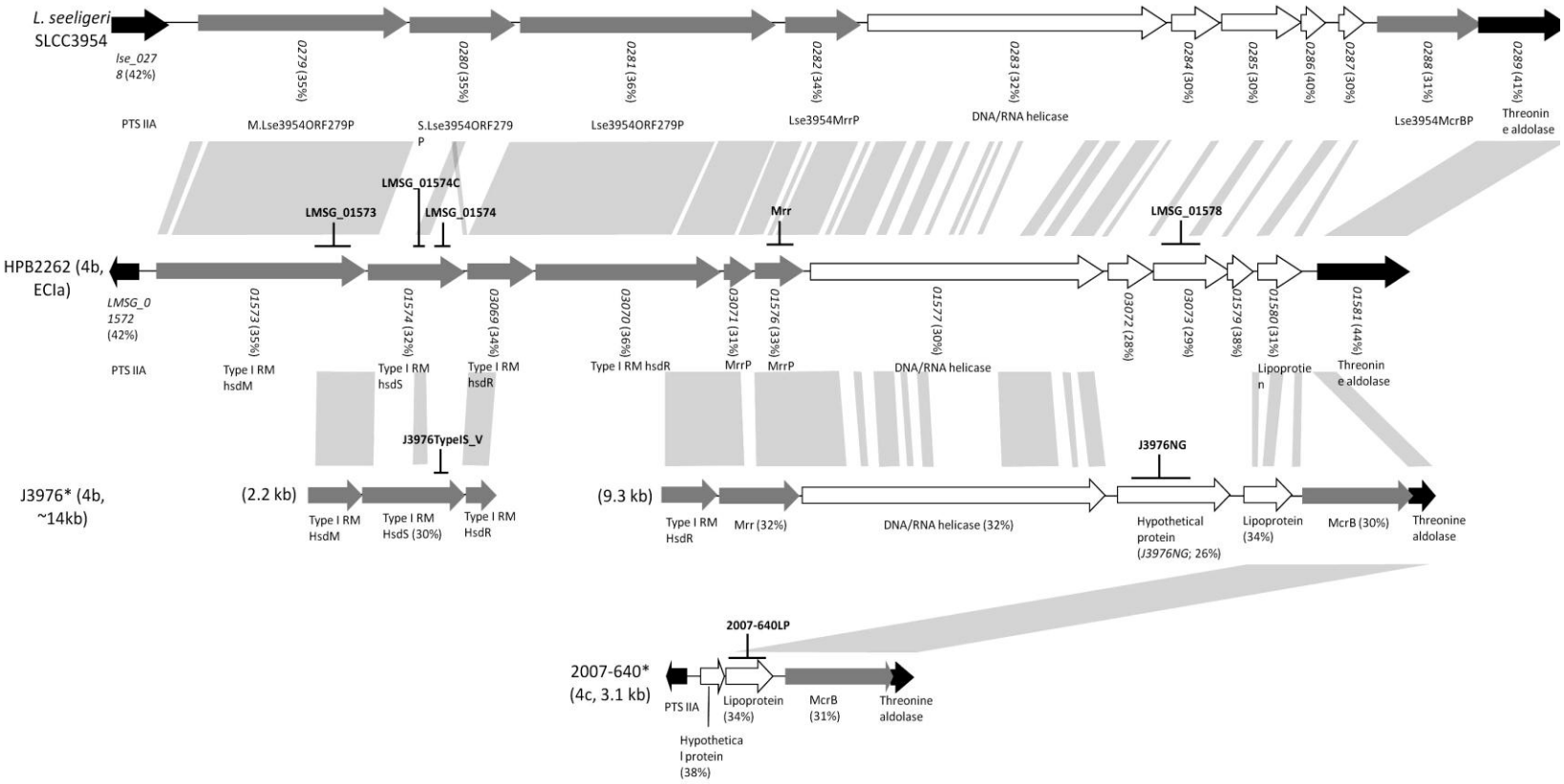
[j] Isolates displaying these hybridization patterns seem to harbor Region 85 similar to FSL J1-175 (Fig. 1C).

**Figure 3.1Genetic organization of Region 85 of different *Listeria* spp. and *L. monocytogenes* strains obtained from bioinformatics analyses and sequencing results.** The strains whose Region 85 was sequenced are marked with asterisk and, next to the stain name, the size of the PCR product with primers annealing to the conserved flanking genes was shown. The conserved flanking genes were colored in black; RM systems were marked in dark grey and annotation on the REBASE database were shown if available; and the other genes in Region 85 were colored in white. Pseudogenes were represented with stippled arrows and their gene names were written inside the parentheses. The grey shade signifies homologous DNA sequence obtained from the analysis with WebACT and the percentage next to the gene names represents the GC content of the gene. For FSL J1-208, the GC content was not calculated due to potentially incorrect identification of ORFs due to sequencing mistakes. The DNA probes were shown as a line over their target genes. A Region 85 harboring Type II RM systems. B Type I RM systems identified in Region 85. C Region 85 harboring Type IV RM systems without Type I RM systems. D Region 85 harboring lipoprotein genes without RM systems. E Region 85 similar to *L. welshimeri* SLCC5334. In each picture, other strains that did not satisfy the above designation were also included based on the similarity of Region 85.

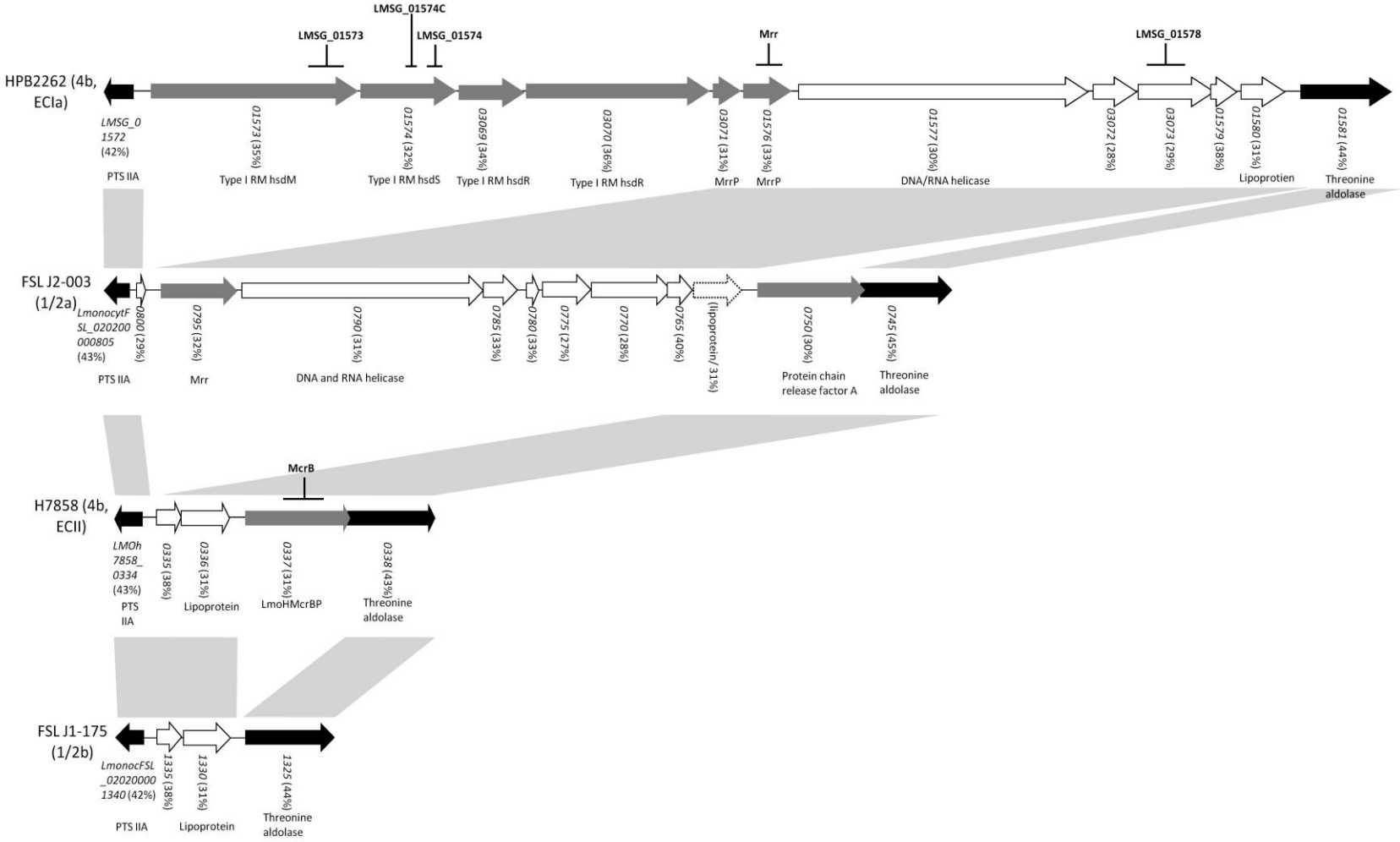
**A.**


J3115Met
J3115* (4b, LmoJ3, 3.7 kb)
PTS IIA
lmoJ3M (32%)
lmoJ3R (29%)
lmoJ3C (32%)
Threonine aldolase
J2479Res
J2479* (4b, LmoJ2, 4.5 kb)
PTS IIA
lmoJ2M (34%)
lmoJ2R (31%)
Threonine aldolase
85M
LP
F236(4b, ECI)
LMOf2365_0322 (43%)
0323 (26%)
(0324) (37%)
0325 (30%)
0326 (33%)
0327 (33%)
0328 (28%)
0329 (31%)
0330 (44%)
PTS IIA
R.FORF327P
DNA binding protein (C protein)
M.FORF327P
Hypotehtical protein (recombinase/ McrA)
Lipoprotein
Threonine aldolase
FSL J1-208* (4a)
PTS IIA
Restriction endonuclease
DNA binding protein
DNA methyltransferase
DNA methyltransferase
Recombinase
LmonocyFSL_0202000068 10
Threonine aldolase
2003-151R* (indeterminate, 1.5 kb)
PTS IIA
Lipoprotein (32%)
Threonine aldolase

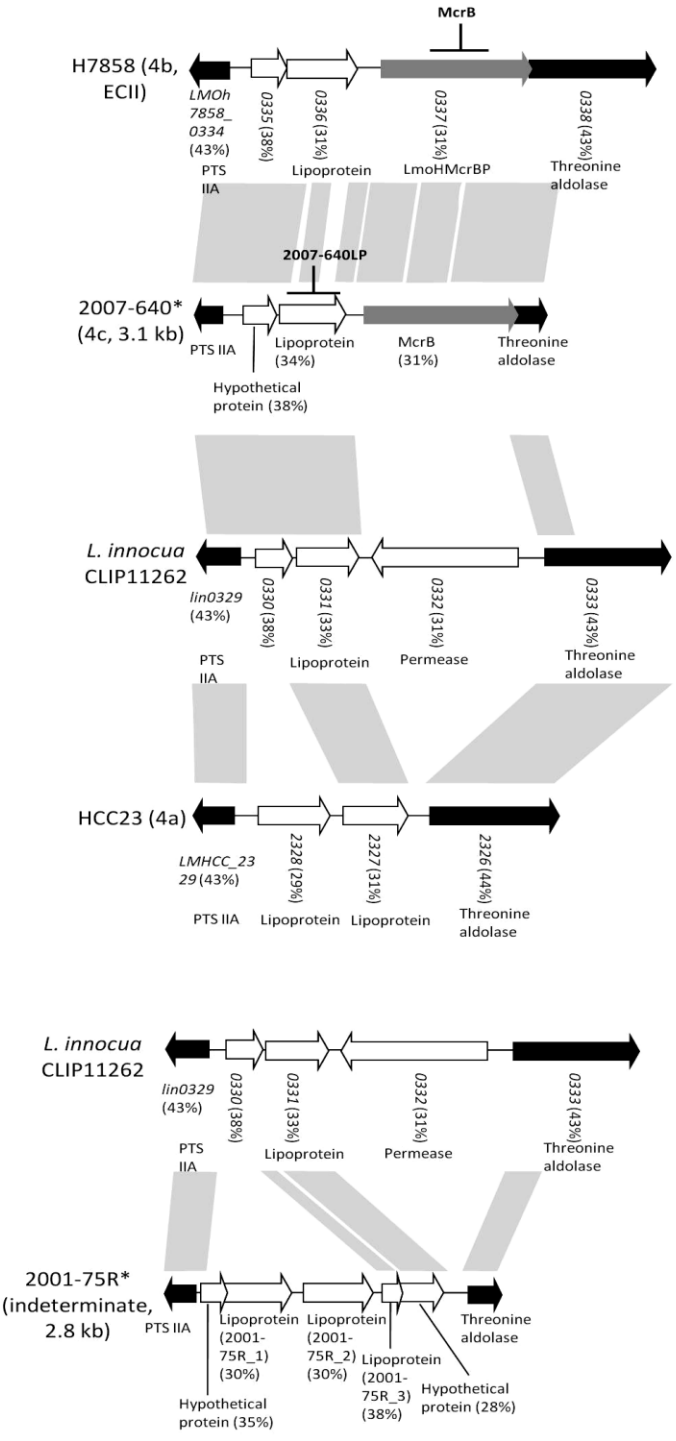
**B.**


L. seeligeri SLCC3954
lse_0278 (42%)
0279 (35%)
0280 (35%)
0281 (36%)
0282 (34%)
0283 (32%)
0284 (30%)
0285 (30%)
0286 (40%)
0287 (30%)
0288 (31%)
0289 (41%)
PTS IIA
M.Lse3954ORF279P
S.Lse3954ORF279P
Lse3954ORF279P
Lse3954MrrP
DNA/RNA helicase
Lse3954McrBP
Threonine aldolase
LMSG_01574C
LMSG_01573
LMSG_01574
Mrr
LMSG_01578
HPB2262 (4b, ECIa)
LMSG_01572 (42%)
01573 (35%)
01574 (32%)
030069 (34%)
030070 (36%)
030071 (31%)
01576 (33%)
01577 (30%)
030072 (28%)
030073 (29%)
01579 (38%)
01580 (31%)
01581 (44%)
PTS IIA
Type I RM hsdM
Type I RM hsdS
Type I RM hsdR
Type I RM hsdR
MrrP
MrrP
DNA/RNA helicase
Lipoprotein
Threonine aldolase
J3976TypeIS_V
J3976NG
J3976* (4b, ~14kb)
(2.2 kb)
(9.3 kb)
Type I RM HsdM
Type I RM HsdS (30%)
Type I RM HsdR
Type I RM HsdR
Mrr (32%)
DNA/RNA helicase (32%)
Hypothetical protein (J3976NG; 26%)
Lipoprotein (34%)
McrB (30%)
Threonine aldolase
2007-640LP
2007-640* (4c, 3.1 kb)
PTS IIA
Lipoprotein (34%)
McrB (31%)
Threonine aldolase
Hypothetical protein (38%)

**C.**


LMSG_01574C
LMSG_01573
LMSG_01574
Mrr
LMSG_01578
HPB2262 (4b, ECIa)
LMSG_01572 (42%)
01573 (35%)
01574 (32%)
03069 (34%)
03070 (36%)
03071 (31%)
01576 (33%)
01577 (30%)
03072 (28%)
03073 (29%)
01579 (38%)
01580 (31%)
01581 (44%)
PTS IIA
Type I RM hsdM
Type I RM hsdS
Type I RM hsdR
Type I RM hsdR
MrrP
MrrP
DNA/RNA helicase
Lipoprotien
Threonine aldolase
FSL J2-003 (1/2a)
LmonocytFSL_020200000805 (43%)
0800 (29%)
0795 (32%)
0790 (31%)
0785 (33%)
0780 (33%)
0775 (27%)
0770 (28%)
0765 (40%)
(lipoprotein/ 31%)
0750 (30%)
0745 (45%)
PTS IIA
Mrr
DNA and RNA helicase
Protein chain release factor A
Threonine aldolase
McrB
H7858 (4b, ECII)
LMOh7858_0334 (43%)
0335 (38%)
0336 (31%)
0337 (31%)
0338 (43%)
PTS IIA
Lipoprotein
LmoHMcrBP
Threonine aldolase
FSL J1-175 (1/2b)
LmonocFSL_02020000 1340 (42%)
1335 (38%)
1330 (31%)
1325 (44%)
PTS IIA
Lipoprotein
Threonine aldolase

**D.**


McrB
H7858 (4b, ECII)
LMOh 7858_ 0334 (43%)
0335 (38%)
0336 (31%)
0337 (31%)
0338 (43%)
PTS IIA
Lipoprotein
LmoHMcrBP
Threonine aldolase
2007-640LP
2007-640* (4c, 3.1 kb)
PTS IIA
Lipoprotein (34%)
McrB (31%)
Threonine aldolase
Hypothetical protein (38%)
L. innocua CLIP11262
lin0329 (43%)
0330 (38%)
0331 (33%)
0332 (31%)
0333 (43%)
PTS IIA
Lipoprotein
Permease
Threonine aldolase
HCC23 (4a)
LMHCC_23 29 (43%)
2328 (29%)
2327 (31%)
2326 (44%)
PTS IIA
Lipoprotein
Lipoprotein
Threonine aldolase
L. innocua CLIP11262
lin0329 (43%)
0330 (38%)
0331 (33%)
0332 (31%)
0333 (43%)
PTS IIA
Lipoprotein
Permease
Threonine aldolase
2001-75R* (indeterminate, 2.8 kb)
PTS IIA
Lipoprotein (2001-75R_1) (30%)
Lipoprotein (2001-75R_2) (30%)
Lipoprotein (2001-75R_3) (38%)
Threonine aldolase
Hypothetical protein (35%)
Hypothetical protein (28%)

**E.**

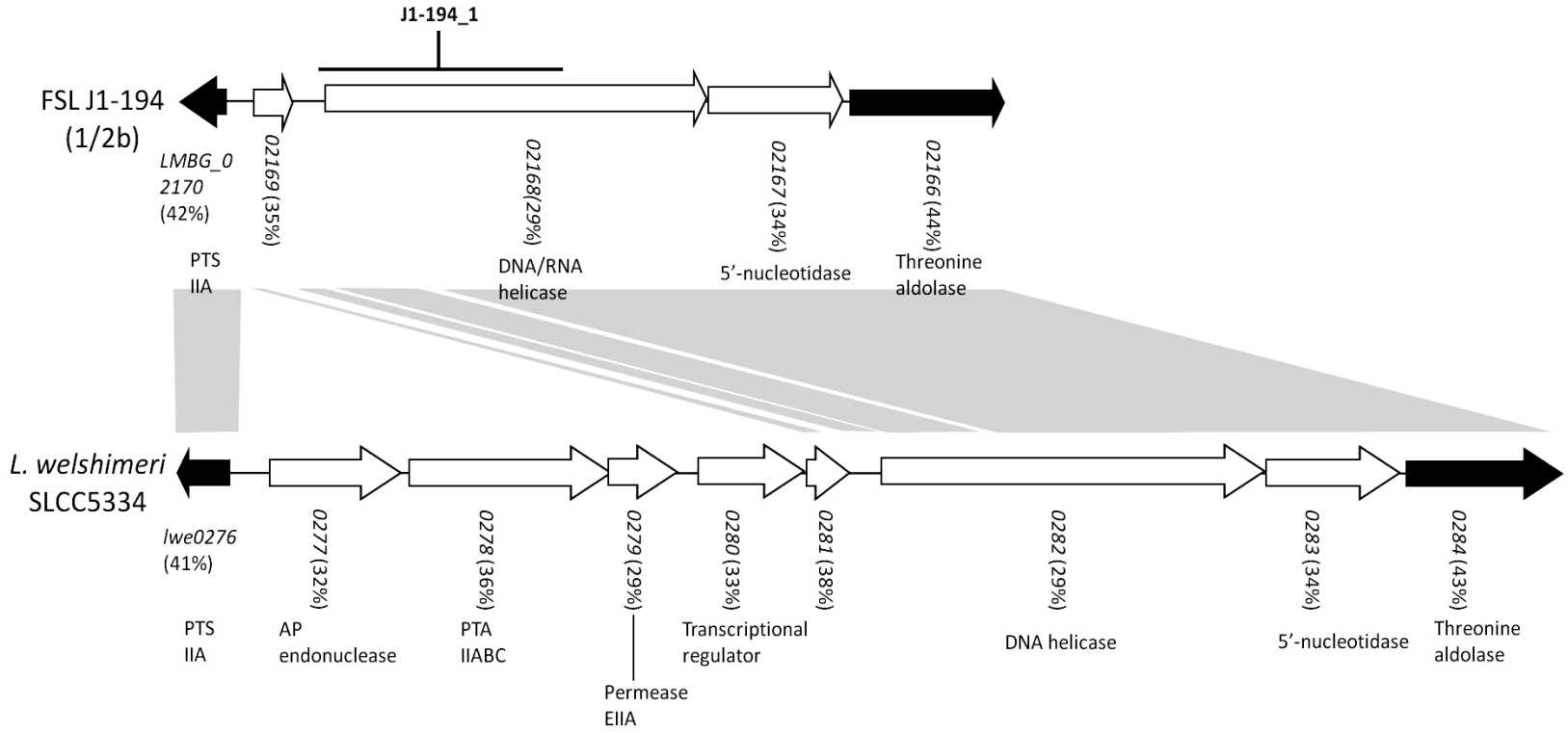

J1-194_1
FSL J1-194
(1/2b)
LMBG_02170
(42%)
02169 (35%)
02168 (29%)
02167 (34%)
02166 (44%)
PTS
IIA
DNA/RNA
helicase
5'-nucleotidase
Threonine
aldolase
L. welshimeri
SLCC5334
lwe0276
(41%)
0277 (32%)
0278 (36%)
0279 (29%)
0280 (33%)
0281 (38%)
0282 (29%)
0283 (34%)
0284 (43%)
PTS
IIA
AP
endonuclease
PTA
IIABC
Permease
EIIA
Transcriptional
regulator
DNA helicase
5'-nucleotidase
Threonine
aldolase

**Figure 3.2 Alignment of the *hsdS* genes identified in Region 85 of J3976 (J3976 *hsdS*), HPB2262 (*LMSG_01574*) and *L. welshimeri* SLCC3954 (*lse_0280*).** Conserved regions were highlighted in grey. The DNA sequence of DNA probe LMSG_01574C was marked in pink whereas the variable region DNA probes (LMSG_01574 and J3976TypeIS_V) were colored in green and yellow, respectively.

```
J3976 hsdS          ATGTCACATATTAACAAGAAAGTTCCTAAAAGAAGATTTCAGGAATTTAATAATACTGAT 60
lse_0280            ATGTCGGATAGTAAAAAGAAGGTGCCTAAACGAAGGTTTGAGGAATTTAGTAATGCTAAT 60
LMSG_01574          GTGTCACATATTAACAAGAGAGTTCCTAAAAGAAGATTTATGGAATTCAATAATGCTGTC 60
                     ****  *** *** ****  ** ****** **** ***  ****** * **** **

J3976 hsdS          GACTGGGAACAGCGTAAGTTAGGGGATGTTACGAATTA-TATAAAGGGATTTGCATTTAA 119
lse_0280            GATTGGGAACTGCGTAAGTTAGGCGGTCTGATGAATAT-TACATCTGTAAAACGAATTCA 119
LMSG_01574          GCTTGGGAACAGCGTAAGTTAGGTGAAATTGCGAATTCCTTCGAATATGGGCTTAATGCA 120
                    *  ******* ************ *   *   ****   *             * *  *

J3976 hsdS          ATCCAGAGACTATGCCGATAAAGGTATAAGAATAATAAGAGTATCAGA---TTTATCAAG 176
lse_0280            TCAATCAGACTGGACTGATAAAGGTGTCCGCTTTTTAAGGGCTCGCGACATTGTATCTGC 179
LMSG_01574          AGTTCTAAAACATA-TGATGGAGAAAATAAATATATTCGAATTACAGA-----TAT-TGA 173
                          * *       ***  **           *  *       **     ***

J3976 hsdS          TGATAAAATAAAAGAATTGAATGACAATGTTTATGTTACTCAACTTGTAGCTAATGAAAA 236
lse_0280            CTCAAAAGGTAAAAATCCTTCTGAATATCTTTATATCTCGAAAAAATTATATGATGAGCA 239
LMSG_01574          TGAAAGTTCTCATGTATTTAATCAAGATAATCTAACCTCCCCAGATATAAGTTTAGATAA 233
                        *      *         * *  **  *       *   *    **  *   **  *

J3976 hsdS          TAAAAAATTTATTATAGAAAAAGAAGATATTATTATTACAACTGTTGGTTCTAAGGCAGA 296
lse_0280            TTCAAAA---ATTTCAGGAAAAG------------TTGGGGTTGGGGATTT---GCTCGT 281
LMSG_01574          CCTAAATCATTACTTATTAGAAG---------------AGGGTGATATTCTCTTAGCGAG 278
                       ***         *  * ***                  **    *

J3976 hsdS          AATGAAAGAATCAGCTGTTGGAAGACCGATAATAGTAAGGCAT-AATAATAACTATCTAT 355
lse_0280            AACTGGGGTAGGGTCTATCGGTATCCCTATGCTTATTAAGCATGAAGAACCACTATATTT 341
LMSG_01574          AACAGGAGCAAGTACTG--GAAAATCTTATTGTTACAATAAAATAGATGGCAAAGTATTT 336
                    **     * *    **   *  *  *  **  *    *   *  *      *   * * *

J3976 hsdS          TAA--ATCAAAATTTAGTTAAAGTAAGAAGTAGTGAAAGTAGTTTTAGCTACTTTATATA 413
lse_0280            CAAGGATGGAAATATCATTTGGTTTCAGAATAAGAAAAATATTGATGGGGGATTTTTCTA 401
LMSG_01574          TTTG-CTGGGTTTCTAATAAGAGCAAAAATCAAACATGAATATAATGTCAGTTTTATCTT 395
                          *     * *  *          *  *   *      *  *      *** * *

J3976 hsdS          CTGT--CAGCTTATTAATCCTAGATACTCTGATTATATTGCCACAATTGAA--AGAGGTA 469
lse_0280            TTAT--TCGTTCAATAGTCATAGTATTCAAAAGTTCATTCGTGATAGTGCT--GGTATTG 457
LMSG_01574          TCAAAGCACTTTAACAG--AAAGGTATAATAATTTTATTCAGGTTACTTCTCAGAGATCA 453
                              * *  *     **        * *  ***      * *

J3976 hsdS          ATGCCAATCAAGCCAATATAGCTATAAGTGATTTATGGGAGTATAGTTTTCTCTCACCTA 529
lse_0280            GAACGGTTGGAACCTACACAATTGATAGCGGGGGAAAGACGCCAA---TTTATTTACCTA 514
LMSG_01574          GGCCAACCAGGAATTA-ATGCTCAAGAGTATGCTAGATTTGCTT----TATATATTCCTA 508
                       *           * *        **      *     *       *       ****

J3976 hsdS          AAAAAGATGAACAGACAAAAATTGGTACCTTCTTCAAACAACTTGATCACACTATCGCTC 589
lse_0280            ATAAAAAAGAACAACAAAGAATTGGCACCTTTTTCAAACAACTCGATAATACTATCGCTC 574
LMSG_01574          AATTAAAAGAACAACAAAAAATAGGTGACTTTTTCAAACAACTAGACAACACTATCGCTC 568
                    *   * * *****   ** *** **   *** *********** **  * **********

J3976 hsdS          TTCATCAACGAAAGCTAGAAAAAATAAAAGCATTAAAAACTGCCTATCTCTCAGAAATGT 649
lse_0280            TTCATCAACGTAAGTTAGAAAAAATAAAAGCATTAAAAACTGCCTATCTGTCAGAAATGT 634
LMSG_01574          TTCATCAACGTAAGCTAGAAAAAATAAAAGCATTAAAAACTGCCTATCTCTCAGAAATGT 628 LMSG_
                    ********** *** ********************************** ********** 01574C

J3976 hsdS          TTCCAGCAGAAGGAGAAACCAAACCAAAACGCCGATTTGCAGGATTTACTGATGCTTGGG 709
lse_0280            TTCCAGCAGAAGGAGAAACCAAACCAAAACGCCGATTTGCAGGATTTACTGATGATTGGG 694
LMSG_01574          TTCCAGCAGAAGGAGAAACCAAACCAAAACGCCGATTTGGAGGATTTACTGATGATTGGG 688
                    ************************************ ************** *****

J3976 hsdS          AAGAGATTAACCTTGGAGATATTATTGA--TTATGAGCAACCAA-------CAAAATATA 760
lse_0280            AACAGCGTAAACTAGATAATTCTATTAAAGTGATGGATGGAGATAGGGGTTCAAACTATC 754
LMSG_01574          AACAGCGTAAGTTTATAGAAATTATTAA-----TCGTTTAAG---------TAAAACCTC 734
                    ** **  ***  *     *   **** *     *                  ***   *
```

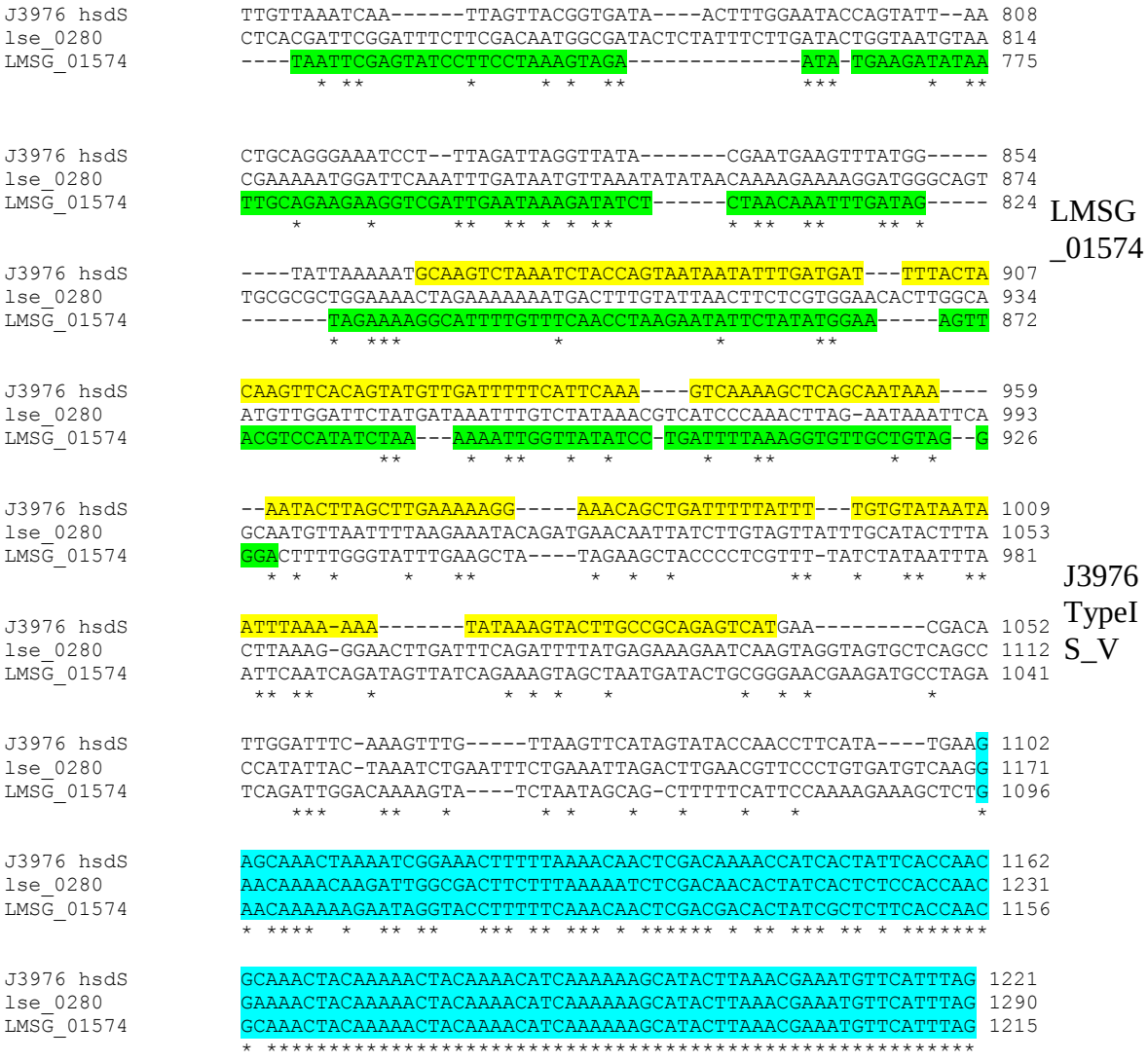
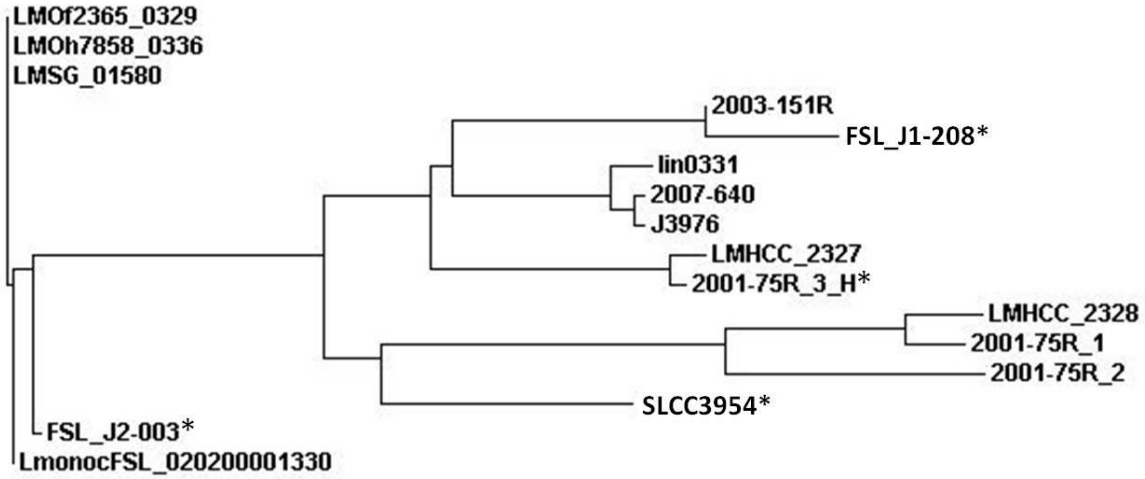
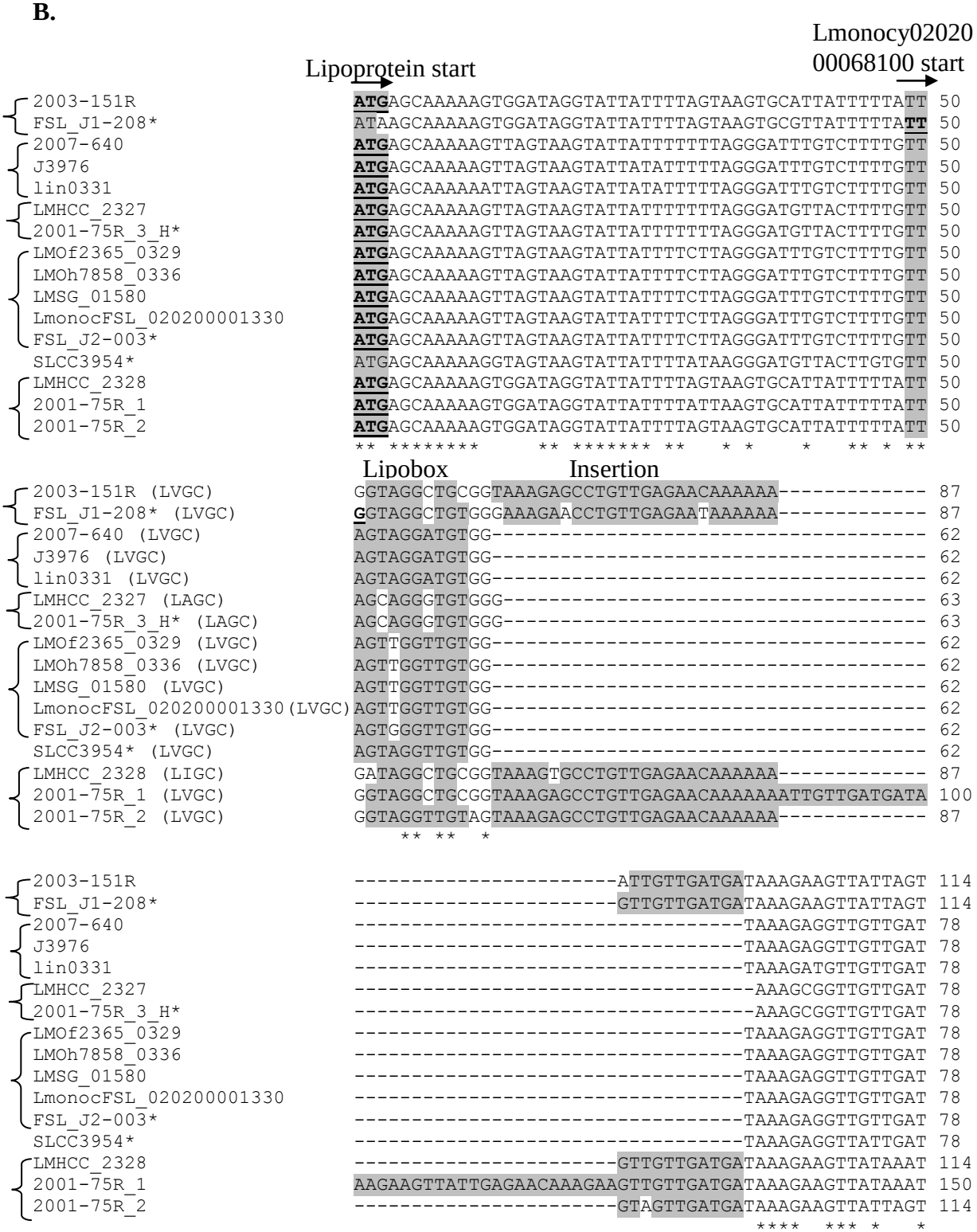

J3976 hsdS
lse_0280
LMSG_01574
LMSG_01574
J3976 TypeI S_V

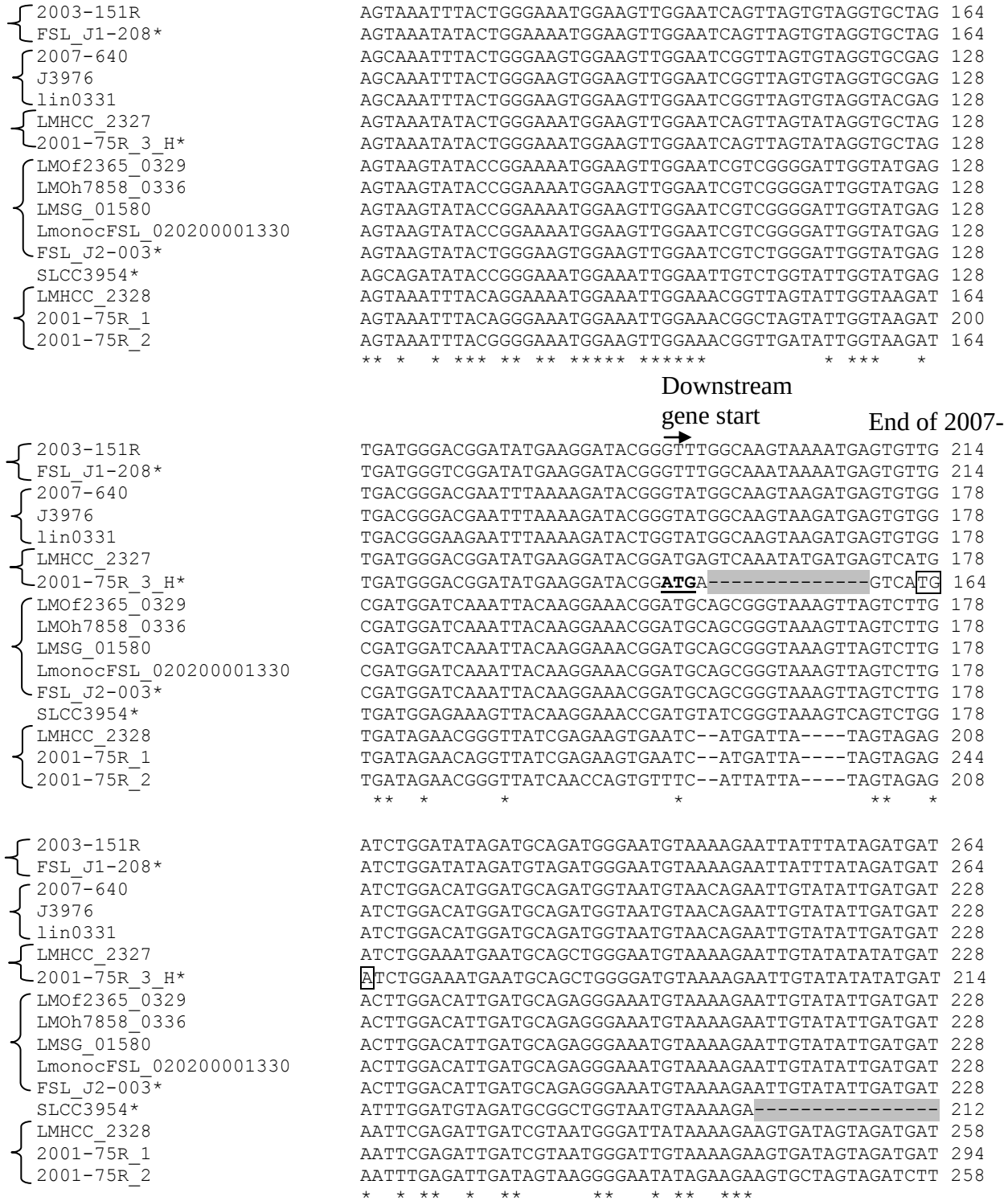
**Figure 3.3 Comparison of the lipoprotein genes and homologous DNA sequences (A and B) and of the proteins encoded in the lipoprotein genes (C).** In A and B, when the Region 85 harbored a homologous DNA sequence to the lipoprotein genes in which an ORF of the comparable length to the other lipoproteins was not identified, the homologous DNA sequence was used for the comparison and marked with the strain name and asterisk. In A, the phylogram of the DNA sequences was drawn with ClustalW. In B, the start codon of a gene was boldfaced and underlined whereas the stop codon was surrounded with the rectangular box. Similar DNA sequences were marked with the brace and important genetic features were marked in grey. In the alignment with the DNA sequence encoding the lipobox, the putative lipobox amino acid sequence was designated next to the DNA or gene name. In C, the lipobox was underlined and conserved sequences were obtained with the WebLogo. Based on the results from the WebLogo, six conserved regions were identified and marked in grey.

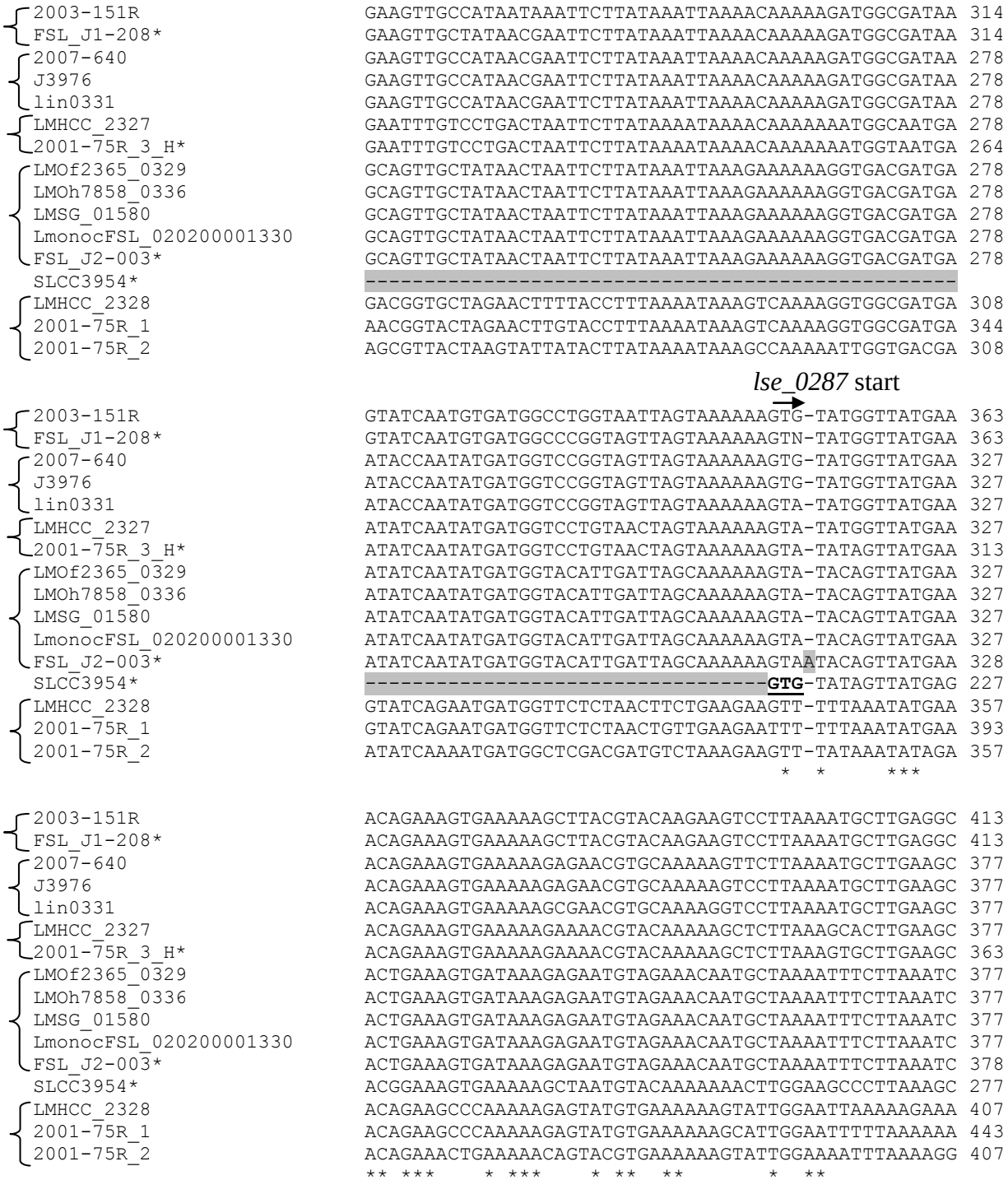
**A.**

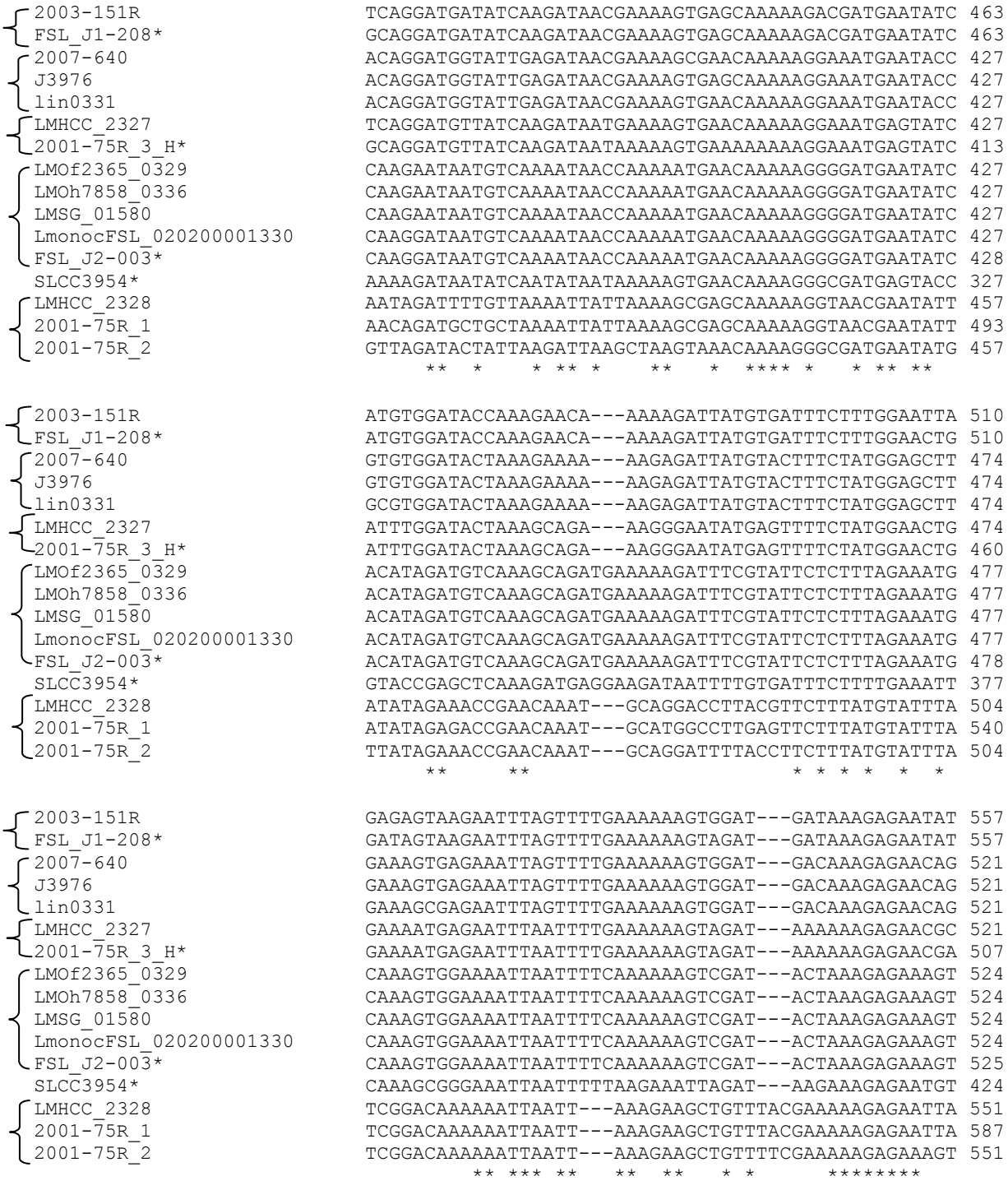

LMOf2365_0329
LMOh7858_0336
LMSG_01580
2003-151R
FSL_J1-208*
lin0331
2007-640
J3976
LMHCC_2327
2001-75R_3_H*
LMHCC_2328
2001-75R_1
2001-75R_2
SLCC3954*
FSL_J2-003*
LmonocFSL_020200001330

**B.**


Lmonocy02020
00068100 start
Lipoprotein start
Lipobox
Insertion
2003-151R
FSL_J1-208*
2007-640
J3976
lin0331
LMHCC_2327
2001-75R_3_H*
LMOf2365_0329
LMOh7858_0336
LMSG_01580
LmonocFSL_020200001330
FSL_J2-003*
SLCC3954*
LMHCC_2328
2001-75R_1
2001-75R_2
2003-151R (LVGC)
FSL_J1-208* (LVGC)
2007-640 (LVGC)
J3976 (LVGC)
lin0331 (LVGC)
LMHCC_2327 (LAGC)
2001-75R_3_H* (LAGC)
LMOf2365_0329 (LVGC)
LMOh7858_0336 (LVGC)
LMSG_01580 (LVGC)
LmonocFSL_020200001330(LVGC)
FSL_J2-003* (LVGC)
SLCC3954* (LVGC)
LMHCC_2328 (LIGC)
2001-75R_1 (LVGC)
2001-75R_2 (LVGC)

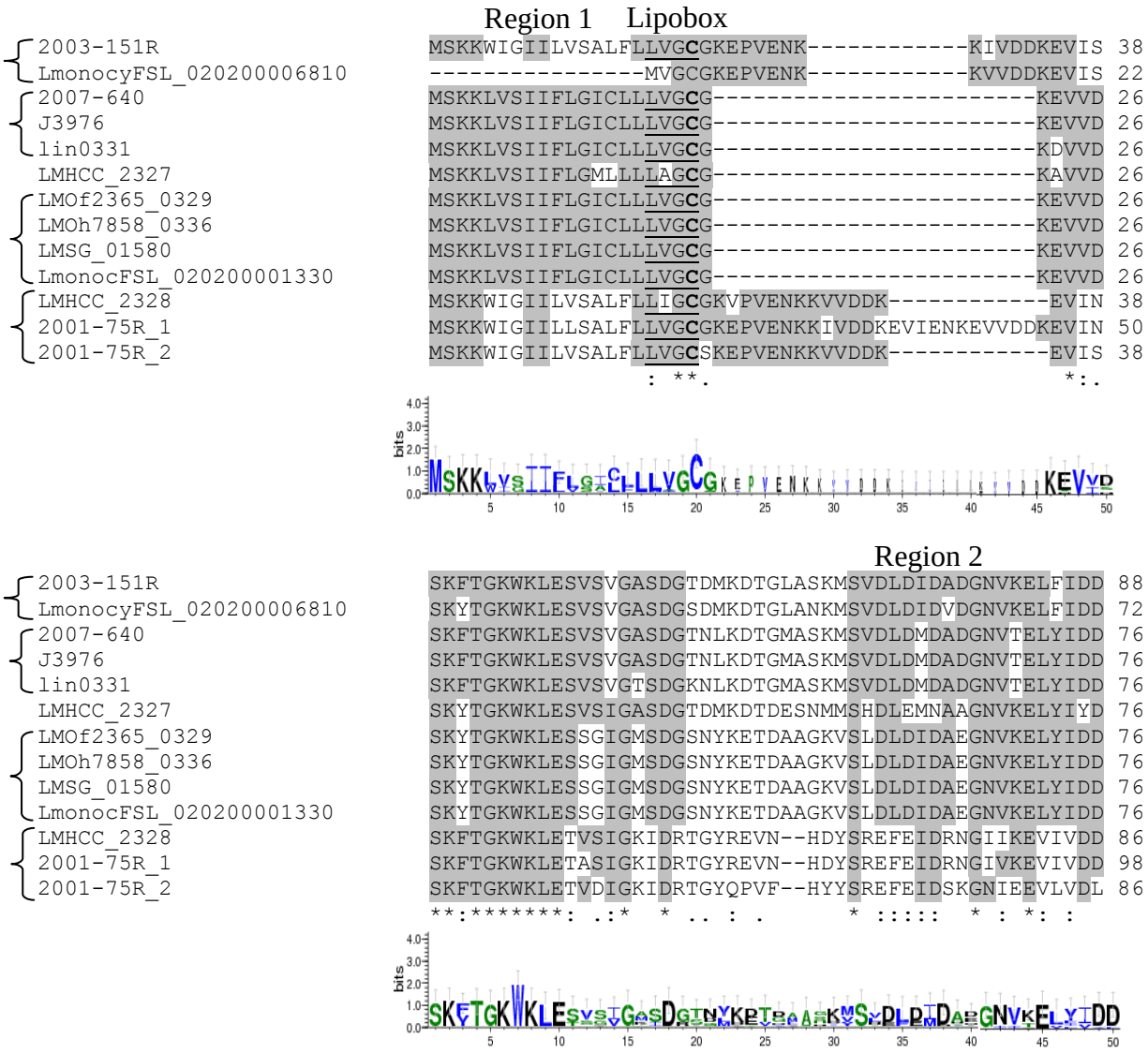

2003-151R AGTAAATTTACTGGGAAATGGAAGTTGGAATCAGTTAGTGTAGGTGCTAG 164
FSL_J1-208* AGTAAATATACTGGAAAATGGAAGTTGGAATCAGTTAGTGTAGGTGCTAG 164
2007-640 AGCAAATTTACTGGGAAGTGGAAGTTGGAATCGGTTAGTGTAGGTGCGAG 128
J3976 AGCAAATTTACTGGGAAGTGGAAGTTGGAATCGGTTAGTGTAGGTGCGAG 128
lin0331 AGCAAATTTACTGGGAAGTGGAAGTTGGAATCGGTTAGTGTAGGTACGAG 128
LMHCC_2327 AGTAAATATACTGGGAAATGGAAGTTGGAATCAGTTAGTATAGGTGCTAG 128
2001-75R_3_H* AGTAAATATACTGGGAAATGGAAGTTGGAATCAGTTAGTATAGGTGCTAG 128
LMOf2365_0329 AGTAAGTATACCGGAAAATGGAAGTTGGAATCGTCGGGGATTGGTATGAG 128
LMOh7858_0336 AGTAAGTATACCGGAAAATGGAAGTTGGAATCGTCGGGGATTGGTATGAG 128
LMSG_01580 AGTAAGTATACCGGAAAATGGAAGTTGGAATCGTCGGGGATTGGTATGAG 128
LmonocFSL_020200001330 AGTAAGTATACCGGAAAATGGAAGTTGGAATCGTCGGGGATTGGTATGAG 128
FSL_J2-003* AGTAAGTATACTGGGAAGTGGAAGTTGGAATCGTCTGGGATTGGTATGAG 128
SLCC3954* AGCAGATATACCGGGAAATGGAAATTGGAATTGTCTGGTATTGGTATGAG 128
LMHCC_2328 AGTAAATTTACAGGAAAATGGAAATTGGAAACGGTTAGTATTGGTAAGAT 164
2001-75R_1 AGTAAATTTACAGGGAAATGGAAATTGGAAACGGCTAGTATTGGTAAGAT 200
2001-75R_2 AGTAAATTTACGGGGAAATGGAAGTTGGAAACGGTTGATATTGGTAAGAT 164
Downstream gene start
End of 2007-
2003-151R TGATGGGACGGATATGAAGGATACGGGTTTGGCAAGTAAAATGAGTGTTG 214
FSL_J1-208* TGATGGGTCGGATATGAAGGATACGGGTTTGGCAAATAAAATGAGTGTTG 214
2007-640 TGACGGGACGAATTTAAAAGATACGGGTATGGCAAGTAAGATGAGTGTGG 178
J3976 TGACGGGACGAATTTAAAAGATACGGGTATGGCAAGTAAGATGAGTGTGG 178
lin0331 TGACGGGAAGAATTTAAAAGATACTGGTATGGCAAGTAAGATGAGTGTGG 178
LMHCC_2327 TGATGGGACGGATATGAAGGATACGGATGAGTCAAATATGATGAGTCATG 178
2001-75R_3_H* TGATGGGACGGATATGAAGGATACGGATGA---------------GTCATG 164
LMOf2365_0329 CGATGGATCAAATTACAAGGAAACGGATGCAGCGGGTAAAGTTAGTCTTG 178
LMOh7858_0336 CGATGGATCAAATTACAAGGAAACGGATGCAGCGGGTAAAGTTAGTCTTG 178
LMSG_01580 CGATGGATCAAATTACAAGGAAACGGATGCAGCGGGTAAAGTTAGTCTTG 178
LmonocFSL_020200001330 CGATGGATCAAATTACAAGGAAACGGATGCAGCGGGTAAAGTTAGTCTTG 178
FSL_J2-003* CGATGGATCAAATTACAAGGAAACGGATGCAGCGGGTAAAGTTAGTCTTG 178
SLCC3954* TGATGGAGAAAGTTACAAGGAAACCGATGTATCGGGTAAAGTCAGTCTGG 178
LMHCC_2328 TGATAGAACGGGTTATCGAGAAGTGAATC--ATGATTA----TAGTAGAG 208
2001-75R_1 TGATAGAACAGGTTATCGAGAAGTGAATC--ATGATTA----TAGTAGAG 244
2001-75R_2 TGATAGAACGGGTTATCAACCAGTGTTTC--ATTATTA----TAGTAGAG 208
2003-151R ATCTGGATATAGATGCAGATGGGAATGTAAAAGAATTATTTATAGATGAT 264
FSL_J1-208* ATCTGGATATAGATGTAGATGGGAATGTAAAAGAATTATTTATAGATGAT 264
2007-640 ATCTGGACATGGATGCAGATGGTAATGTAACAGAATTGTATATTGATGAT 228
J3976 ATCTGGACATGGATGCAGATGGTAATGTAACAGAATTGTATATTGATGAT 228
lin0331 ATCTGGACATGGATGCAGATGGTAATGTAACAGAATTGTATATTGATGAT 228
LMHCC_2327 ATCTGGAAATGAATGCAGCTGGGAATGTAAAAGAATTGTATATATATGAT 228
2001-75R_3_H* ATCTGGAAATGAATGCAGCTGGGGATGTAAAAGAATTGTATATATATGAT 214
LMOf2365_0329 ACTTGGACATTGATGCAGAGGGAAATGTAAAAGAATTGTATATTGATGAT 228
LMOh7858_0336 ACTTGGACATTGATGCAGAGGGAAATGTAAAAGAATTGTATATTGATGAT 228
LMSG_01580 ACTTGGACATTGATGCAGAGGGAAATGTAAAAGAATTGTATATTGATGAT 228
LmonocFSL_020200001330 ACTTGGACATTGATGCAGAGGGAAATGTAAAAGAATTGTATATTGATGAT 228
FSL_J2-003* ACTTGGACATTGATGCAGAGGGAAATGTAAAAGAATTGTATATTGATGAT 228
SLCC3954* ATTTGGATGTAGATGCGGCTGGTAATGTAAAAGA---------------- 212
LMHCC_2328 AATTCGAGATTGATCGTAATGGGATTATAAAAGAAGTGATAGTAGATGAT 258
2001-75R_1 AATTCGAGATTGATCGTAATGGGATTGTAAAAGAAGTGATAGTAGATGAT 294
2001-75R_2 AATTTGAGATTGATAGTAAGGGGAATATAGAAGAAGTGCTAGTAGATCTT 258


2003-151R GAAGTTGCCATAATAAATTCTTATAAATTAAAACAAAAAGATGGCGATAA 314
FSL_J1-208* GAAGTTGCTATAACGAATTCTTATAAATTAAAACAAAAAGATGGCGATAA 314
2007-640 GAAGTTGCCATAACGAATTCTTATAAATTAAAACAAAAAGATGGCGATAA 278
J3976 GAAGTTGCCATAACGAATTCTTATAAATTAAAACAAAAAGATGGCGATAA 278
lin0331 GAAGTTGCCATAACGAATTCTTATAAATTAAAACAAAAAGATGGCGATAA 278
LMHCC_2327 GAATTTGTCCTGACTAATTCTTATAAAATAAAACAAAAAAATGGCAATGA 278
2001-75R_3_H* GAATTTGTCCTGACTAATTCTTATAAAATAAAACAAAAAAATGGTAATGA 264
LMOf2365_0329 GCAGTTGCTATAACTAATTCTTATAAATTAAAGAAAAAAGGTGACGATGA 278
LMOh7858_0336 GCAGTTGCTATAACTAATTCTTATAAATTAAAGAAAAAAGGTGACGATGA 278
LMSG_01580 GCAGTTGCTATAACTAATTCTTATAAATTAAAGAAAAAAGGTGACGATGA 278
LmonocFSL_020200001330 GCAGTTGCTATAACTAATTCTTATAAATTAAAGAAAAAAGGTGACGATGA 278
FSL_J2-003* GCAGTTGCTATAACTAATTCTTATAAATTAAAGAAAAAAGGTGACGATGA 278
SLCC3954* ----------
LMHCC_2328 GACGGTGCTAGAACTTTTACCTTTAAAATAAAGTCAAAAGGTGGCGATGA 308
2001-75R_1 AACGGTACTAGAACTTGTACCTTTAAAATAAAGTCAAAAGGTGGCGATGA 344
2001-75R_2 AGCGTTACTAAGTATTATACTTATAAAATAAAGCCAAAAATTGGTGACGA 308
lse_0287 start
2003-151R GTATCAATGTGATGGCCTGGTAATTAGTAAAAAAGTG-TATGGTTATGAA 363
FSL_J1-208* GTATCAATGTGATGGCCCGGTAGTTAGTAAAAAAGTN-TATGGTTATGAA 363
2007-640 ATACCAATATGATGGTCCGGTAGTTAGTAAAAAAGTG-TATGGTTATGAA 327
J3976 ATACCAATATGATGGTCCGGTAGTTAGTAAAAAAGTG-TATGGTTATGAA 327
lin0331 ATACCAATATGATGGTCCGGTAGTTAGTAAAAAAGTA-TATGGTTATGAA 327
LMHCC_2327 ATATCAATATGATGGTCCTGTAACTAGTAAAAAAGTA-TATGGTTATGAA 327
2001-75R_3_H* ATATCAATATGATGGTCCTGTAACTAGTAAAAAAGTA-TATAGTTATGAA 313
LMOf2365_0329 ATATCAATATGATGGTACATTGATTAGCAAAAAAGTA-TACAGTTATGAA 327
LMOh7858_0336 ATATCAATATGATGGTACATTGATTAGCAAAAAAGTA-TACAGTTATGAA 327
LMSG_01580 ATATCAATATGATGGTACATTGATTAGCAAAAAAGTA-TACAGTTATGAA 327
LmonocFSL_020200001330 ATATCAATATGATGGTACATTGATTAGCAAAAAAGTA-TACAGTTATGAA 327
FSL_J2-003* ATATCAATATGATGGTACATTGATTAGCAAAAAAGTAATACAGTTATGAA 328
SLCC3954* ----------GTG-TATAGTTATGAG 227
LMHCC_2328 GTATCAGAATGATGGTTCTCTAACTTCTGAAGAAGTT-TTTAAATATGAA 357
2001-75R_1 GTATCAGAATGATGGTTCTCTAACTGTTGAAGAATTT-TTTAAATATGAA 393
2001-75R_2 ATATCAAAATGATGGCTCGACGATGTCTAAAGAAGTT-TATAAATATAGA 357
2003-151R ACAGAAAGTGAAAAAGCTTACGTACAAGAAGTCCTTAAAATGCTTGAGGC 413
FSL_J1-208* ACAGAAAGTGAAAAAGCTTACGTACAAGAAGTCCTTAAAATGCTTGAGGC 413
2007-640 ACAGAAAGTGAAAAAGAGAACGTGCAAAAAGTTCTTAAAATGCTTGAAGC 377
J3976 ACAGAAAGTGAAAAAGAGAACGTGCAAAAAGTCCTTAAAATGCTTGAAGC 377
lin0331 ACAGAAAGTGAAAAAGCGAACGTGCAAAAGGTCCTTAAAATGCTTGAAGC 377
LMHCC_2327 ACAGAAAGTGAAAAAGAAAACGTACAAAAAGCTCTTAAAGCACTTGAAGC 377
2001-75R_3_H* ACAGAAAGTGAAAAAGAAAACGTACAAAAAGCTCTTAAAGTGCTTGAAGC 363
LMOf2365_0329 ACTGAAAGTGATAAAGAGAATGTAGAAACAATGCTAAAATTTCTTAAATC 377
LMOh7858_0336 ACTGAAAGTGATAAAGAGAATGTAGAAACAATGCTAAAATTTCTTAAATC 377
LMSG_01580 ACTGAAAGTGATAAAGAGAATGTAGAAACAATGCTAAAATTTCTTAAATC 377
LmonocFSL_020200001330 ACTGAAAGTGATAAAGAGAATGTAGAAACAATGCTAAAATTTCTTAAATC 377
FSL_J2-003* ACTGAAAGTGATAAAGAGAATGTAGAAACAATGCTAAAATTTCTTAAATC 378
SLCC3954* ACGGAAAGTGAAAAAGCTAATGTACAAAAAAACTTGGAAGCCCTTAAAGC 277
LMHCC_2328 ACAGAAGCCCAAAAAGAGTATGTGAAAAAAGTATTGGAATTAAAAAGAAA 407
2001-75R_1 ACAGAAGCCCAAAAAGAGTATGTGAAAAAAGCATTGGAATTTTTAAAAAA 443
2001-75R_2 ACAGAAACTGAAAAACAGTACGTGAAAAAAGTATTGGAAAATTTAAAAGG 407


2003-151R TCAGGATGATATCAAGATAACGAAAAGTGAGCAAAAAGACGATGAATATC 463
FSL_J1-208* GCAGGATGATATCAAGATAACGAAAAGTGAGCAAAAAGACGATGAATATC 463
2007-640 ACAGGATGGTATTGAGATAACGAAAAGCGAACAAAAAGGAAATGAATACC 427
J3976 ACAGGATGGTATTGAGATAACGAAAAGTGAGCAAAAAGGAAATGAATACC 427
lin0331 ACAGGATGGTATTGAGATAACGAAAAGTGAACAAAAAGGAAATGAATACC 427
LMHCC_2327 TCAGGATGTTATCAAGATAATGAAAAGTGAACAAAAAGGAAATGAGTATC 427
2001-75R_3_H* GCAGGATGTTATCAAGATAATAAAAAGTGAAAAAAAAGGAAATGAGTATC 413
LMOf2365_0329 CAAGAATAATGTCAAAATAACCAAAAATGAACAAAAAGGGGATGAATATC 427
LMOh7858_0336 CAAGAATAATGTCAAAATAACCAAAAATGAACAAAAAGGGGATGAATATC 427
LMSG_01580 CAAGAATAATGTCAAAATAACCAAAAATGAACAAAAAGGGGATGAATATC 427
LmonocFSL_020200001330 CAAGGATAATGTCAAAATAACCAAAAATGAACAAAAAGGGGATGAATATC 427
FSL_J2-003* CAAGGATAATGTCAAAATAACCAAAAATGAACAAAAAGGGGATGAATATC 428
SLCC3954* AAAAGATAATATCAATATAATAAAAAGTGAACAAAAGGGCGATGAGTACC 327
LMHCC_2328 AATAGATTTTGTTAAAATTATTAAAAGCGAGCAAAAAGGTAACGAATATT 457
2001-75R_1 AACAGATGCTGCTAAAATTATTAAAAGCGAGCAAAAAGGTAACGAATATT 493
2001-75R_2 GTTAGATACTATTAAGATTAAGCTAAGTAAACAAAAGGGCGATGAATATG 457
** * * ** * ** * **** * * ** **
2003-151R ATGTGGATACCAAAGAACA---AAAAGATTATGTGATTTCTTTGGAATTA 510
FSL_J1-208* ATGTGGATACCAAAGAACA---AAAAGATTATGTGATTTCTTTGGAACTG 510
2007-640 GTGTGGATACTAAAGAAAA---AAGAGATTATGTACTTTCTATGGAGCTT 474
J3976 GTGTGGATACTAAAGAAAA---AAGAGATTATGTACTTTCTATGGAGCTT 474
lin0331 GCGTGGATACTAAAGAAAA---AAGAGATTATGTACTTTCTATGGAGCTT 474
LMHCC_2327 ATTTGGATACTAAAGCAGA---AAGGGAATATGAGTTTTCTATGGAACTG 474
2001-75R_3_H* ATTTGGATACTAAAGCAGA---AAGGGAATATGAGTTTTCTATGGAACTG 460
LMOf2365_0329 ACATAGATGTCAAAGCAGATGAAAAAGATTTCGTATTCTCTTTAGAAATG 477
LMOh7858_0336 ACATAGATGTCAAAGCAGATGAAAAAGATTTCGTATTCTCTTTAGAAATG 477
LMSG_01580 ACATAGATGTCAAAGCAGATGAAAAAGATTTCGTATTCTCTTTAGAAATG 477
LmonocFSL_020200001330 ACATAGATGTCAAAGCAGATGAAAAAGATTTCGTATTCTCTTTAGAAATG 477
FSL_J2-003* ACATAGATGTCAAAGCAGATGAAAAAGATTTCGTATTCTCTTTAGAAATG 478
SLCC3954* GTACCGAGCTCAAAGATGAGGAAGATAATTTTGTGATTTCTTTTGAAATT 377
LMHCC_2328 ATATAGAAACCGAACAAAT---GCAGGACCTTACGTTCTTTATGTATTTA 504
2001-75R_1 ATATAGAGACCGAACAAAT---GCATGGCCTTGAGTTCTTTATGTATTTA 540
2001-75R_2 TTATAGAAACCGAACAAAT---GCAGGATTTTACCTTCTTTATGTATTTA 504
** ** * * * * * *
2003-151R GAGAGTAAGAATTTAGTTTTGAAAAAAGTGGAT---GATAAAGAGAATAT 557
FSL_J1-208* GATAGTAAGAATTTAGTTTTGAAAAAAGTAGAT---GATAAAGAGAATAT 557
2007-640 GAAAGTGAGAAATTAGTTTTGAAAAAAGTGGAT---GACAAAGAGAACAG 521
J3976 GAAAGTGAGAAATTAGTTTTGAAAAAAGTGGAT---GACAAAGAGAACAG 521
lin0331 GAAAGCGAGAATTTAGTTTTGAAAAAAGTGGAT---GACAAAGAGAACAG 521
LMHCC_2327 GAAAATGAGAATTTAATTTTGAAAAAAGTAGAT---AAAAAAGAGAACGC 521
2001-75R_3_H* GAAAATGAGAATTTAATTTTGAAAAAAGTAGAT---AAAAAAGAGAACGA 507
LMOf2365_0329 CAAAGTGGAAAATTAATTTTCAAAAAAGTCGAT---ACTAAAGAGAAAGT 524
LMOh7858_0336 CAAAGTGGAAAATTAATTTTCAAAAAAGTCGAT---ACTAAAGAGAAAGT 524
LMSG_01580 CAAAGTGGAAAATTAATTTTCAAAAAAGTCGAT---ACTAAAGAGAAAGT 524
LmonocFSL_020200001330 CAAAGTGGAAAATTAATTTTCAAAAAAGTCGAT---ACTAAAGAGAAAGT 524
FSL_J2-003* CAAAGTGGAAAATTAATTTTCAAAAAAGTCGAT---ACTAAAGAGAAAGT 525
SLCC3954* CAAAGCGGGAAATTAATTTTTAAGAAATTAGAT---AAGAAAGAGAATGT 424
LMHCC_2328 TCGGACAAAAAATTAATT---AAAGAAGCTGTTTACGAAAAAGAGAATTA 551
2001-75R_1 TCGGACAAAAAATTAATT---AAAGAAGCTGTTTACGAAAAAGAGAATTA 587
2001-75R_2 TCGGACAAAAAATTAATT---AAAGAAGCTGTTTTCGAAAAAGAGAAAGT 551
** *** ** ** ** * * ********

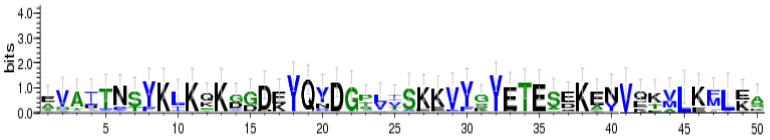
2003-151R TTTGATTAAGGAAACGTACGAACAAAATTAA 588
FSL_J1-208* TTTGATTCAGGAAACGTACGAACAAAATTAA 588
2007-640 TTTGATTCAGGAAACATATAAGAAAAATTAA 552
J3976 TTTGATTCAGGAAACATATAAGAAAAATTAA 552
lin0331 TTTGATTAAGGAAACATATAAGAAAAATTAA 552
LMHCC_2327 TATGGTTAAAGAAACGTATAAGAAAAATTAA 552
2001-75R_3_H* TATGGTTAAAGAAACGTATAAGAAAAATTAA 538
LMOf2365_0329 TTTGATTAAGGAAACCTACAAGAAAAATTAA 555
LMOh7858_0336 TTTGATTAAGGAAACCTACAAGAAAAATTAA 555
LMSG_01580 TTTGATTAAGGAAACCTACAAGAAAAATTAA 555
LmonocFSL_020200001330 TTTGATTAAGGAAACCTACAAGAAAAATTAA 555
FSL_J2-003* TTTGATTAAGGAAACCTACAAGAAAAATTAA 556
SLCC3954* TTTGATTAAGCAAACATATACGAAACAATAA 455
LMHCC_2328 TTCTAGTAAAGATATATATATTAAAAAATAA 582
2001-75R_1 TTCTGGTAAAGATATATATATTAAAAAATAA 618
2001-75R_2 TTCGAGTAAAGACACATTTATTAAACAATGA 582

**C.**


Region 1
Lipobox
2003-151R
LmonocyFSL_020200006810
2007-640
J3976
lin0331
LMHCC_2327
LMOf2365_0329
LMOh7858_0336
LMSG_01580
LmonocFSL_020200001330
LMHCC_2328
2001-75R_1
2001-75R_2
Region 2
bits

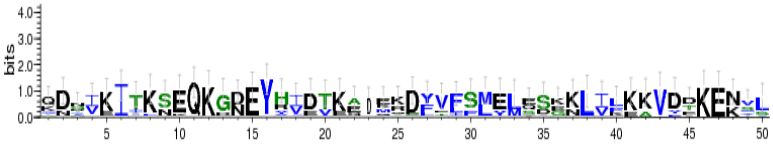
Region 3 Region 4

```
2003-151R                    EVAIINSYKLKQKDGDKYQCDGLVISKKVYGYETESEKAYVQEVLKMLEA 138
LmonocyFSL_020200006810      EVAITNSYKLKQKDGDKYQCDGPVVSKKVYGYETESEKAYVQEVLKMLEA 122
2007-640                     EVAITNSYKLKQKDGDKYQYDGPVVSKKVYGYETESEKENVQKVLKMLEA 126
J3976                        EVAITNSYKLKQKDGDKYQYDGPVVSKKVYGYETESEKENVQKVLKMLEA 126
lin0331                      EVAITNSYKLKQKDGDKYQYDGPVVSKKVYGYETESEKANVQKVLKMLEA 126
LMHCC_2327                   EFVLTNSYKIKQKNGNEYQYDGPVTSKKVYGYETESEKENVQKALKALEA 126
LMOf2365_0329                AVAITNSYKLKKKGDDEYQYDGTLISKKVYSYETESDKENVETMLKFLKS 126
LMOh7858_0336                AVAITNSYKLKKKGDDEYQYDGTLISKKVYSYETESDKENVETMLKFLKS 126
LMSG_01580                   AVAITNSYKLKKKGDDEYQYDGTLISKKVYSYETESDKENVETMLKFLKS 126
LmonocFSL_020200001330       AVAITNSYKLKKKGDDEYQYDGTLISKKVYSYETESDKENVETMLKFLKS 126
LMHCC_2328                   DGARTFTFKIKSKGGDEYQNDGSLTSEEVFKYETEAQKEYVKKVLELKRK 136
2001-75R_1                   NGTRTCTFKIKSKGGDEYQNDGSLTVEEFFKYETEAQKEYVKKALEFLKK 148
2001-75R_2                   SVTKYYTYKIKPKIGDEYQNDGSTMSKEVYKYRTETEKQYVKKVLENLKG 136
                               .   ::*:* * .::** **    ::.: *.**::*  *:  *:  .
```



Region 5 Region 6

```
2003-151R                    QDDIKITKSEQKDDEYHVDTKE-QKDYVISLELESKNLVLKKVDDKENIL 187
LmonocyFSL_020200006810      QDDIKITKSEQKDDEYHVDTKE-QKDYVISLELDSKNLVLKKVDDKENIL 171
2007-640                     QDGIEITKSEQKGNEYRVDTKE-KRDYVLSMELESEKLVLKKVDDKENSL 175
J3976                        QDGIEITKSEQKGNEYRVDTKE-KRDYVLSMELESEKLVLKKVDDKENSL 175
lin0331                      QDGIEITKSEQKGNEYRVDTKE-KRDYVLSMELESENLVLKKVDDKENSL 175
LMHCC_2327                   QDVIKIMKSEQKGNEYHLDTKA-EREYEFSMELENENLILKKVDKKENAM 175
LMOf2365_0329                KNNVKITKNEQKGDEYHIDVKADEKDFVFSLEMQSGKLIFKKVDTKEKVL 176
LMOh7858_0336                KNNVKITKNEQKGDEYHIDVKADEKDFVFSLEMQSGKLIFKKVDTKEKVL 176
LMSG_01580                   KNNVKITKNEQKGDEYHIDVKADEKDFVFSLEMQSGKLIFKKVDTKEKVL 176
LmonocFSL_020200001330       KDNVKITKNEQKGDEYHIDVKADEKDFVFSLEMQSGKLIFKKVDTKEKVL 176
LMHCC_2328                   IDFVKIIKSEQKGNEYYIETEQ-MQDLTFFMYLSDKKLIKEAVYEKENYS 185
2001-75R_1                   TDAAKIIKSEQKGNEYYIETEQ-MHGLEFFMYLSDKKLIKEAVYEKENYS 197
2001-75R_2                   LDTIKIKLSKQKGDEYVIETEQ-MQDFTFFMYLSDKKLIKEAVFEKEKVS 185
                              :  :*  .:**.:** ::.:   :   : : :.. :*: : *  **:
```



```
2003-151R                    IKETYEQN 195
LmonocyFSL_020200006810      IQETYEQN 179
2007-640                     IQETYKKN 183
J3976                        IQETYKKN 183
lin0331                      IKETYKKN 183
LMHCC_2327                   VKETYKKN 183
LMOf2365_0329                IKETYKKN 184
LMOh7858_0336                IKETYKKN 184
LMSG_01580                   IKETYKKN 184
LmonocFSL_020200001330       IKETYKKN 184
LMHCC_2328                   SKDIYIKK 193
2001-75R_1                   GKDIYIKK 205
2001-75R_2                   SKDTFIKQ 193
                              :: : ::
```

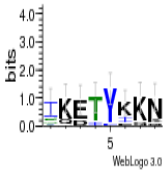
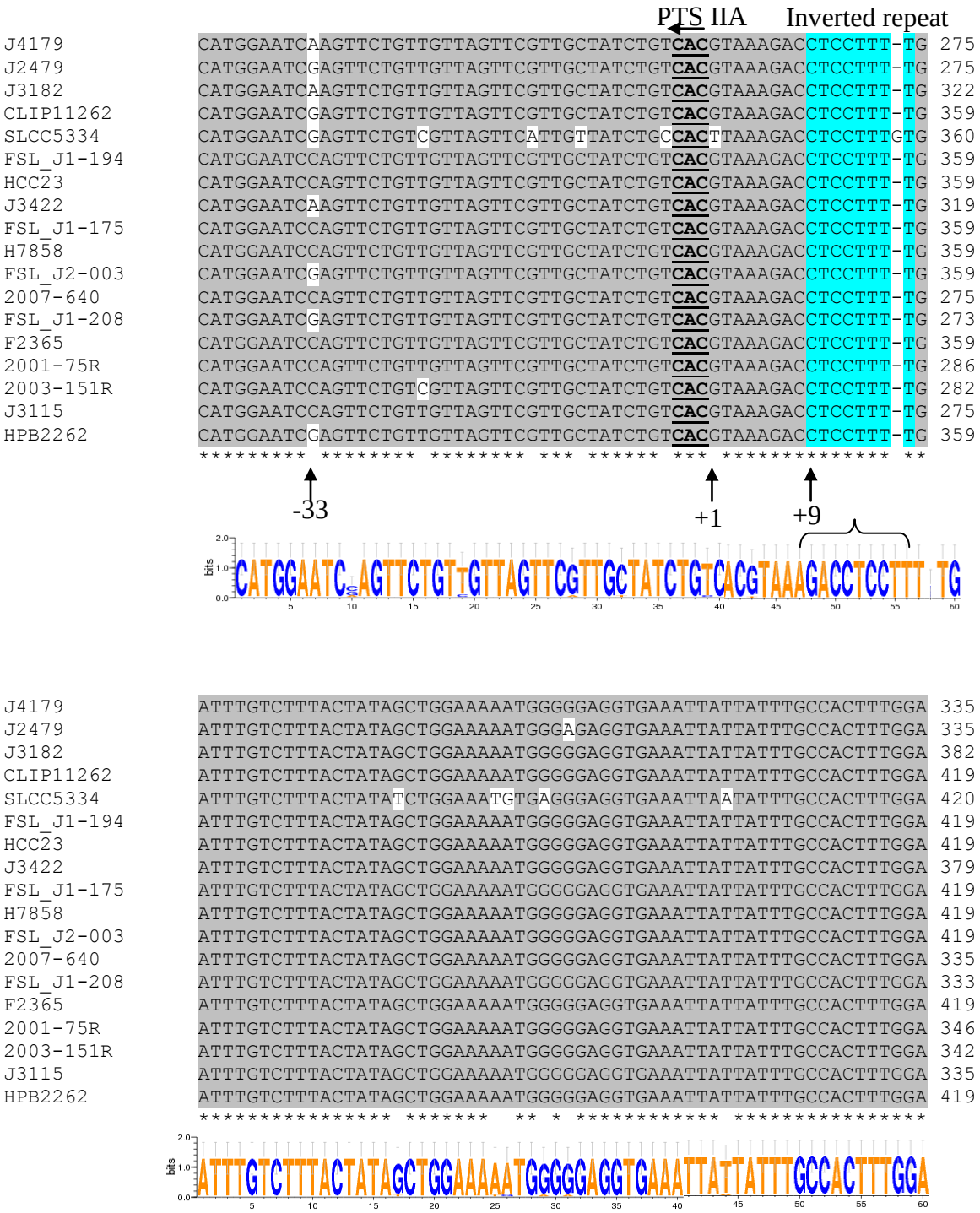
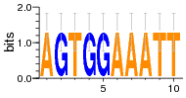


**Figure 3.4 Comparison of the upstream (A) and downstream (B) flanking regions of Region 85 in which different gene contents reside.** In both comparisons, the DNA sequences were obtained from the sequenced strains representing different forms of Region 85 and other strains whose hybridization patterns with Region 85 probes suggest a novel form of Region 85. The start codon of the genes was boldfaced and underlined and the conserved regions were colored in grey. The conserved sequence was determined with WebLogo and the inverted repeat was marked in blue. For the downstream flanking region comparison, the reverse complementary strand was used.

**A.**


PTS IIA
Inverted repeat
J4179 CATGGAATCAAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 275
J2479 CATGGAATCGAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 275
J3182 CATGGAATCAAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 322
CLIP11262 CATGGAATCGAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
SLCC5334 CATGGAATCGAGTTCTGTCGTTAGTTCATTGTTATCTGCCACTTAAAGACCTCCTTTGTG 360
FSL_J1-194 CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
HCC23 CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
J3422 CATGGAATCAAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 319
FSL_J1-175 CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
H7858 CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
FSL_J2-003 CATGGAATCGAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
2007-640 CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 275
FSL_J1-208 CATGGAATCGAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 273
F2365 CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
2001-75R CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 286
2003-151R CATGGAATCCAGTTCTGTCGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 282
J3115 CATGGAATCCAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 275
HPB2262 CATGGAATCGAGTTCTGTTGTTAGTTCGTTGCTATCTGTCACGTAAAGACCTCCTTT-TG 359
-33
+1
+9
bits
J4179 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 335
J2479 ATTTGTCTTTACTATAGCTGGAAAAATGGGAGAGGTGAAATTATTATTTGCCACTTTGGA 335
J3182 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 382
CLIP11262 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
SLCC5334 ATTTGTCTTTACTATATCTGGAAATGTGAGGGAGGTGAAATTAATTTGCCACTTTGGA 420
FSL_J1-194 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
HCC23 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
J3422 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 379
FSL_J1-175 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
H7858 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
FSL_J2-003 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
2007-640 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 335
FSL_J1-208 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 333
F2365 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
2001-75R ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 346
2003-151R ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 342
J3115 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 335
HPB2262 ATTTGTCTTTACTATAGCTGGAAAAATGGGGGAGGTGAAATTATTATTTGCCACTTTGGA 419
bits

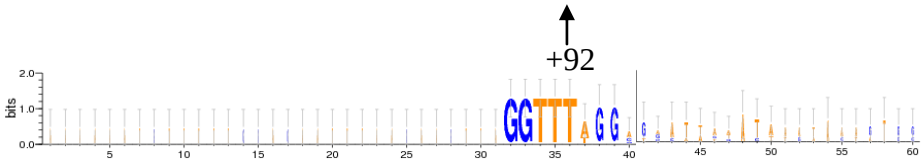
```
J4179           AGTGGAAATT---------------------------------------------------- 345
J2479           AGTGGAAATT---------------------------------------------------- 345
J3182           AGTGGAAATT---------------------------------------------------- 392
CLIP11262       AGTGGAAATT---------------------------------------------------- 429
SLCC5334        AGTGGAAATTAATTTTTGCTTGAGTATTGTATTTTTTTTTACTGGCTAATTCAAGAAATAT 480
FSL_J1-194      AGTGGAAATT---------------------------------------------------- 429
HCC23           AGTGGAAATT---------------------------------------------------- 429
J3422           AGTGGAAATT---------------------------------------------------- 389
FSL_J1-175      AGTGGAAATT---------------------------------------------------- 429
H7858           AGTGGAAATT---------------------------------------------------- 429
FSL_J2-003      AGTGGAAATT---------------------------------------------------- 429
2007-640        AGTGGAAATT---------------------------------------------------- 345
FSL_J1-208      AGTGGAAATT---------------------------------------------------- 343
F2365           AGTGGAAATT---------------------------------------------------- 429
2001-75R        AGTGGAAATT---------------------------------------------------- 356
2003-151R       AGTGGAAATT---------------------------------------------------- 352
J3115           AGTGGAAATT---------------------------------------------------- 345
HPB2262         AGTGGAAATT---------------------------------------------------- 429
                **********
```



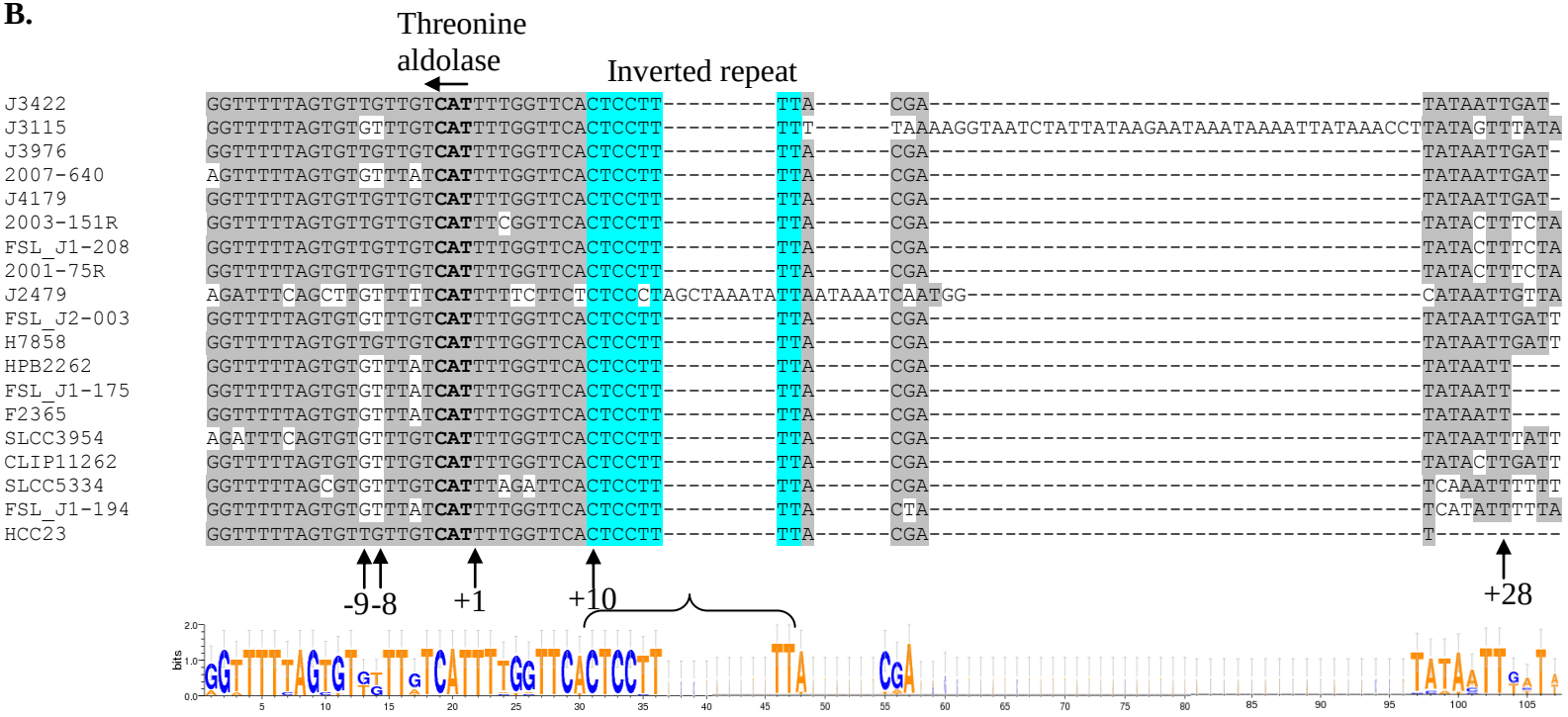

```
J4179           ------------------------------------------------------------
J2479           ------------------------------------------------------------
J3182           ------------------------------------------------------------
CLIP11262       ------------------------------------------------------------
SLCC5334        AGGGCTTTTTTGGCGAATTTAAATGAAAATATATGAAATAAATTGAAATAAAATGAAAAA 540
FSL_J1-194      ------------------------------------------------------------
HCC23           ------------------------------------------------------------
J3422           ------------------------------------------------------------
FSL_J1-175      ------------------------------------------------------------
H7858           ------------------------------------------------------------
FSL_J2-003      ------------------------------------------------------------
2007-640        ------------------------------------------------------------
FSL_J1-208      ------------------------------------------------------------
F2365           ------------------------------------------------------------
2001-75R        ------------------------------------------------------------
2003-151R       ------------------------------------------------------------
J3115           ------------------------------------------------------------
HPB2262         ------------------------------------------------------------
```

```
J4179          ------------------------------------------------------------
J2479          ------------------------------------------------------------
J3182          ------------------------------------------------------------
CLIP11262      ------------------------------------------------------------
SLCC5334       ATATGATTGTAAGCGATTGATTTCTGTGATATGCTTATTTCAACTTATAAAATATGGAGG 600
FSL_J1-194     ------------------------------------------------------------
HCC23          ------------------------------------------------------------
J3422          ------------------------------------------------------------
FSL_J1-175     ------------------------------------------------------------
H7858          ------------------------------------------------------------
FSL_J2-003     ------------------------------------------------------------
2007-640       ------------------------------------------------------------
FSL_J1-208     ------------------------------------------------------------
F2365          ------------------------------------------------------------
2001-75R       ------------------------------------------------------------
2003-151R      ------------------------------------------------------------
J3115          ------------------------------------------------------------
HPB2262        ------------------------------------------------------------


J4179          ------------------------------------------------------------
J2479          ------------------------------------------------------------
J3182          ------------------------------------------------------------
CLIP11262      ------lwe0277------------------------------------------------
SLCC5334       GAATTAAATGAAGTTAGCTACAAGAATCAATTCGTTTTTATCAGTTTATGATAATGATTT 660
FSL_J1-194     ------------------------------------------------------------
HCC23          ------------------------------------------------------------
J3422          ------------------------------------------------------------
FSL_J1-175     ------------------------------------------------------------
H7858          ------------------------------------------------------------
FSL_J2-003     ------------------------------------------------------------
2007-640       ------------------------------------------------------------
FSL_J1-208     ------------------------------------------------------------
F2365          ------------------------------------------------------------
2001-75R       ------------------------------------------------------------
2003-151R      ------------------------------------------------------------
J3115          ------------------------------------------------------------
HPB2262        ------------------------------------------------------------
```

```
J4179          --------------------------------GGTTT------------------------ 350
J2479          --------------------------------GGTTT------------------------ 350
J3182          --------------------------------GGTTTAGG--------------------- 400
CLIP11262      --------------------------------GGTTTTGGG-------------------- 438
SLCC5334       AAAAAATGTTTTTGGAGAATTTAAGAAGTTAGGTTTAGGCTACATTGACTTGAACTATCC 720
FSL_J1-194     --------------------------------GGTTTAGG--------------------- 437
HCC23          --------------------------------GGTTTA------------------------ 435
J3422          --------------------------------GGTTTAGG--------------------- 397
FSL_J1-175     --------------------------------GGTTT------------------------ 434
H7858          --------------------------------GGTTT------------------------ 434
FSL_J2-003     --------------------------------GGTTTTGGAGCATAATATAAATAAAGTGG 458
2007-640       --------------------------------GGTTT------------------------ 350
FSL_J1-208     --------------------------------GGTTTTGGAGGATTAAATA---------- 362
F2365          --------------------------------GGTTTAGGCGAATTAAATA---------- 448
2001-75R       --------------------------------GGTTT------------------------ 361
2003-151R      --------------------------------GGTTT------------------------ 357
J3115          --------------------------------GGTTTAGG---------TT---------- 355
HPB2262        --------------------------------GGTTTAGGAGGATAGTATAAATAAAGTGG 458
                                               *****
```



**B.**


Threonine aldolase
Inverted repeat
J3422 GGTTTTTAGTGTTGTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATTGAT-
J3115 GGTTTTTAGTGTGTTTGTCATTTGGTTCACTCCTT--------TTT-----TAAAAGGTAATCTATTATAAGAATAAATAAAATTATAAACCTTATAGTTTATA
J3976 GGTTTTTAGTGTTGTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATTGAT-
2007-640 AGTTTTTAGTGTGTTTATCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATTGAT-
J4179 GGTTTTTAGTGTTGTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATTGAT-
2003-151R GGTTTTTAGTGTTGTTGTCATTCGGTTCACTCCTT--------TTA-----CGA----------TATACTTTCTA
FSL_J1-208 GGTTTTTAGTGTTGTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATACTTTCTA
2001-75R GGTTTTTAGTGTTGTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATACTTTCTA
J2479 AGATTTCAGCTTGTTTTTCATTTTTCTTCTCTCCCTAGCTAAATATTAATAAATCAATGG----------CATAATTGTTA
FSL_J2-003 GGTTTTTAGTGTGTTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATTGATT
H7858 GGTTTTTAGTGTTGTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATTGATT
HPB2262 GGTTTTTAGTGTGTTTATCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATT----
FSL_J1-175 GGTTTTTAGTGTGTTTATCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATT----
F2365 GGTTTTTAGTGTGTTTATCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATT----
SLCC3954 AGATTTCAGTGTGTTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATAATTTATT
CLIP11262 GGTTTTTAGTGTGTTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------TATACTTGATT
SLCC5334 GGTTTTTAGCGTGTTTGTCATTTAGATTCACTCCTT--------TTA-----CGA----------TCAAATTTTTT
FSL_J1-194 GGTTTTTAGTGTGTTTATCATTTGGTTCACTCCTT--------TTA-----CTA----------TCATATTTTTA
HCC23 GGTTTTTAGTGTTGTTGTCATTTGGTTCACTCCTT--------TTA-----CGA----------T----------
-9
-8
+1
+10
+28
bits

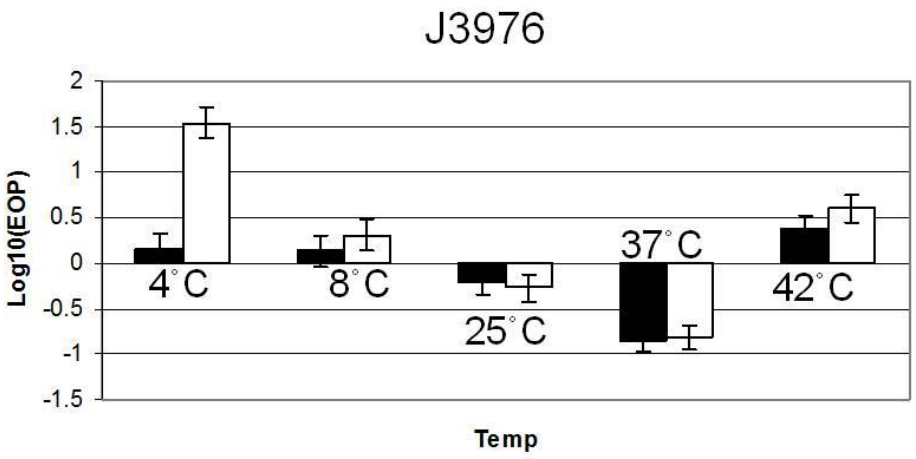



**Figure 3.5 Phage infection assay results.** The infections with 20422-1propagated in the reference strain and LmoJ3 isolate J3115 (20422-1MJ3) were marked with black and white bars, respectively.

## REFERENCES


1. **Aertsen, A., and C. W. Michiels.** 2005. Mrr instigates the SOS response after high pressure stress in *Escherichia coli*. Mol. Microbiol. **58:**1381-1391.
2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403-410.
3. **Buchrieser, C.** 2007. Biodiversity of the species *Listeria monocytogenes* and the genus *Listeria*. Microbes Infect. **9:**1147-1155.
4. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics. **21:**3422-3423.
5. **Cheng, Y., R. M. Siletzky, and S. Kathariou.** 2008. Genomic Divisions/Lineages, Epidemic Clones, and Population Structure, p. 337-358. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.
6. **Cheng, Y., J. W. Kim, S. Lee, R. M. Siletzky, and S. Kathariou.** 2010. DNA probes for unambiguous identification of *Listeria monocytogenes* epidemic clone II strains. Appl. Environ. Microbiol. **76:**3061-3068.
7. **Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, and J. D. Thompson.** 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. **31:**3497-3500.
8. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. Genome Res. **14:**1188-1190.

9. **den Bakker, H. C., X. Didelot, E. D. Fortes, K. K. Nightingale, and M. Wiedmann.** 2008. Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. BMC Evol. Biol. **8:**277.

10. **den Bakker, H. C., E. D. Fortes, and M. Wiedmann.** 2010. Multilocus sequence typing of outbreak-associated *Listeria monocytogenes* isolates to identify epidemic clones. Foodborne Pathog. Dis. **7:**257-265.

11. **Doumith, M., C. Buchrieser, P. Glaser, C. Jacquet, and P. Martin.** 2004. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. J. Clin. Microbiol. **42:**3819-3822.

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

13. **Eifert, J. D., P. A. Curtis, M. C. Bazaco, R. J. Meinersmann, M. E. Berrang, S. Kernodle, C. Stam, L. A. Jaykus, and S. Kathariou.** 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog. Dis. **2:**192-200.

14. **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:**2383-2390.

15. **Garcia-del Portillo, F., and P. Cossart.** 2007. An important step in *Listeria* lipoprotein research. J. Bacteriol. **189:**294-297.

16. **Graves, L. M., and B. Swaminathan.** 2001. PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. Int. J. Food Microbiol. **65:**55-62.

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

18. **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:**3972-3979.

19. **Hoskisson, P. A., and M. C. Smith.** 2007. Hypervariation and phase variation in the bacteriophage 'resistome'. Curr. Opin. Microbiol. **10:**396-400.

20. **Kathariou, S.** 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J. Food Prot. **65:**1811-1829.

21. **Kathariou, S., L. Graves, C. Buchrieser, P. Glaser, R. M. Siletzky, and B. Swaminathan.** 2006. Involvement of closely related strains of a new clonal group of *Listeria monocytogenes* in the 1998-99 and 2002 multistate outbreaks of foodborne listeriosis in the United States. Foodborne Pathog. Dis. **3:**292-302.

22. **Kibbe, W. A.** 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res. **35:**W43-6.

23. **Kim, J. W., R. M. Siletzky, and S. Kathariou.** 2008. Host ranges of *Listeria*-specific bacteriophages from the turkey processing plant environment in the United States. Appl. Environ. Microbiol. **74:**6623-6630.

24. **MacDonald, P. D., R. E. Whitwam, J. D. Boggs, J. N. MacCormack, K. L. Anderson, J. W. Reardon, J. R. Saah, L. M. Graves, S. B. Hunter, and J. Sobel.** 2005. Outbreak of listeriosis among Mexican immigrants as a result of consumption of illicitly produced Mexican-style cheese. Clin. Infect. Dis. **40:**677-682.

25. **Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng, and S. H. Bryant.** 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. **39:**D225-9.

26. **Meinersmann, R. J., R. W. Phillips, M. Wiedmann, and M. E. Berrang.** 2004. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes reveals clonal and recombination histories of three lineages. Appl. Environ. Microbiol. **70:**2193-2203.

27. **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:**337-351.

28. **Mullapudi, S., R. M. Siletzky, and S. Kathariou.** 2008. Heavy-metal and benzalkonium chloride resistance of *Listeria monocytogenes* isolates from the environment of turkey-processing plants. Appl. Environ. Microbiol. **74:**1464-1468.

29. **Murray, N. E.** 2002. 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. Microbiology. **148:**3-20.

30. **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:**2386-2395.

31. **Orsi, R. H., Q. Sun, and M. Wiedmann.** 2008. Genome-wide analyses reveal lineage specific contributions of positive selection and recombination to the evolution of *Listeria monocytogenes*. BMC Evol. Biol. **8:**233.

32. **Ragon, M., T. Wirth, F. Hollandt, R. Lavenir, M. Lecuit, A. Le Monnier, and S. Brisse.** 2008. A new perspective on *Listeria monocytogenes* evolution. PLoS Pathog. **4:**e1000146.

33. **Raleigh, E. A.** 1992. Organization and function of the *mcrBC* genes of *Escherichia coli* K-12. Mol. Microbiol. **6:**1079-1086.

34. **Rice, P., I. Longden, and A. Bleasby.** 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. **16:**276-277.

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

36. **Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin.** 2011. Foodborne illness acquired in the United States-major pathogens. Emerg. Infect. Dis. **17:**7-15.

37. **Sibley, M. H., and E. A. Raleigh.** 2004. Cassette-like variation of restriction enzyme genes in Escherichia coli C and relatives. Nucleic Acids Res. **32:**522-534.

38. **Sperry, K. E., S. Kathariou, J. S. Edwards, and L. A. Wolf.** 2008. Multiple-locus variable-number tandem-repeat analysis as a tool for subtyping *Listeria monocytogenes* strains. J. Clin. Microbiol. **46:**1435-1450.

39. **Srikhanta, Y. N., K. L. Fox, and M. P. Jennings.** 2010. The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. Nat. Rev. Microbiol. **8:**196-206.

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

41. **Thomas, C. M., and K. M. Nielsen.** 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. **3:**711-721.

42. **Tock, M. R., and D. T. Dryden.** 2005. The biology of restriction and anti-restriction. Curr. Opin. Microbiol. **8:**466-472.

43. **U.S. Food and Drug Administration.** 2006. Food additives permitted for direct addition to food for human consumption; bacteriophage preparation. Fed. Regist. **71:**47729-47732.

44. **Wagner, M., and J. MacLauchlin.** 2008. Biology, p. 3-25. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

45. **Ward, T. J., T. F. Ducey, T. Usgaard, K. A. Dunn, and J. P. Bielawski.** 2008. Multilocus genotyping assays for single nucleotide polymorphism-based subtyping of *Listeria monocytogenes* isolates. Appl. Environ. Microbiol. **74:**7629-7642.

46. **Ward, T. J., L. Gorski, M. K. Borucki, R. E. Mandrell, J. Hutchins, and K. Pupedis.** 2004. Intraspecific phylogeny and lineage group identification based on the *prfA* virulence gene cluster of *Listeria monocytogenes*. J. Bacteriol. **186:**4994-5002.

47. **Ward, T. J., T. Usgaard, and P. Evans.** 2010. A targeted multilocus genotyping assay for lineage, serogroup, and epidemic clone typing of *Listeria monocytogenes*. Appl. Environ. Microbiol. **76:**6680-6684.

48. **Warriner, K., and A. Namvar.** 2009. What is the hysteria with *Listeria*? Trends in Food Science & Technology. **20:**245-254.

49. **Yildirim, S., D. Elhanafi, W. Lin, A. D. Hitchins, R. M. Siletzky, and S. Kathariou.** 2010. Conservation of genomic localization and sequence content of Sau3AI-like restriction-modification gene cassettes among *Listeria monocytogenes* epidemic clone I and selected strains of serotype 1/2a. Appl. Environ. Microbiol. **76:**5577-5584.

50. **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:**4158-4164.

51. **Zheng, Y., D. Cohen-Karni, D. Xu, H. G. Chin, G. Wilson, S. Pradhan, and R. J. Roberts.** 2010. A unique family of Mrr-like modification-dependent restriction endonucleases. Nucleic Acids Res. **38:**5527-5534.

# Chapter 4 Two Novel Type II Restriction-Modification (RM) Systems Occupying a Genomically Equivalent Location on the Chromosome of *Listeria monocytogenes*


## ABSTRACT

*Listeria monocytogenes* is responsible for the potentially life-threatening foodborne disease listeriosis. One epidemic-associated clonal group of *L. monocytogenes*, epidemic clone I (ECI), harbors a Sau3AI-like restriction-modification (RM) system that also lies in the same genomic region in certain strains of other lineages. In this study, we identified and characterized two novel Type II RM systems, LmoJ2 and LmoJ3, in this locus. Sequence analysis suggested that LmoJ2 and LmoJ3 recognize GCWGC (W=A or T) and GCNGC, respectively. Both RM systems consisted of genes with GC content below the genome average and were harbored in same genetic region in strains of different serotypes and lineages, suggesting potential horizontal gene transfer. Genomic DNA from LmoJ2 and LmoJ3 strains grown at various temperatures (4-42ºC) in liquid or on agar was resistant to digestion with restriction enzymes recognizing GCWGC or GCNGC, indicating that the methyltransferases were expressed under these conditions, including at low temperatures. Phages propagated in a LmoJ2 strain exhibited a moderately increased (up to 7.9 fold) infectivity at 4 and 8ºC, while phages propagated in a LmoJ3 strain revealed the dramatically increased (up to $10^{5.9}$-fold) infectivity at all temperatures. Among the sequenced *Listeria* phages, lytic phages possessed significantly fewer recognition sites for these RM systems than lysogenic phages, suggesting that sequence content in lytic phages has evolved towards reduced susceptibility to such RM systems. The ability of LmoJ2 and LmoJ3 to protect

against phages may affect the efficiency of phages as biocontrol agent for *L. monocytogenes* strains harboring these RM systems.

## INTRODUCTION

*Listeria monocytogenes* is the Gram-positive bacterium that can cause severe foodborne disease (listeriosis) in humans and animals. Symptoms of listeriosis include septicemia, meningitis, encephalitis, stillbirths, and abortions (36). *L. monocytogenes* remains a leading contributor to severe foodborne disease and death in the United States and other industrialized nations. In the United States, listeriosis is estimated to result in ca. 1600 cases and 250 deaths annually, with fatality rate estimated at ca. 16% (33). Control of *L. monocytogenes* has been challenging as this microorganism is ubiquitously present in the environment and exhibits a number of unique environmental adaptations including ability to grow at refrigeration temperatures, to form biofilms, and to resist disinfectants (4, 10, 18, 36).

A small number of clusters of genetically related strains of *L. monocytogenes* has been responsible for multiple chronologically and geographically unrelated outbreaks (5, 12, 13). One of these clusters, designated epidemic clone I (ECI), has been most extensively studied due to involvement in numerous outbreaks and sporadic cases in North America and Europe (5, 12). The genome of the ECI strain implicated in the 1985 outbreak of listeriosis in California involving Mexican-style cheese has been sequenced (26).

One of the characteristic attributes of ECI strains is that their genomic DNA exhibits resistance to the restriction endonuclease Sau3AI due to cytosine methylation at GATC sites (9, 22). Genome sequence analysis revealed a putative Sau3AI-like restriction-modification (RM) system consisting of three genes, encoding a cytosine-5 methyltransferase, a DNA-binding protein and a restriction endonuclease (Fig. 1) (26, 41). The involvement of this gene

cassette in Sau3AI resistance was subsequently confirmed with deletion mutagenesis (40). In addition to ECI (which, along with most other serotype 4b strains, belongs to lineage I), this RM cassette was also harbored in the same genomic region by certain strains of other lineages, including strains of serotypes 1/2a (lineage II), 1/2b (lineage II), 4a and 4c (lineage III) ((40); Chapter 3).

There is evidence for several additional RM systems in the genomic region that harbors the Sau3AI-like RM system in ECI and other strains. Some of these RM systems were identified via *in silico* analysis in our laboratory and include multiple Type I RM systems and two Type IV RM systems (McrB and Mrr) which will be described in detail in a separate presentation. In this study, we will focus on two novel Type II RM systems (LmoJ2 and LmoJ3) that inhabit this location and that were identified in the process of investigating genomic content and diversity in this region.

In the presence of active RM systems, the restriction endonuclease cuts foreign, typically unmethylated DNA molecules while the methyltransferase attaches methyl groups to the genomic DNA so that it is protected from restriction (35). Hence, RM systems can defend host cells against phage infection and reduce horizontal gene transfer (34, 35). Thus, RM system-rich regions in the genome may play important roles in the ecology and evolution of bacteria, including *L. monocytogenes*. However, only limited information is currently available regarding the functionality of the RM systems residing in the aforementioned RM system-rich *L. monocytogenes* genomic region. Even for the Sau3AI-like RM system, we lack conclusive evidence that its restriction endonuclease is actually expressed (40). In this study, we describe the genes of LmoJ2 and LmoJ3 as well as the genetic relationships of the

strains possessing these RM systems, and provide evidence for site-specific methylation of genomic DNA and for relative protection against phage infection in *L. monocytogenes* strains harboring these RM systems.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** *L. monocytogenes* isolates analyzed in this study are listed in Table 1. Bacterial cells were grown either in brain heart infusion (BHI; Becton, Dickinson and Company, Sparks, MD) or on BHI plate containing1.2% agar (Becton, Dickinson and Company) until stationary phase: overnight at 25, 37, and 42˚C; 7 days at 8˚C; and approximately 27 days at 4˚C.

**Extraction of bacterial genomic DNA and Polymerase Chain Reaction (PCR).** Genomic DNA was extracted from broth or plate cultures with DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA) following the directions provided by the manufacturer. PCR was performed with exTaq (Takara, Otsu, Shiga, Japan) in a thermocycler (Biometra, Goettingen, Germany) using the genomic DNA as template. The PCR reaction was initiated at 95˚C for 5 min and was followed by 31 cycles (each 95˚C for 1 min, 52˚C for 1 min, and 72˚C for 1 min) with a final extension at 72˚C for 10 min. When necessary, annealing temperatures and extension times were adjusted according to the melting temperatures of primers and the expected amplicon size, respectively. Primers employed in this study (Table 2) were purchased from Eurofins MWG Operon (Huntsville, AL) and are listed in Table 2. PCR amplicons were visualized on Tris/Borate/EDTA gels containing 0.8% agarose (Apex BioResearch Products, Research Triangle Park, NC).

**DNA-DNA dot blots.** For DNA-DNA hybridizations, we employed the nonradioactive DIG system (Roche, Florence, SC). Membranes were prepared by spotting bacterial genomic DNA on the nylon membrane (GE Water &Process Technologies, Trevose, PA) with a replicator (V&P Scientific, Inc., San Diego, CA); denaturing and neutralizing with denaturation solution (0.5M NaOH and 1.5M NaCl) and neutralization solution (1M Tris and 1.5M NaCl), and immobilizing the DNA onto the membrane using a UV crosslinker (UV Stratalinker; Agilent Technologies, Santa Clara, CA). To prepare probes, PCR amplicons were purified from the gel with the QIAquick gel extraction kit (QIAGEN), denatured for 10 min in boiling water, and DIG-labeled (Roche). The LmoJ2 and LmoJ3 probes (J2479Res and J3115Met, respectively) were prepared by PCR using primers J2479Res_F/ J2479_3R and J3115_2F/ J3115Met_R, respectively (Table 2).

Hybridization was implemented primarily as suggested by the manufacturer with some modifications. Specifically, a membrane was prehybridized in the prehybridization solution (5XSSC, 50% formamide, 0.1% N-lauroyl sarcosine, 0.02% SDS and 2% blocking agent) for three hours, and hybridized overnight at 42˚C with the DIG-labeled DNA probe. After hybridization, the membrane was subjected to three washes at low, medium, and high stringencies (0.1% SDS containing 2XSSC, 0.5XSSC, or 0.1XSSC, respectively), each for 20 min at 42˚C. Following a brief wash with maleic acid washing buffer (0.1M maleic acid, 0.15M NaCl and 0.3% Tween-20), the membrane was blocked with 1% blocking reagent (Roche) solution in maleic acid buffer (0.1M maleic acid and 0.15M NaCl). Then, antibodies in the antibody solution (Roche) were allowed to bind to the membrane and excess antibodies were removed during two washing steps with maleic acid washing buffer. The membrane

was incubated at 37˚C for 15 min in the presence of 1% CSPD (Roche) diluted in the detection buffer (0.1M Tris and 0.1M NaCl) and hybridization results were visualized on X-ray films after various exposure times. Membranes were stripped and re-probed up to three or four times depending on the concentration of the DNA spotted on the membrane.

**DNA sequencing and analysis.** PCR products were obtained from *L. monocytogenes* strains J2479 and J3115 using primers annealing to the flanking genes (H7858_0334R and H7858_0338F) (Table 2). The PCR amplicons were gel-purified with the QIAquick gel extraction kit (QIAGEN) and sequenced by GeneWiz (South Plainfield, NJ). The sequencing results were manually assembled and annotated with open reading frame (ORF) finder (www.ncbi.nlm.nih.gov/projects/gorf/), BLASTp (1) and conserved domain searches (21) provided by the National Center for Biotechnology Information (NCBI). Recognition sites for the RM systems were inferred based on the specificity sequences of the homologs identified through the BLAST program (http://tools.neb.com/~vincze/BLAST/index.php) on the REBASE database (http://rebase.neb.com/rebase/rebase.html) (31). Promoters were predicted with Neural Network Promoter Prediction provided by the Berkeley Drosophila Genome Project (http://www.fruitfly.org/seq_tools/promoter.html) (37). Putative Rho-independent terminators were identified with the model of Lesnik et al. (19) and RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) to detect stem-loop structures. Ribosome binding sites (RBS) were identified on the basis of location and similarity to the consensus sequence AGGAGGTG (15). Conserved domains of cytosine-5 methyltransferases were identified based on domains of M.HhaI described by Kumar et al. (17).

To examine the frequency of restriction sites for LmoJ2 and LmoJ3 (GCWGC (W=A or T) and GCNGC, respectively) in *L. monocytogenes* chromosomal fragments and *Listeria* phages, sequences listed in Table 3 were retrieved from the NCBI database. The number of restriction sites for LmoJ2 and LmoJ3 was determined by enumerating the sites for ApeKI (recognition sequence: G↓CWGC) and Fnu4HI (recognition sequence: GC↓NGC), respectively, with NEBcutter (http://tools.neb.com/NEBcutter2/). The frequency of the restriction sites was calculated by dividing the number of the restriction sites by the length of the analyzed DNA in kb.

**Enzyme digestion assays.** Genomic DNA was digested with restriction endonucleases ApeKI and Fnu4HI under conditions specified by the manufacturer (New England BioLabs, Ipswich, MA). Digested genomic DNAs were compared with uncut genomic DNA as well as with genomic DNA digested with HindIII (New England BioLabs), for quality control, on 0.8% agarose gels.

**Phage infection assays.** For phage assays, we utilized *Listeria* phage 20422-1 isolated from the environment of a food processing facility (14). This phage was propagated in *L. monocytogenes* DP-L862 as described (14). In addition, methylated phage derivatives were obtained by propagating 20422-1 at least twice in strains J2479 and J3115 that harbor LmoJ2 and LmoJ3, respectively.

For bacteriophage amplification, bacteria were grown overnight at 37˚C in 5 ml of BHI broth, and 100 µl was blended with 100 µl of the phage solution diluted with Luria Bertani medium (LB; Becton, Dickinson and Company). Phages were allowed to adhere to bacterial cells for five minutes. The phage-host mixture was added into 4 ml of molten soft

agar (0.75 % agar) in LB supplemented with 10 mM $CaCl_2$, and evenly distributed onto regular LB plates (1.2% agar) containing 10 mM $CaCl_2$. The plates were incubated at 37˚C for 24-48 hr. To collect phage particles, plates were flooded with 5 ml of SM buffer (0.1M NaCl, 0.2% $MgSO_4 \cdot 7H_2O$, 0.05M Tris-Cl, 0.01% gelatin, pH 7.5) and incubated 24-48 hr at 4˚C. The liquid was filtered through a 0.2 nm filter (Fischer Scientific, Pittsburgh, PA) and the filtrate was stored at 4˚C.

Phage susceptibility assays followed the same protocol and plaques were enumerated. Efficiency of plaquing (EOP) was defined as the ratio of plaque forming units (pfu) per ml obtained from a specific strain to pfu/ml obtained with the reference strain DP-L862. Phage experiments were conducted at least three times and statistically analyzed with SAS (SAS Institute Inc., Cary, NC) using the mixed ANOVA model.

**Pulsed-field gel electrophoresis (PFGE), multilocus genotyping (MLGT) and targeted MLGT (TMLGT).** PFGE was conducted with AscI (New England Biolabs) and ApaI (Roche) as described by Graves et al. (11). PFGE dendrograms based on dice coefficient were obtained with BioNumerics (Applied Maths, Austin, TX). Optimization level and position tolerance were set at 1.5% for ApaI and at 2.0% for AscI. MLGT and TMLGT were performed as described (8, 38, 39).

**Nucleotide sequence accession number.** The nucleotide sequences of LmoJ2 from *L. monocytogenes* J2749 and LmoJ3 from *L. monocytogenes* J3115 have been submitted to GenBank (accession nos. JN235993 and JN235992, respectively).

## RESULTS

**Sequence analysis of LmoJ2.** The novel RM systems were localized between the PTS IIA and aldolase genes on the chromosome of *L. monocytogenes* (Fig. 1) and were designated LmoJ2 and LmoJ3 after the strains in which they were first identified (*L. monocytogenes* J2749 and J3115, respectively). Sequence analysis of the 4.5 kb PCR amplicon from *L. monocytogenes* strain J2479 revealed LmoJ2 consisting of two ORFs, *lmoJ2M* (1144 bp) and *lmoJ2R* (1659 bp). The GC contents of *lmoJ2M* and *lmoJ2R* were 34% and 31%, respectively, significantly lower than the genome average of 38% (26).

The deduced polypeptide encoded by *lmoJ2M* exhibited high similarity with DNA methyltransferases from different bacteria, including *Enterococcus faecalis* T11 (ZP_05596805.1; 75% identity), *Lactobacillus delbrueckii* subsp. bulgaricus ATCC BAA-365 (YP_813118.1; 72% identity), *Bacillus cellulosilyticus* DSM 2522 (ZP_06364416.1; 59% identity) and *Geobacillus* sp. Y412MC52 (ZP_04393775.1; 59% identity). Conserved domain searches suggested that this protein (designated M.LmoJ2P) was a cytosine-5 methyltransferase (pfam00145). Nine of the ten conserved cytosine-5 methyltransferase motifs were identified in M.LmoJ2P, with one (motif IV) missing (data not shown).

The deduced polypeptide encoded by the second ORF, *lmoJ2R*, was similar to proteins encoded by various bacteria, including *E. faecalis* T11 (ZP_05596807.1; 59% identity), *B. cellulosilyticus* DSM 2522 (ZP_06364415.1; 47% identity) and *Geobacillus* sp. Y412MC52 (ZP_04393774.1; 45% identity). Notably, the *lmoJ2R* homologs in these bacteria were located adjacent to *lmoJ2M* homologs, suggesting that these two genes were harbored on one cassette in these genomes. Conserved domain searches suggested that the product of

*lmoJ2R* (designated R.LmoJ2P) belonged to the protein family (pfam09491) represented by the AlwI restriction endonuclease (recognition sequence: GGATC$(N)_4$↓).

The LmoJ2 recognition site was predicted to be GCWGC (W=A or T) by BLAST searches against REBASE. Proteins with highest homology to M.LmoJ2P in REBASE were M.Gsp412ORF3572P (YP_003254598.1) in *Geobacillus* sp. Y412MC61 plasmid pGYMC6101, and M.Bce98ORF752P (YP_001374089.1) in *B. cereus* subsp. cytotoxis NVH 391-98. For R.LmoJ2P, the REBASE database homologs included Gsp412ORF3572P (YP_003254599.1) and Bce98ORF752P (YP_001374090.1), both of which were adjacent to their homologs to M.LmoJ2P. The predicted recognition site for all these putative methyltransferases and restriction endonucleases was GCWGC (W=A or T); hence, an inference was drawn that LmoJ2 might also recognize GCWGC.

Sequence analysis revealed a putative promoter upstream of *lmoJ2R* along with a putative Rho-independent terminator in the intergenic region between *lmoJ2R* and the threonine aldolase gene (Fig. 1). Putative ribosome-binding sites were detected in front of *lmoJ2M* and *lmoJ2R*. Noticeably, the ribosome-binding site for *lmoJ2R* (AGGCTACT) deviated considerably from the consensus sequence.

Recognition sites in the vicinity of the promoter region play an important role in the control of some RM systems by reducing transcription when the sites are methylated (25). The sole LmoJ2 recognition site (GCAGC) in LmoJ2 and its proximal region was located within the methyltransferase gene (Fig. 1) and thus would not be expected to be involved in the regulation of LmoJ2.

**Sequence analysis of LmoJ3.** LmoJ3 was identified from the analysis of the sequence of the 3.7 kb PCR product obtained from *L. monocytogenes* strain J3115. Three ORFs were identified, designated *lmoJ3M* (984 bp), *lmoJ3R* (1170 bp), and *lmoJ3C* (384 bp) (Fig. 1). The GC contents of these ORFs were 29% (*lmoJ3M*) and 32% (*lmoJ3R* and *lmoJ3C*). As described above for LmoJ2, this was significantly below the genome average and suggests horizontal gene transfer.

The deduced product of *lmoJ3M* shared similarity with DNA methyltransferases such as Bsp6IM in *Bacillus* spp. (P43420.1; 65% identity) and *Staphylococcus aureus* RF122 (YP_417819.1; 58% identity). Conserved domain searches of the deduced polypeptide (designated M.LmoJ3P) assigned it to the cytosine 5-methyltransferease protein family (pfam00145). Comparisons with M.HhaI revealed that M.LmoJ3P harbored all ten conserved motifs of cytosine 5-methyltransferases (data not shown).

The second ORF, *lmoJ3R*, was predicted to encode a protein homologous to hypothetical proteins in several other bacteria including *Finegoldia magna* ATCC 29328 (YP_001691161.1; 46% identity), *Flavobacterium johnsoniae* UW01 (YP_001193007.1; 44%) and *B. weihenstephanensis* KBAB4 (YP_001643707.1; 43% identity). The deduced *lmoJ3R* product (designated R.LmoJ3P) belonged to the protein family (pfam09565) represented by the restriction endonuclease NgoFVII (recognition sequence: GCSGC, where S=C or G) (25).

The deduced product of the third ORF, *lmoJ3C*, exhibited homology to proteins from several bacteria, including hypothetical proteins of *Clostridium ramosum* DSM 1402 (ZP_02426636.1; 44% identity) and *Parvimonas micra* ATCC 33270 (ZP_02094636.1; 43%

identity), and a predicted transcriptional regulator in *Gordonibacter pamelaeae* 7-10-1b (CBL03716.1; 40% identity). The deduced protein (designated C.LmoJ3P) harbored a helix-turn-helix motif (pfam01381), suggesting that it is a DNA-binding protein that may function as control (C) protein in the regulation of the RM cassette (22).

BLAST analysis against REBASE identified M.LmoJ3P homologs with specificity for GCNGC, including M.Bsp6I in *Bacillus* spp. plasmid pXH13 (CAA57293.1) and M.MagORF4250P in *Mycoplasma agalactiae* strain 5632 (YP_003515594.1). Several homologs of R.LmoJ3P were also identified with specificity for GCNGC, such as Bli1904II in *B. licheniformis* SJ1904 (NCBI number unknown) and FjoUWORF652P in *F. johnsoniae* UW101 (ZP_01247103.1). These results suggested that R.LmoJ3 might also recognize GCNGC.

A putative promoter was identified upstream of *lmoJ3M* as well as upstream of *lmoJ3C* and bidirectional Rho-independent terminators were identified between *lmoJ3M* and *lmoJ3R* (Fig. 1).These predictions suggest that two transcripts may be generated from LmoJ3: one for *lmoJ3M* and a divergent transcript for *lmoJ3C* and *lmoJ3R*. Putative ribosome-binding sites were identified in front of each of the three ORFs of LmoJ3.

Four GCNGC sites were identified in the LmoJ3 region (Fig. 1). Two of these sites (GCGGC and GCTGC) were located close to each other between the putative *lmoJ3M* promoter and the start of *lmoJ3M* (Fig. 1), raising the possibility that they might be involved in the regulation of the RM system.

**LmoJ2 and LmoJ3 are harbored in the same genomic region by additional isolates of *L. monocytogenes*.** Hybridization of genomic DNA from other *L. monocytogenes*

isolates in our strain collection with DNA probes derived from LmoJ2 and LmoJ3 identified several that putatively harbored these RM systems. Evidence for LmoJ2 was obtained for isolates of serotypes 4b, 1/2a (or 3a), and 4c, as well as isolates the serotype of which could not be determined based on the multiplex PCR assay employed (Table 1). Evidence for LmoJ3 was obtained for two additional isolates of serotypes 4b and 1/2a, respectively (Table 1).

PCR with primers derived from flanking genes (*LMOf2365_0322* and *LMOf2365_0330*) and genes internal to LmoJ2 and LmoJ3 indicated that LmoJ2 in OLM 115 and OLM 116, and LmoJ3 in 2010-0072B were present in the same genomic location as in J2479 and J3115 (Fig. 2). Identical results were obtained with all the other isolates harboring LmoJ2 or LmoJ3 (data not shown). The size of the PCR products was also conserved among different isolates harboring LmoJ2 or lmoJ3, suggesting lack of detectable insertions or deletions in this region (Fig. 2).

Relatedness of these isolates was examined via PFGE (Fig. 3), MLGT and TMLGT (Table 1). These analyses revealed that LmoJ2 was harbored by several strains of unique genotypes (including two lineage III strains of serotype 4b) as well as by strains of two distinct clusters (one consisting of the lineage I serotype 4b isolates OLM 80, OLM 115 and OLM 116, and the other composed of four lineage II serotype 1/2a (or 3a) isolates from foods and the environment of processing plants) (Fig. 3). The two serotype 4b strains harboring LmoJ3 were of lineage I and clearly distinct from each other (Fig. 3). Thus, LmoJ2 as well as LmoJ3 were harbored by a genetically heterogeneous group of isolates. It was noteworthy, however, that 7/14 of the LmoJ2-bearing strains were of lineage III, including

isolates of serotype 4b and 4c. These lineage III strains exhibited PFGE profiles markedly different form each other and from those of other isolates (Fig. 3).

**Genomic DNA from strains harboring LmoJ2 and LmoJ3 was methylated at GCWGC and GCNGC, respectively, when bacteria were grown planktonically or on agar at different temperatures, including 4˚C.** To verify the recognition site of the novel RM systems and obtain evidence for expression of the DNA methyltransferases, genomic DNA was extracted from LmoJ2 or LmoJ3 isolates and digested with ApeKI (recognition sequence: G↓CWGC) and Fnu4HI (recognition sequence: GC↓NGC). In addition, DNAs were digested with HindIII (recognition sequence: A↓AGCTT) to verify that they were digestible by an enzyme not affected by the novel RM systems.

Genomic DNA from LmoJ2 and LmoJ3 isolates grown at 37˚C on agar was resistant to ApeKI and Fnu4HI, respectively. This resistance was consistently exhibited by all LmoJ2 and LmoJ3 (Fig. 4 and data not shown). Similar results were observed when J2479, J3115 and other representative isolates with LmoJ2 and LmoJ3 were grown at 37˚C in liquid and at other temperatures (4, 8, 25, and 42˚C) in liquid or on agar (data not shown). To confirm that LmoJ2 was specific for GCWGC and not GCNGC we also digested genomic DNA from LmoJ2 isolates with Fnu4HI. DNA from these cultures exhibited susceptibility to this enzyme (Fig. 4 and data not shown). Taken together, these findings support the hypothesis that LmoJ2 and LmoJ3 recognize GCWGC and GCNGC, respectively. The findings also indicate that the methyltransferases of these RM systems were expressed at all tested temperatures, including 4 and 8˚C, and regardless of whether the bacteria were grown on agar or planktonically.

**Evidence for expression of LmoJ2 and LmoJ3-associated restriction endonucleases.** Expression of the putative restriction endonucleases associated with LmoJ2 and LmoJ3 was assessed by determining whether *Listeria* phage 20422-1 propagated in LmoJ2 or LmoJ3 strains possessed higher EOP values in these strains than when it was propagated in a reference strain deficient in these RM systems (DP-L862), and would thus lack modification at sites recognized by the putative restriction endonucleases in LmoJ2 or LmoJ3. The derivatives of 20422-1 propagated in the LmoJ2-harboring strain J2479 and the LmoJ3-harboring strain J3115 were designated 20422-1MJ2 and 20422-MJ3, respectively.

When the LmoJ2 strain J2479 was infected with 20422-1, EOP was ca. 2.5-7.9 fold lower than that obtained with the reference strain DP-L862 at all tested growth temperatures (4, 8, 25, 37 and 42ºC) ($p<0.05$) (Fig. 5). When J2479 was infected with 20422-1MJ2, EOP increased (by $10^{0.7}$ to $10^{0.9}$, i.e., ca. 5.0-7.9 fold) in cells grown at 4 or 8ºC ($p=0.001$ and 0.007, respectively); however, no significant change in EOP was noted when cells were grown at 25, 37 or 42ºC (Fig. 5). These data suggest that the LmoJ2 restriction endonuclease was moderately expressed, and only at low temperatures, conferring a degree of protection from phage attack. Infection of J2479 with 20422-1MJ3 did not confer additional protection in comparison with 20422-1MJ2 (Fig. 5).

Infection of the LmoJ3 strain J3115 with 20422-1 revealed that the EOP was $10^{-3.4}$-$10^{-5.8}$, significantly lower than with the reference strain ($p<0.0001$), at all tested growth temperatures. A pronounced ($10^{3.3}$ to$10^{5.9}$fold) increase in EOP was noted upon infection with 20422-1MJ3 ($p<0.0001$) (Fig. 5). The extent of the increase was lower when bacteria were grown at 25ºC ($10^{3.3}$fold increase than at other temperatures) ($p<0.0001$). A significant,

albeit smaller, increase in EOP was also observed following infection of J3115 with 20422-1MJ2 ($p<0.0001$) (Fig. 5). Similar findings were obtained with an additional strain harboring LmoJ3 (2007-584) grown at 37ºC (data not shown).

**LmoJ2 and LmoJ3 restriction sites are significantly less frequent in the genomes of lytic compared with temperate *Listeria* phages.** The number of recognition sites for a restriction endonuclease is correlated to the expected susceptibility of the phage to the corresponding restriction system (23). In fact, reduction of the target sites is often adopted by bacteriophages as a tactic to decrease the impact of RM systems (32). Analysis of sequenced *Listeria* phage genomes revealed that the genome of lytic phages harbored these recognition sequences at markedly lower frequencies than those of temperate phages, regardless of phage family (*Siphoviridae* vs. *Myoviridae*) or genome size (Table 3). An extreme case of this trend was observed in the *Siphoviridae* lytic phages P35 and P40 that possessed only one and two GCWGC sites, respectively (Table 3). In contrast, temperate phages harbored LmoJ2 and LmoJ3 recognition sites at frequencies similar to those of chromosomal fragments of *L. monocytogenes* F2365 (Table 3).

To assess relative susceptibility to LmoJ2 vs. LmoJ3 we calculated the GCWGC/GCNGC ratio in phage genomes and *L. monocytogenes* F2365 chromosomal segments (Table 3). In *Siphoviridae* temperate phages, approximately half of the LmoJ3 recognition sites were GCWGC, similarly to the chromosomal fragments. In contrast, the GCWGC/GCNGC ratio was lower (<25%) in *Siphoviridae* lytic phages and higher (ca. 85%) in *Myoviridae* phages.

Close scrutiny of the distribution of the GCWGC and GCNGC sites revealed that they were unevenly distributed in the phage genomes. In many phage genomes such regions included a gene encoding tape measurement protein (Table 3). In *Siphoviridae* lytic phages, all GCWGC and the majority of GCNGC sites were located in this gene, and in temperate phages, this gene also harbored considerably more GCWGC and GCNGC sites than the average for the phage genome (Table 3). In *Myoviridae* lytic phages, most GCWGC and GCNGC sites were located in tail lysin genes. In both the tape measurement protein and tail lysin genes, the GCWGC/GCNGC ratio was similar to that observed for the phage genomes (Table 3).

## DISCUSSION

In this study, we presented evidence for two novel Type II RM systems (LmoJ2 and LmoJ3) in the same locus of *L. monocytogenes* in which ECI strains harbor a Sau3AI-like RM system. The LmoJ2 and LmoJ3 RM systems consisted of genes of unusually low GC content and were identified in genetically diverse strains of multiple serotypes, belonging to different lineages. Such findings imply that these RM systems may have been acquired through site-specific horizontal gene transfer, probably via multiple, independent events. Alternatively, the novel RM systems might have been acquired by an ancestral strain of *L. monocytogenes* but were maintained only in certain strains and (in the case of LmoJ2) certain strain clusters. The serotype 4b cluster with LmoJ2 was of special interest, as it consisted exclusively of isolates from human listeriosis cases from the relatively distant past; isolates of this strain type were not identified among the >140 serotype 4b clinical isolates from

recent years (2003-2008) (Chapter 7). Our bioinformatics data and methylation assays suggested that the recognition sites of LmoJ2 were included within those of LmoJ3 and that LmoJ2 and LmoJ3 recognized GCWGC and GCNGC, respectively. In addition, for LmoJ3-harboring strains, phage DNA methylation by LmoJ2 increased infection efficiency compared with the unmethylated phage, albeit to a lesser degree than methylation by LmoJ3, providing further evidence for LmoJ2 recognition sites being subset of those for LmoJ3. Cytosine methylation by LmoJ2 and LmoJ3 is expected to render DNA resistant to ApeKI and Fnu4HI, respectively, and these enzymes are both sensitive to cytosine methylation at GpC sites (3), suggesting that LmoJ2 and LmoJ3 likely methylate cytosine at GpC sites within their recognition sites.

The recognition sites of these RM systems attract attention due to three features. One is their relative GC-richness, considering that the genome of *L. monocytogenes* exhibits only 38% of GC content (26). Nonetheless, we found that the *L. monocytogenes* chromosome harbored an average of at least one site per kb. Abundant in G and C nucleotides as these sites are, their small size may result in this high frequency. Another noteworthy attribute is the inclusion of LmoJ2 restriction sites within those for LmoJ3. These overlapping recognition sites might have ecological implications for strains bearing each RM system and occupying the same environmental niche, and may account for the observation that, even though LmoJ3 conferred the larger protection against phage infection, LmoJ2-harboring strains were more common. Although LmoJ2 might provide only moderate protection against phage and under limited conditions, this RM system may convert its host to a biological reservoir for methylated phages that can counteract, albeit incompletely, restriction by

LmoJ3, thus potentially offsetting the competitive edge conferred by LmoJ3. These interactions impact the balance between LmoJ2 and LmoJ3-harboring strains, thus contributing to coexistence of strains harboring these two RM systems. Lastly, cytosine methylation-susceptible restriction endonucleases recognizing GCWGC and GCNGC have been employed to identify cytosine methylation of CpG and GpC sites in eukaryotic DNA within trinucleotide repeat tracks associated with human neurodegenerative and neuromuscular diseases such as Huntington's disease and fragile X syndrome (3, 30). ApeKI and Fnu4HI are examples of such enzymes (3, 30) and the novel RM systems described here might expand the repertoire of tools to investigate biomedically relevant epigenetic changes in the human genome.

Genomic DNA digestions suggested that methyltransferases encoded by LmoJ2 and LmoJ3 methylate the genomic DNA at all tested temperatures, including 4 and 8ºC. Since cells could die if even one site of the genomic DNA is not methylated in the presence of an active restriction enzyme (16, 25), it is reasonable that the methyltransferases would be expressed constitutively. However, phage infection assays revealed that methylation of the phage DNA by propagation in an LmoJ2-harboring strain enhanced infectivity only moderately, and only at low temperature. Since such phage particles exhibited increase in infectivity in LmoJ3 strains, their DNA seems to be methylated; thus, we suspect that the LmoJ2 restriction enzyme was expressed at low levels, and only at low temperatures. On the other hand, when phage DNA was methylated by multiplication in a strain harboring LmoJ3, infection efficiency increased dramatically at all tested temperatures, suggesting that the

LmoJ3 restriction endonuclease was expressed at all tested temperatures, including 4 and 8ºC.

Activity of the restriction endonucleases at low temperatures and accompanying impact on resistance to phage infection might reflect a strategy against *Listeria*-specific phages in low-temperature environments. Thus, the novel RM systems may confer a fitness advantage to the strains harboring them, especially at low temperature. Given that refrigeration is widely utilized in food processing and distribution, this advantage may promote dissemination of LmoJ2 and LmoJ3-bearing strains in the food chain. Conditions other than temperature may also be involved in the expression of the novel RM systems and it is possible that under certain conditions, a higher level of restriction by LmoJ2 may be observed. Elucidation of such conditions could facilitate our understanding on the impact of these RM systems on the ecology of *L .monocytogenes*.

Influence of temperature on RM system expression was previously reported for LlaJI on lactococcal plasmid pNP40, which conferred more protection against phage infection at 19ºC than at higher temperatures (28). Transcriptional level of the LlaJI was higher at 37ºC than at 25 or 30ºC, which does not correspond to the phage infection data (28). Overexpression of the LlaJI restriction endonuclease shifted the temperature for optimal protection against phage to 30ºC, suggesting that temperature-dependence might be related to the intracellular concentration of the active restriction endonucleases (29). Although it remains to be identified what mechanism is responsible for the impact of temperature on the activity of LmoJ2 and LmoJ3, these previous findings are indicative of delicate control mechanisms that these temperature-sensitive RM systems might be subjected to.

Lytic phages tend to possess fewer recognition sites for RM systems than temperate phages (32). Analysis of the frequency of LmoJ2 and LmoJ3 recognition sites within the genomes of temperate and lytic *Listeria* phages indeed revealed significantly lower frequency in the latter. This trend is worth contemplating given that the majority of *Listeria* phages are following the temperate lifestyle (20), and that their expected higher susceptibility to the LmoJ2 and LmoJ3 restriction enzymes may reduce horizontal gene transfer frequently mediated by prophages and often involving genes for toxins and virulence factors (20, 27). The relative scarcity of LmoJ2 and LmoJ3 recognition sites in lytic phages might arise from greater selective pressure to reduce the restriction sites in comparison with temperate phages that can survive within the cell once the genome is integrated into the chromosome.

Meanwhile, it should be mentioned that the lytic phages are only "relatively" well equipped to defend themselves from the novel RM systems. In fact, the phage used in this study, 20422-1, belongs to lytic phages and was restricted by LmoJ2 and LmoJ3 (Fig. 4), suggesting that lytic phages are still susceptible to these RM systems. Since 20422-1 displayed a high homology with the lytic phages P100 and A511 (14), these phages are expected to be similarly affected by LmoJ2 and LmoJ3.

When the ratio of LmoJ2 restriction sites (GCWGC) to those of LmoJ3 (GCNGC) was analyzed, GCWGC accounted for ca. 50% of GCNGC sites in *Siphoviridae* temperate phages. GCWGC sites were under-represented in the *Siphoviridae* lytic phages and over-represented in *Myoviridae* phages. The low GCWGC/ GCNGC ratio observed for *Siphoviridae* lytic phages might be associated with the relatively high GC content of these phages. However, except for these phages, the overall GC content range of 35-37% does not

seem to be great enough to explain these trends; thus, we speculate that the main driving force behind these phenomena might be how often in the past a phage encountered RM systems recognizing GCWGC compared with those recognizing GCNGC. For instance, relatively fewer GCWGC sites in *Myoviridae* phages might reflect that these phages might have more often infected cells harboring a RM system recognizing GCWGC than cells bearing a GCNGC-recognizing RM system.

All the phage genomes analyzed in this study possessed some regions with a high density of GCWGC and GCNGC sites, for instance, the tape measurement protein gene in all the temperate phages and *Siphoviridae* lytic phages and the tail lysin gene in *Myoviridae* lytic phages. It remains unknown why these restriction sites were densely packed in a certain region. However, the evolutionary process of reducing restriction sites as an attempt to counteract RM systems might exert a greater impact on sequence content of these genes due to the mutations accompanying this defense mechanism against LmoJ2 and LmoJ3. Also, mutagenesis toward fewer restriction sites might proceed more promptly in the region rich in restriction sites since adjacent or overlapping restriction sites could be more efficiently eliminated with fewer nucleotide changes.

In conclusion, we have demonstrated that two novel RM systems (LmoJ2 and LmoJ3) are present in the same genetic location of various strains of *L. monocytogenes*, and may have been acquired through horizontal gene transfer. The genomic region where LmoJ2 and LmoJ3 are localized also harbors the Sau3AI-like RM system in ECI strains, suggesting a site-specific mechanism for acquisition of these Type II RM systems in *L. monocytogenes*. LmoJ2 and LmoJ3 recognize the closely related DNA sequences GCWGC and GCNGC,

respectively, and they methylate genomic DNA and confer at least partial protection against phage infection. Both RM systems appeared to be active at low temperatures, which might be important in the ecology of the bacteria through the food chain. Further studies on LmoJ2, LmoJ3 and other RM systems in this locus will enhance our understanding of the ecological and evolutionary roles of genetic elements that bestow DNA methylation and phage resistance to *L. monocytogenes* and will provide information needed to assess the effectiveness of *Listeria* phages as biocontrol agent.

**Table 4.1 Bacterial strains used in this study**

| Strain | Isolation year | Origin | Serotype[a] | Lineage[b] (MLGT) | Source[c] |
|---|---|---|---|---|---|
| **<LmoJ2>** | | | | | |
| J2479 | 2003 | Clinical | 4b | III (Lm3.42) | CDC |
| (HI) 267 | 1999 | Clinical | 4b | III (Lm3.49) | NCSU |
| NRRL B-33191 | 2001 | Animal | 4b | III (Lm3.13) | (38) |
| OLM 80 | 1954 | Clinical | 4b | I (1.70_4b) | NCSU |
| OLM 115 | 1955 | Clinical | 4b | I (1.70_4b) | NCSU |
| OLM 116 | 1956 | Clinical | 4b | I (1.70_4b) | NCSU |
| #191A[d] | 2004 | Processing plant | 1/2a (or 3a) | II (2.30_1/2a_T189) | (24) |
| L1522a[e] | 2005 | Processing plant | 1/2a (or 3a) | II (2.30_1/2a_T189) | (24) |
| LW-A75 | 2003 | Food (cold-smoked salmon) | 1/2a (or 3a) | II (2.20_1/2a_T492) | FDA |
| LW-A97 | 2004 | Food (avocado puree) | 1/2a (or 3a) | II (2.20_1/2a_T492) | FDA |
| 2008-392 | 2008 | Clinical | 4c | III (Lm3.46) | NCDHHS |
| NRRL B-33330 | 2002 | Food (liquid unpasteurized whole egg) | Indeterminate | III (Lm3.17) | (38) |
| OLM 62 | 1953 | Animal | Indeterminate | III (Lm3.42) | NCSU |
| NRRL B-33372 | 2004 | Animal | Indeterminate | III (Lm3.26) | (38) |
| **<LmoJ3>** | | | | | |
| J3115 | 2004 | Clinical | 4b | I (1.73_4b) | CDC |
| 2007-584 | 2007 | Clinical | 4b | I (1.74_4b) | NCDHHS |
| 2010-0072B | 2010 | Clinical | 1/2a (or 3a) | II (2.92_1/2a) | NCDHHS |

**Table 4.1 Continued**

[a] Serotype was determined either by serological assay or by the multiplex PCR described by Doumith et al. (7). "Indeterminate" indicates that no serotype designation could be assigned based on the multiplex PCR (7).

[b] Lineage was designated based on TMLGT (39).

[c] CDC, Centers for Disease Control and Prevention; NCSU, *Listeria* Strain collection at North Carolina State University; FDA, US Food and Drug Administration; NCDHHS, North Carolina Department of Health and Human Services.

[d] #191A was named 90 in the study by Mullapudi et al. (24).

[e] L1522a was named 1747 in the study by Mullapudi et al. (24).

**Table 4.2 DNA primers used in this study**

| Primer | Sequence | Target |
|---|---|---|
| J2479Res_F | GACCTAGTAAAGCAGGTGCT | *lmoJ2R* |
| J2479_3R | AGGTACCCATTCGATAGTCG | *lmoJ2R* |
| J3115_2F | TATCAGGCTTTCCCTGTCAA | *lmoJ3M* |
| J3115Met_R | GGTACTACAACCGAATTCCC | *lmoJ3M* |
| H7858_0334R | GTTCCCGAATCATTTCCAC | *LMOf2365_0322* |
| H7858_0338F | CTCGTGAATCTCCAAATGCG | *LMOf2365_0330* |

**Table 4.3 Frequency of LmoJ2 and LmoJ3 recognition sites (GCWGC and GCNGC, respectively)**

| DNA | Family | Host serotype | Life cycle | Size (kb) | GC% | GCNGC per kb | GCWGC per kb | GCNGC and GCWGC-rich gene | GC% | GCNGC per kbp | GCWGC per kb | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| F2365 fragment 1[a] | | | | 130.0 | 38 | 2.35 | 1.28 (55%)[b] | | | | | (26) |
| F2365 fragment 2 | | | | 130.0 | 38 | 2.61 | 1.57 (60%) | | | | | (26) |
| F2365 fragment 3 | | | | 130.0 | 37 | 2.32 | 1.51 (65%) | | | | | (26) |
| F2365 fragment 4 | | | | 130.0 | 40 | 3.07 | 1.62 (53%) | | | | | (26) |
| F2365 fragment 5 | | | | 130.0 | 39 | 2.52 | 1.52 (60%) | | | | | (26) |
| Average | | | | | | 2.57 | 1.50 (58%) | | | | | |
| A500 | *Siphoviridae* | 4 | Temperate | 38.9 | 37 | 2.60 | 1.47 (57%) | LiPA500_gp016 (tmp[c]) | 40 | 8.20 | 3.73 (45%) | (6) |
| A118 | *Siphoviridae* | ½ | Temperate | 40.8 | 36 | 2.45 | 1.25 (51%) | A118p16 (tmp[c]) | 39 | 7.43 | 4.46 (60%) | (2) |
| A006 | *Siphoviridae* | ½ | Temperate | 38.1 | 36 | 2.13 | 0.94 (44%) | LiPA006_gp15 (tmp[c]) | 37 | 3.96 | 2.50 (63%) | (6) |
| B025 | *Siphoviridae* | 5, 6 | Temperate | 42.7 | 35 | 1.83 | 1.15 (62%) | LiPB025_gp14 (tmp[c]) | 38 | 5.69 | 3.66 (64%) | (6) |
| PSA | *Siphoviridae* | 4 | Temperate | 37.6 | 35 | 1.30 | 0.80 (61%) | 2389gp14 (tmp[c]) | 38 | 4.22 | 2.92 (69%) | (42) |
| B054 | *Myoviridae* | 5, 6 | Temperate | 48.2 | 36 | 1.97 | 1.58 (80%) | LiPB054_gp18 (tmp[c]) | 39 | 2.54 | 2.12 (83%) | (6) |
| P100 | *Myoviridae* | ½, 4, 5 | Lytic | 131.4 | 36 | 0.26 | 0.22 (84%) | LBPV100_gp028 (tail lysin) | 42 | 3.84 | 3.29 (86%) | (2) |
| A511 | *Myoviridae* | ½, 4 | Lytic | 137.6 | 36 | 0.24 | 0.21 (88%) | LiPA511_gp028 (tail lysin) | 42 | 3.84 | 3.29 (86%) | (15) |
| P40 | *Siphoviridae* | ½, 4, 5, 6 | Lytic | 35.6 | 41 | 0.25 | 0.06 (24%) | P40_gp14 (tmp[c]) | 45 | 2.74 | 1.09 (40%) | (6) |
| P35 | *Siphoviridae* | ½ | Lytic | 35.8 | 39 | 0.25 | 0.03 (12%) | LiPP35_gp14 (tmp[c]) | 44 | 3.18 | 0.53 (27%) | (6) |

**Table 4.3 Continued**

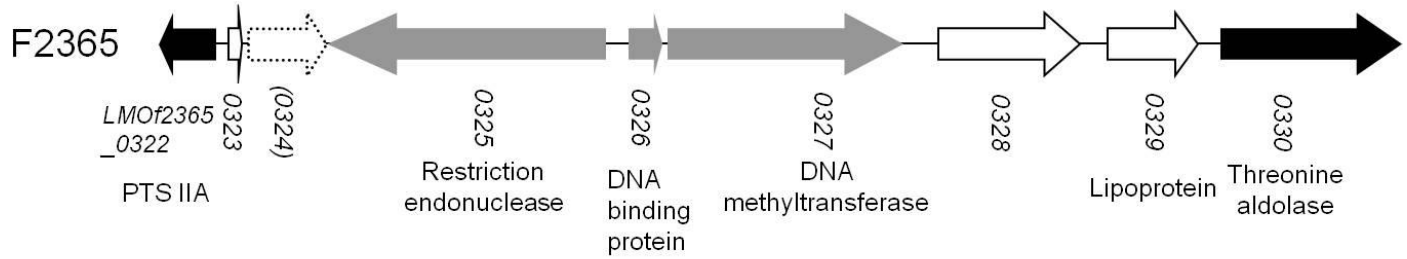
[a] F2365 fragments 1-5 correspond to randomly chosen DNA sequences, in this case, between 70,000-200,000, 650,000-780,000, 1230,000-1360,000, 1810,000-1940,000, and 2390, 000-2520,000 nt of F2365 genome (NC_002973.6), respectively.

[b] This percentage represents the proportion of GCWGC sites within GCNGC sites.

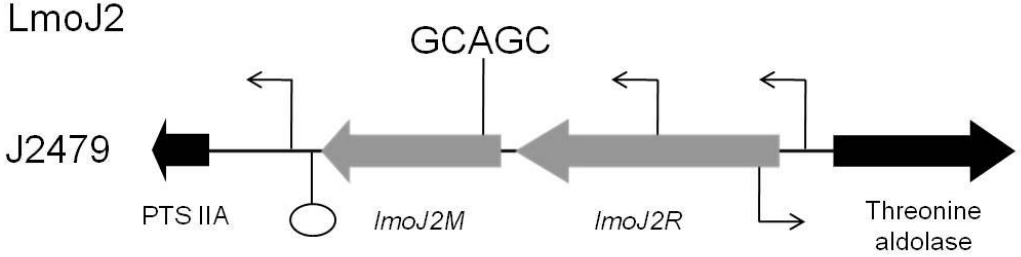
[c] Tmp, tape measurement protein.

**Figure 4.1 Organization of the genomic region harboring LmoJ2 and LmoJ3.** Gray arrows indicate genes comprising each RM system. Conserved flanking genes are shown as black arrows, and other genes unrelated to RM systems are indicated with white arrows. Stippled arrow indicates the pseudogene in strain F2365 (40). The location and direction of a predicted promoter is marked with a bent arrow, and putative Rho-independent terminators are shown as lollipop symbols. Locations of recognition sites for LmoJ2 and LmoJ3 are indicated and those that may be important in transcriptional control are underlined.

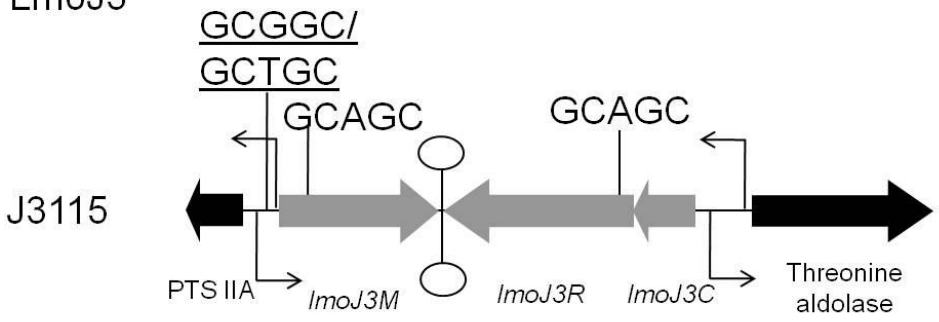
A. ECI Sau3AI-like R-M




B. LmoJ2




C. LmoJ3



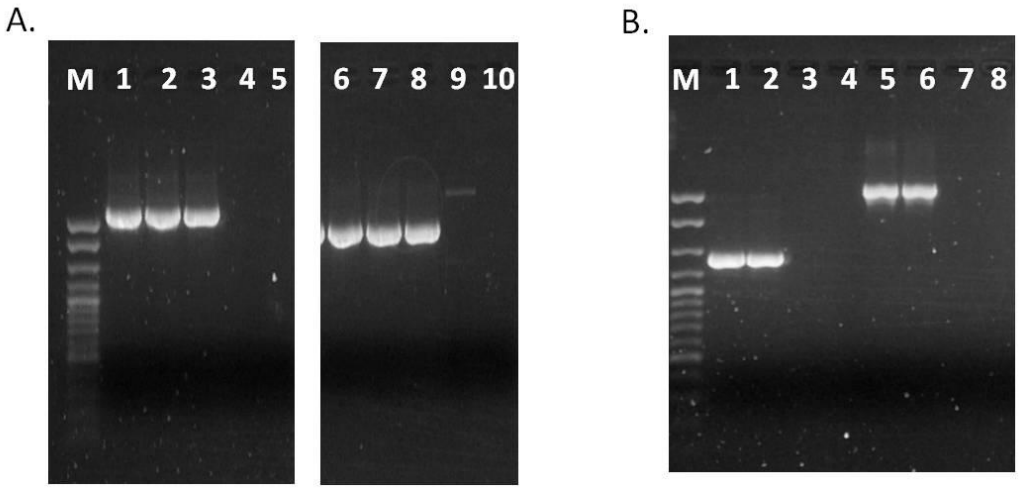



**Figure 4.2 LmoJ2 and LmoJ3 are present in the same genomic locations in isolates of diverse serotypes.** Confirmation of genomic location of LmoJ2 (A) was performed using PCR with primers H7858_0334R/ J2479_3R (lanes 1-5) and J2479Res_F/ H7858_0338F (lanes 6-10). Lanes 1 and 6, OLM 115; lanes 2 and 7, OLM 116; lanes 3 and 8, J2479 (used as the positive control); lanes 4, 5, 8 and 9, negative controls. Confirmation of genomic location of LmoJ3 (B) was done using PCR with primers H7858_0334R/ J3115Met_R (lanes 1-4) and J3115_2F/ H7858_0338F (lanes 5-8). Lanes 1 and 5, 2010-0072B; lanes 2 and 6, J3115, used as the positive control; and lanes 3, 4, 7 and 8, negative controls.

**Figure 4.3 PFGE dendrogram of isolates harboring LmoJ2 and LmoJ3.** PFGE with ApaI and AscI was performed as described in Materials and Methods. Isolates harboring LmoJ2 and LmoJ3 are indicated with filled and open circles, respectively. Serotype 4b strains F2365 (ECI), H7550 (ECII) and WS1 (ECV) were included for reference.

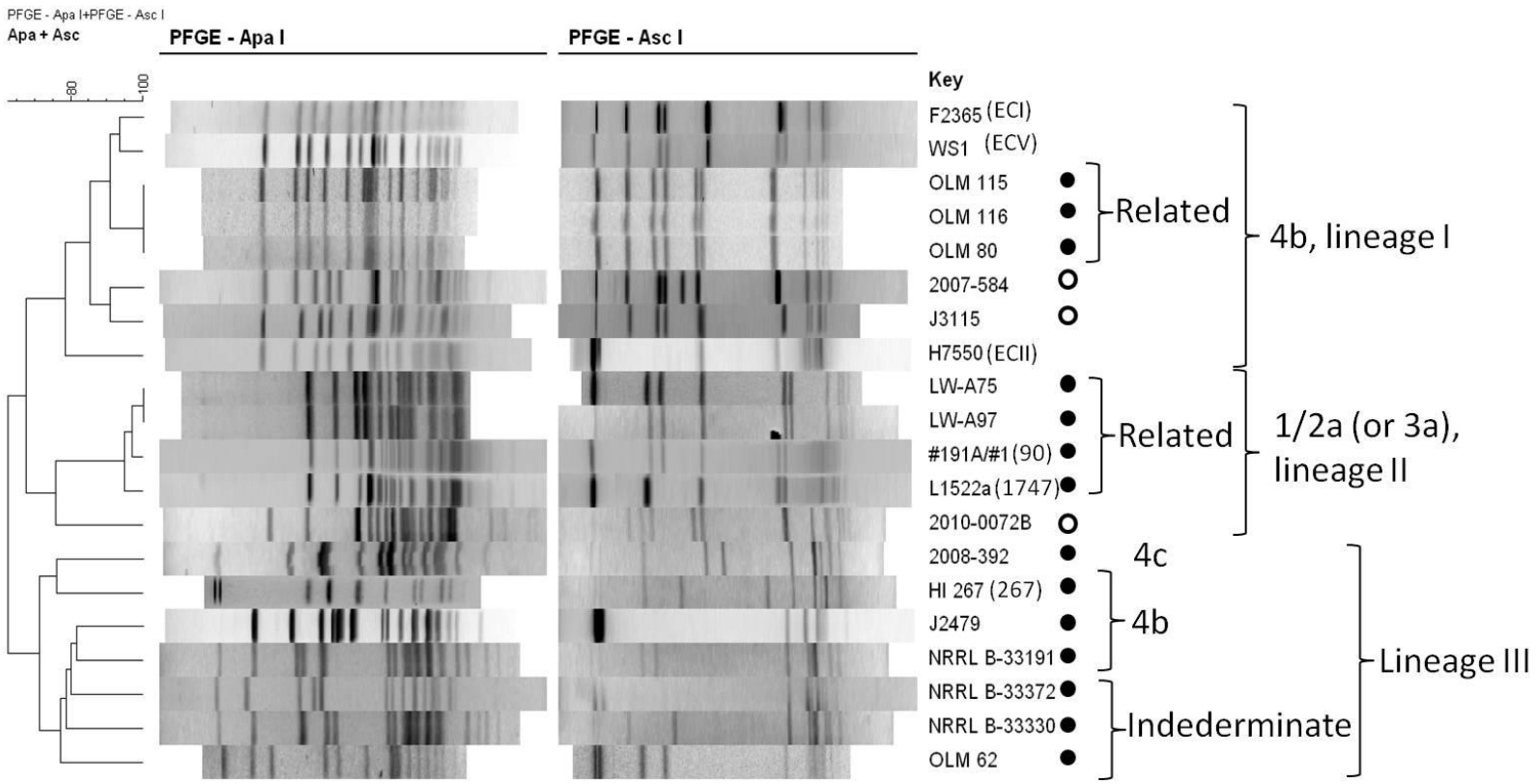
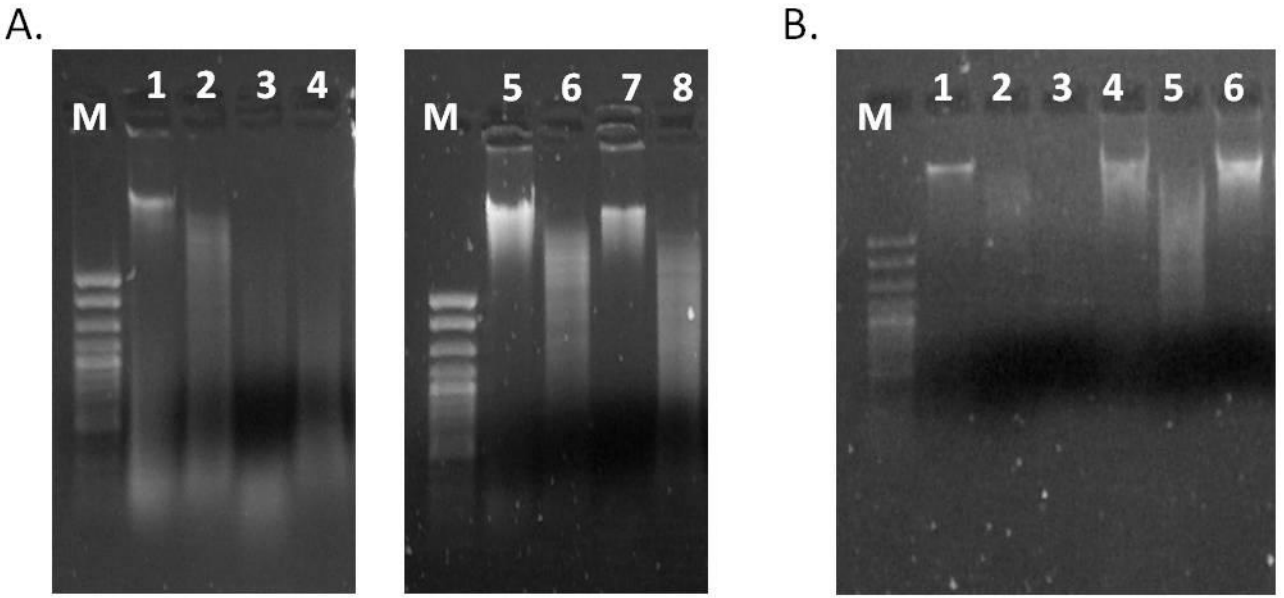

PFGE - Apa I+PFGE - Asc I
Apa + Asc
PFGE - Apa I
PFGE - Asc I
80
100
Key
F2365 (ECI)
WS1 (ECV)
OLM 115
OLM 116
OLM 80
2007-584
J3115
H7550 (ECII)
LW-A75
LW-A97
#191A/#1 (90)
L1522a (1747)
2010-0072B
2008-392
HI 267 (267)
J2479
NRRL B-33191
NRRL B-33372
NRRL B-33330
OLM 62
Related
4b, lineage I
Related
1/2a (or 3a), lineage II
4c
4b
Indederminate
Lineage III

**Figure 4.4 LmoJ2 and LmoJ3 are associated with resistance to digestion by restriction enzymes recognizing GCWGC and GCNGC, respectively.** Genomic DNA of strain LW-A75 and J3115, harboring LmoJ2 (A) or LmoJ3 (B), respectively, was digested with ApeKI (recognition sequence: G↓CWGC) and Fnu4HI (recognition sequence: GC↓NGC) as well as with HindIII (quality control). DNA was extracted from cells grown on BHI agar at 37˚C. The ECI strain F2365 that lacks LmoJ2 or LmoJ3 was included as a negative control. A. lane 1, undigested F2365 genomic DNA; lanes 2-4, F2365 genomic DNA digested with HindIII, ApeKI and Fnu4HI, respectively; lane 5, intact genomic DNA from LmoJ2 isolate LW-A75; and lanes 6-8, LW-A75 genomic DNA digested with HindIII, ApeKI and Fnu4HI, respectively. B. lane 1, undigested F2365 genomic DNA; lanes 2 and 3 , F2365 genomic DNA digested with HindIII and Fnu4HI; lane 4, intact genomic DNA from LmoJ3 isolate J3115; and lanes 5 and 6, J3115 genomic DNA treated with HindIII and Fnu4HI. M, exACTGene cloning DNA ladder (Fisher Scientific, Fair Lawn, NJ).


A.
M
1 2 3 4
M
5 6 7 8
B.
M
1 2 3 4 5 6

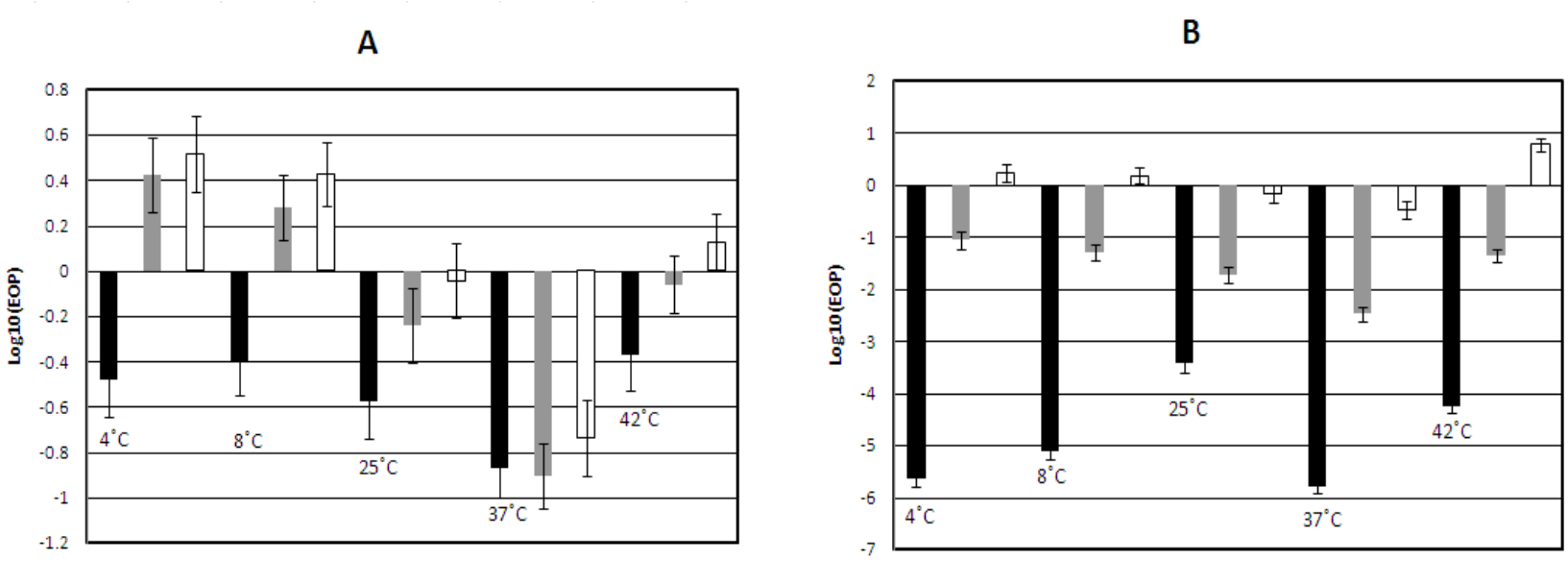



**Figure 4.5 Phage infection assay results of J2479 (A) and J3115 (B).** Black, grey and white bars represent the means of EOP values following infection with *Listeria* phage 20422-1 (propagated in a strain free of LmoJ2 or LmoJ3), 20422-1MJ2 (propagated in the LmoJ2-harboring strain J2479), and 20422-1MJ3 (propagated in the LmoJ3-harboring strain J3115), respectively. Error bars represent the standard error calculated from the statistical analysis.

## REFERENCES


1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403-410.

2. **Carlton, R. M., W. H. Noordman, B. Biswas, E. D. de Meester, and M. J. Loessner.** 2005. Bacteriophage P100 for control of *Listeria monocytogenes* in foods: genome sequence, bioinformatic analyses, oral toxicity study, and application. Regul. Toxicol. Pharmacol. **43:**301-312.

3. **Castel, A. L., M. Nakamori, C. A. Thornton, and C. E. Pearson.** 2011. Identification of restriction endonucleases sensitive to 5-cytosin methylation at non-CpG sites, including expanded (CAG)n/(CTG)n repeats. Epigenetics. **6:**1-5.

4. **Chan, Y. C., and M. Wiedmann.** 2009. Physiology and genetics of *Listeria monocytogenes* survival and growth at cold temperatures. Crit. Rev. Food Sci. Nutr. **49:**237-253.

5. **Cheng, Y., R. M. Siletzky, and S. Kathariou.** 2008. Genomic divisions/ lineages, epidemic clones, and population structure, p. 337-358. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

6. **Dorscht, J., J. Klumpp, R. Bielmann, M. Schmelcher, Y. Born, M. Zimmer, R. Calendar, and M. J. Loessner.** 2009. Comparative genome analysis of *Listeria* bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. J. Bacteriol. **191:**7206-7215.

7. **Doumith, M., C. Buchrieser, P. Glaser, C. Jacquet, and P. Martin.** 2004. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. J. Clin. Microbiol. **42:**3819-3822.

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

9. **Eifert, J. D., P. A. Curtis, M. C. Bazaco, R. J. Meinersmann, M. E. Berrang, S. Kernodle, C. Stam, L. A. Jaykus, and S. Kathariou.** 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog. Dis. **2:**192-200.

10. **Elhanafi, D., V. Dutta, and S. Kathariou.** 2010. Genetic characterization of plasmid-associated benzalkonium chloride resistance determinants in a *Listeria monocytogenes* strain from the 1998-1999 outbreak. Appl. Environ. Microbiol. **76:**8231-8238.

11. **Graves, L. M., and B. Swaminathan.** 2001. PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. Int. J. Food Microbiol. **65:**55-62.

12. **Kathariou, S.** 2003. Foodborne outbreaks of listeriosis and epidemic-associated lineages of *Listeria monocytogenes*, p. 243-256. *In* M. E. Torrence and R. E. Isaacson (eds.), Microbial food safety in animal agriculture, 1st ed. Iowa State Press, Ames, IA.

13. **Kathariou, S.** 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J. Food Prot. **65:**1811-1829.

14. **Kim, J. W., R. M. Siletzky, and S. Kathariou.** 2008. Host ranges of *Listeria*-specific bacteriophages from the turkey processing plant environment in the United States. Appl. Environ. Microbiol. **74:**6623-6630.

15. **Klumpp, J., J. Dorscht, R. Lurz, R. Bielmann, M. Wieland, M. Zimmer, R. Calendar, and M. J. Loessner.** 2008. The terminally redundant, nonpermuted genome of *Listeria* bacteriophage A511: a model for the SPO1-like myoviruses of gram-positive bacteria. J. Bacteriol. **190:**5753-5765.

16. **Kobayashi, I.** 2001. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. **29:**3742-3756.

17. **Kumar, S., X. Cheng, S. Klimasauskas, S. Mi, J. Posfai, R. J. Roberts, and G. G. Wilson.** 1994. The DNA (cytosine-5) methyltransferases. Nucleic Acids Res. **22:**1-10.

18. **Lemon, K. P., D. E. Higgins, and R. Kolter.** 2007. Flagellar motility is critical for *Listeria monocytogenes* biofilm formation. J. Bacteriol. **189:**4418-4424.

19. **Lesnik, E. A., R. Sampath, H. B. Levene, T. J. Henderson, J. A. McNeil, and D. J. Ecker.** 2001. Prediction of rho-independent transcriptional terminators in *Escherichia coli*. Nucleic Acids Res. **29:**3583-3594.

20. **Loessner, M. J., and E. D. R. Catherine.** 2005. *Listeria* phages: basics and applications, p. 362-279. *In* K. M. Waldor, D. I. Friedman, and S. L. Adhya (eds.), Phages: their role in bacterial pathogenesis and biotechnology. ASM Press, Washington, DC.

21. **Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M.**

**Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng, and S. H. Bryant.** 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. **39:**D225-9.

22. **McClelland, M., M. Nelson, and E. Raschke.** 1994. Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases. Nucleic Acids Res. **22:**3640-3659.
23. **Moineau, S., S. Pandian, and T. R. Klaenhammer.** 1993. Restriction/Modification systems and restriction endonucleases are more effective on lactococcal bacteriophages that have emerged recently in the dairy industry. Appl. Environ. Microbiol. **59:**197-202.
24. **Mullapudi, S., R. M. Siletzky, and S. Kathariou.** 2008. Heavy-metal and benzalkonium chloride resistance of *Listeria monocytogenes* isolates from the environment of turkey-processing plants. Appl. Environ. Microbiol. **74:**1464-1468.
25. **Nagornykh, M. O., E. S. Bogdanova, A. S. Protsenko, M. V. Zakharova, A. S. Solonin, and K. V. Severinov.** 2008. Regulation of gene expression in type II restriction-modification system. Genetika. **44:**606-615.
26. **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:**2386-2395.

27. **Ochman, H., J. G. Lawrence, and E. A. Groisman.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature. **405:**299-304.
28. **O'Driscoll, J., F. Glynn, O. Cahalane, M. O'Connell-Motherway, G. F. Fitzgerald, and D. van Sinderen.** 2004. Lactococcal plasmid pNP40 encodes a novel, temperature-sensitive restriction-modification system. Appl. Environ. Microbiol. **70:**5546-5556.
29. **O'Driscoll, J., D. F. Heiter, G. G. Wilson, G. F. Fitzgerald, R. Roberts, and D. van Sinderen.** 2006. A genetic dissection of the LlaJI restriction cassette reveals insights on a novel bacteriophage resistance system. BMC Microbiol. **6:**40.
30. **Ramsahoye, B. H., A. K. Burnett, and C. Taylor.** 1997. Restriction endonuclease isoschizomers ItaI, BsoFI and Fsp4HI are characterised by differences in their sensitivities to CpG methylation. Nucleic Acids Res. **25:**3196-3198.
31. **Roberts, R. J., T. Vincze, J. Posfai, and D. Macelis.** 2010. REBASE--a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res. **38:**D234-6.
32. **Rocha, E. P., A. Danchin, and A. Viari.** 2001. Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. Genome Res. **11:**946-958.

33. **Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin.** 2011. Foodborne illness acquired in the United States-major pathogens. Emerg. Infect. Dis. **17:**7-15.

34. **Thomas, C. M., and K. M. Nielsen.** 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. **3:**711-721.

35. **Tock, M. R., and D. T. Dryden.** 2005. The biology of restriction and anti-restriction. Curr. Opin. Microbiol. **8:**466-472.

36. **Wagner, M., and J. MacLauchlin.** 2008. Biology, p. 3-26. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

37. **Waibel, A. H., T. Hanazawa, G. E. Hinton, K. Shikano, and K. J. Lang.** 1989. Phoneme Recognition Using Time-Delay Neural Networks. IEEE Transactions on Acoustic, Speech, and Signal Processing. **37:**328-339.

38. **Ward, T. J., T. F. Ducey, T. Usgaard, K. A. Dunn, and J. P. Bielawski.** 2008. Multilocus genotyping assays for single nucleotide polymorphism-based subtyping of *Listeria monocytogenes* isolates. Appl. Environ. Microbiol. **74:**7629-7642.

39. **Ward, T. J., T. Usgaard, and P. Evans.** 2010. A targeted multilocus genotyping assay for lineage, serogroup, and epidemic clone typing of *Listeria monocytogenes*. Appl. Environ. Microbiol. **76:**6680-6684.

40. **Yildirim, S., D. Elhanafi, W. Lin, A. D. Hitchins, R. M. Siletzky, and S. Kathariou.** 2010. Conservation of genomic localization and sequence content of Sau3AI-like restriction-modification gene cassettes among *Listeria monocytogenes* epidemic clone I and selected strains of serotype 1/2a. Appl. Environ. Microbiol. **76:**5577-5584.

41. **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:**4158-4164.

42. **Zimmer, M., E. Sattelberger, R. B. Inman, R. Calendar, and M. J. Loessner.** 2003. Genome and proteome of *Listeria monocytogenes* phage PSA: an unusual case for programmed + 1 translational frameshifting in structural protein synthesis. Mol. Microbiol. **50:**303-317.

# Chapter 5 Functional Characterization of a Putative Lipoprotein in an Immigration Control Region of *Listeria monocytogenes*


## ABSTRACT

*Listeria monocytogenes* is the Gram-positive foodborne pathogen that causes listeriosis in humans and animals. An epidemic-associated clonal group of *L. monocytogenes* (Epidemic Clone I [ECI]) is resistant to Sau3AI digestion due to a Sau3AI-like restriction-modification (RM) system. Downstream to this RM system, we identified a putative lipoprotein gene (*LMOf2365_0329*) that was often absent or diversified in non-pathogenic *Listeria* species and animal isolates of *L. monocytogenes* while being largely present in isolates from human cases of illness and those of environmental/food origin. We characterized the function of this gene by constructing a deletion mutant in the background of the ECI strain F2365. The deletion did not influence cell morphology, hemolysis, growth at various temperatures, and phage infectivity. Also, no difference was observed with regards to resistance to ethanol, salt, $H_2O_2$, and disinfectants such as benzalkonium chloride (BC) and 1-hexadecylpyridinium chloride monohydrate (CPC). However, the deletion mutant exhibited (1) a moderate reduction in swarming on soft agar; (2) decreased MIC for nisin at 10 and 37ºC; (3) a higher MIC for polymyxin B at 42ºC; and (4) an increased viability in the presence of a lethal level of SDS. Most of these phenotypic differences were moderate or small, and genetic complementation did not restore swarming to wildtype levels. Further studies are warranted to assess the possible role of this gene, including its potential involvement in virulence.

## INTRODUCTION

*Listeria monocytogenes* is the Gram-positive bacterium responsible for listeriosis in humans and animals (25). Clinical symptoms of human listeriosis encompass mild flu-like symptoms and invasive manifestations such as septicemia, meningitis, encephalitis, and abortions (6). As a foodborne disease, listeriosis occurs at a relatively low incidence but its hospitalization rate and fatality are the second and third highest, respectively, among agents for foodborne illnesses in the United States (22). Meanwhile, *L. monocytogenes* inhabits a broad range of environmental niches including food-processing facilities and possesses the ability to form biofilm and endure various environmental stresses such as acid, low temperatures and high salt (14, 23, 27). These features render the control of this microorganism particularly challenging (27).

Of 13 serotypes of *L. monocytogenes* (9), over 95% of clinical cases are attributed to serotypes 1/2a, 1/2b and 4b (25). In particular, serotype 4b has been involved in numerous listeriosis outbreaks (25). Previous studies on outbreak strains belonging to this serotype revealed a closely related group of *L. monocytogenes* strains (Epidemic Clone I, ECI) that was implicated in multiple outbreaks that were not linked to each other geographically or chronologically (4). An exemplary outbreak caused by an ECI strain occurred in 1985 in Los Angeles due to contaminated cheese (17). The genome of the strain involved in this outbreak (F2365) has been sequenced to completion (17). One major characteristic of ECI strains is that their genomic DNA is not cleaved by the restriction endonuclease Sau3AI due to methylated cytosine at GATC sites (7). The genome of strain F2365 indeed harbors a Sau3AI-like restriction-modification (RM) system encoded by *LMOf2365_0325* to

*LMOf2365_0327* (Fig. 1A) evidenced to be mediating DNA resistance to Sau3AI (17, 29). Although certain strains of serotypes 1/2a and 4a harbor a highly homologous RM system (29), this RM system was identified only in ECI strains among serotype 4b; thus, a DNA probe derived from this RM system has been employed as a genetic marker for ECI (29).

In the ECI genome, the Sau3AI-like RM system is flanked at one side by a gene encoding a putative lipoprotein (*LMOf2365_0329*) (Fig. 1A). In Gram-positive bacteria, lipoproteins are involved in various cellular functions including antimicrobial resistance, nutrient transport, adhesion, signaling, protein secretion and conjugation (24). In *L. monocytogenes*, lipoproteins have attracted special attention as the most abundant group of surface proteins and due to their pronounced diversity among different strains (3, 5).

Previous data from our laboratory (Chapter 3) revealed that the lipoprotein gene *LMOf2365_0329* had highly conserved homologs in completed *L. monocytogenes* genomes of serotypes 1/2a and 4b including other epidemic-associated clonal groups ECIa, II and IIII (97-100% identity at the nucleotide sequence level). Diversity was observed with an apparent duplication in the serotype 4a strain HCC23. In the non-pathogenic *Listeria* species, the gene exhibited divergence in *L. innocua* (77% identity) whereas *L. welshimeri* did not harbor a homolog. Hybridization with the DNA probe based on the lipoprotein gene against *L. monocytogenes* strains revealed that this gene was harbored by a majority of clinical and environmental/food isolates but under-represented among animal isolates, suggesting that this lipoprotein gene might provide advantages for clinical and environmental/food isolates, potentially with regards to stress tolerance and/or pathogenicity. In this study, we constructed

the deletion mutant of the lipoprotein gene in the ECI strain F2365 and attempted to characterize the function of this putative lipoprotein in growth and stress tolerance.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** Strains used in this study are listed in Table 1. All *Escherichia coli* strains were routinely grown overnight at 37ºC in Luria Bertani medium (LB; Becton, Dickinson and Company, Sparks, MD) or on LB plates supplemented with 1.2% agar (Becton, Dickinson and Company). *L. monocytogenes* strains were routinely cultured overnight at 37ºC in brain heart infusion (BHI; Becton, Dickinson and Company) or on BHI plates containing 1.2% agar. When necessary, growth media were supplemented with antibiotics or other agents, as indicated, at the designated concentrations. For long-term storage, overnight culture grown on agar plates were collected, resuspended in BHI with 20% glycerol (Fisher Scientific, Fair Lawn, NJ), and stored at -80ºC.

**DNA and protein sequence analysis.** The DNA and protein sequences were retrieved from the National Center for Biotechnology Information (NCBI) database. Restriction sites were examined with Webcutter 2.0 (http://bio.lundberg.gu.se/cutter2/) or NEBcutter 2.0 (http://tools.neb.com/NEBcutter2/). Promoters were predicted with the Neural Network Promoter Prediction provided by the Berkeley Drosophila Genome Project (http://www.fruitfly.org/seq_tools/promoter.html) (28). Potential Rho-independent transcription terminators in *L. monocytogenes* F2365 were retrieved from TransTermHP (http://transterm.cbcb.umd.edu/) (12). The protein family was predicted with NCBI Conserved Domain Search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) (16),

Pfam sequence search (http://pfam.sanger.ac.uk/) and Kyoto Encyclopedia of Genes and Genomes (http://www.genome.jp/kegg/).

**Genomic DNA extraction, Polymerase Chain Reaction (PCR) and sequencing.** Genomic DNA was extracted from overnight liquid cultures with the DNeasy Blood and Tissue Kit (QIAGEN, Valencia, CA) according to the manufacturer's protocol. Primers employed in this study were purchased from Eurofins MWG Operon (Huntsville, AL) and are listed in Table 2. PCR conditions were conducted at 95ºC for 5 min followed by 31 or 33 cycles of 95ºC for 1min, 52ºC for 1 min, and 72ºC for 1 min, with a final extension at 72ºC for 10 min. PCR employed exTaq (Takara, Otsu, Shiga, Japan) in a Biometra thermocycler (Biometra, Goettingen, Germany). Annealing temperatures and extension times were adjusted based on the melting temperatures of the primers and size of the expected amplicon, respectively. For sequencing, PCR products were purified using the QIAquick Gel Extraction Kit (QIAGEN) and sequenced by Davis Sequencing (Davis, CA).

**Southern hybridization.** To prepare a membrane for Southern hybridization, we loaded on the agarose gel genomic DNA digested with EcoRI (New England BioLabs, Ipswich, MA) in parallel with 4 µl of Digioxigenin (DIG)-labeled DNA molecular weight marker II (Roche, Florence, SC). After gel electrophoresis, the gel was submerged in 250 mM HCl (Fisher Scientific) for 10 min with gentle shaking and rinsed with tap water. Then, the DNA on the gel was denatured in denaturation solution (0.5M NaOH and 1.5M NaCl) for 30 min and neutralized in neutralization solution (1M Tris and 1.5M NaCl) for 30 min. In the meantime, the nylon membrane (GE Water &Process Technologies, Trevose, PA) was equilibrated serially in distilled water and 2XSSC. The DNA on the gel was transferred to the

nylon membrane through the movement of 10XSSC. A DNA probe was prepared by DIG-labeling the gel-purified PCR amplicon. This probe was hybridized with the genomic DNA immobilized on the membrane as described in Chapter 4.

**Construction of a lipoprotein gene deletion mutant.** An in-frame deletion mutant of the lipoprotein gene *LMOf2365_0329* was constructed via the integration-excision procedure described by Yildirim et al. (29). Specifically, two PCR amplicons were generated from the genomic DNA of F2365 that correspond to the flanking regions upstream and downstream to the lipoprotein gene. The upstream fragment was amplified with primers LPUF and LPUR that contain EcoRI and BamHI sites, respectively, near their 5' termini (Table 2). The downstream region was synthesized via PCR with primers LPDF2 and LPDU whose 5' ends include BamHI and HindIII sites, respectively (Table 2). These PCR amplicons were digested with EcoRI and BamHI (New England BioLabs) for the upstream fragment and with BamHI and HindIII (New England BioLabs) for the downstream fragment. The shuttle vector pCON-1 (2) was extracted with the QIAprep Spin Miniprep Kit (QIAGEN) from overnight cultures grown in LB broth containing 100 µl/ml ampicillin (Fisher Scientific). The purified plasmid was digested with EcoRI and HindIII and the phosphoryl groups at the ends of the cleaved plasmid were eliminated with Shrimp Alkaline Phosphatase (Promega, Madison, WI). After each DNA manipulation step, the PCR amplicons and plasmid were purified with QIAquick PCR purification kit (QIAGEN). Subsequently, the PCR amplicons and pCON-1 were ligated with T4 ligase (Promega) and electroporated into competent cells of *E. coli* SM10 with an electroporator (Electroporator 2510; Eppendorf, Hauppauge, NY). Transformants were selected on LB plates supplemented

with 100 µl/ml of ampicillin and the presence of the recombinant plasmid pCON-1 was confirmed with colony PCR that employed LPUF/LPDU.

The transformant *E. coli* SL5 was used as the donor of the recombinant plasmid in the conjugation with *L. monocytogenes* F2365. For this experiment, 1.5 ml of the overnight culture of F2365 was incubated at 45ºC for 10 min and mixed with 3 ml of *E. coli* SL5 overnight culture grown in LB with 100 µl/ml of ampicillin. This mixture was filtered onto a sterile 0.45µm HA-type filter of 22 mm diameter (Millipore, Billerica, MA). The filter was washed with 10 ml of BHI, transferred onto the BHI agar plate, and incubated overnight at 30ºC at which pCON-1 is capable of replicating in *L. monocytogenes*. The cells on the filter were resuspended in 3ml of BHI and 100 µl of this culture was inoculated into 30 ml BHI supplemented with 20 µl/ml nalixidic acid (Fisher Scientific) and 6 µl/ml chloramphenicol (ACROS Organics, NJ). Nalixidic acid was added into media to select only for chloramphenicol-resistant *L. monocytogenes* since *E. coli* is susceptible to nalixidic acid but *L. monocytogenes* is naturally resistant to this antibiotic. After overnight incubation at 30ºC, 50 µl of the culture was spread on the BHI agar plate containing nalixidic acid (20 µl/ml) and chloramphenicol (6 µl/ml). Seven transformants were selected for colony PCR that employed LPUF/LPDU and hlyF/hlyR to check the presence of the recombinant DNA and *L. monocytogenes*-specific *hly* gene, respectively. Two confirmed transconjugants (SL5-L15 and SL5-L16) were subcultured on BHI agar plates supplemented with nalixidic acid (20 µl/ml) and chloramphenicol (6 µl/ml) and used further for mutant construction.

To induce the recombinant plasmid to incorporate into the chromosome, SL5-L15 and SL5-L16 were grown overnight at 30ºC in BHI containing chloramphenicol (6 µl/ml) and

went through four consecutive passages in 5 ml of BHI supplemented with chloramphenicol (6 µl/ml) at 42˚C, the temperature at which pCON-1 cannot replicate in *L. monocytogenes*. After the 4$^{th}$ passage, the culture was spread onto the BHI agar plate with chloramphenicol (6 µl/ml) and grown overnight at 42ºC. Two putative integrants (SL5-L15-7 and SL5-L16-5) were selected for the excision procedure. Excision was facilitated by transferring the putative integrants over four consecutive passages in BHI at 30ºC. Potential plasmid replication was prevented by incubating these integrants overnight in BHI at 42ºC after the last passage.

The culture grown at 42ºC was spread on BHI agar plate and incubated at 37ºC. Multiple single colonies were streaked on BHI agar plates with and without chloramphenicol (6 µl/ml). Following overnight growth at 37ºC, 21 colonies sensitive to chloramphenicol were selected and examined with colony PCR employing LPUF/LPDU and the internal primers of the target gene (LPF/LPR) (Table 2). The deletion mutant (SL5-L16-5-5) was confirmed both with Southern hybridization with the DNA probe amplified with primers LPF/LPR and via sequencing with primers LPseqF/LPseqR (Table 2).

**Construction of a complementation mutant.** Complementation of the lipoprotein gene was conducted with pPL2 as described by Lauer et al. (13). For the complementation mutant, PCR was conducted with primers LP_Com_F and LP_Com_R containing SalI and NotI sites, respectively, at their 5' terminal ends (Table 2) to amplify the DNA fragment that encompasses the lipoprotein gene and its upstream and downstream intergenic regions. This PCR amplicon and pPL2 were sequentially restricted by NotI (NEB) and SalI (Promega). Then, the DNA fragment was cloned into pPL2 generating the recombinant plasmid pPL2C. This plasmid was transformed into *E. coli* SM10 and selected on LB agar plates

supplemented with 25 µl/ml chloramphenicol. The presence of the DNA insert was examined with colony PCR that employed primers LP_ComF and LP_ComR (Table 2). Consequently, one transformant (Com1) was chosen for further genetic manipulation. The recombinant plasmid pPL2C was transferred from Com1 to the deletion mutant SL5-L16-5-5 via conjugation and the transconjugants were selected on BHI agar plates with 25 µg/ml nalixidic acid and 6.25 µg/ml chloramphenicol. Incorporation of the recombinant plasmid was ensured with colony PCR using primers NC16 and PL95 (Table 2), and one complementation mutant (Com1S1) was used for further study. In the meantime, plasmid pPL2 was transformed into SM10 and conjugated into the deletion mutant in the identical manner, resulting in PPL2S1, the mutant harboring an integrated empty vector.

**Swarming test.** Swarming was assessed by stabbing a single colony from the overnight culture, or by spotting 5 µl of an overnight liquid culture or the resuspension of a single colony onto BHI plates containing 0.25% or 0.4% agar. Motility was also examined by stabbing the overnight liquid culture in tubes containing BHI soft agar (0.25 or 0.4%). Soft gar plates or tubes were incubated at 25ºC and 37ºC for at least 48 hr and photographed at various time points. When agar plates were used, the plates were sealed with plastic wrap during the incubation to prevent dehydration.

**Cell morphology and hemolysis.** Cell morphology was determined by observing the stationary phage cultures under a phase-contrast microscope (Leica Microsystems, Bannockburn, IL). Cultures were growth in various combinations of liquid growth media and incubation temperatures including BHI at room temperature (RT), 30ºC and 37ºC; in tryptic soy broth (TSB; Becton, Dickinson and Company) at RT and 37ºC; and in TSB containing

0.7% yeast extract (Becton, Dickinson and Company) (TSBYE) at RT and 37ºC. We also used cell suspensions of single colonies grown on BHI agar plates and tryptic soy agar (TSA) plates with 5% sheep blood (Remel, Lenexa, KS) at 4ºC and single colonies on tryptic soy agar with 0.7% yeast extract (TSAYE) at RT and 37ºC. Hemolysis was checked by streaking cells on blood agar plates and incubating them overnight at 37ºC.

**Growth on agar plates.** Cell growth was visually inspected after streaking cells onto the agar plates and incubating them under various conditions. To observe growth on the BHI agar plates at different temperatures, cells were streaked onto the BHI agar plates and incubated at 4, 25, 37 and 42ºC. Growth under the microaerobic condition was examined after incubating cells overnight on BHI agar plates in the reduced oxygen environment generated by Oxoid Gas Generating Kit (Oxoid Ltd, Basingstoke, Hampshire, England). Also, growth in the presence of copper was determined by incubating cells overnight on TSAYE supplemented with 4µM cupric sulfate.

Alternatively, growth on the agar plates was examined by serially diluting the overnight culture by 2, 4, and 8-fold and spotting 5µl of the cultures onto the agar plates. This method was applied to assess growth at different temperatures (4, 25, 37 and 42ºC) on BHI agar plates supplemented with ethanol (PHARMCO-AAPER, Brookfield, CT) (3, 4, 5, 6 and 7%), NaCl (Fisher Scientific) (4, 5, 6, 7 and 8%) and $H_2O_2$ (0.003, 0.006, 0.009, 0.012, and 0.015%). Plates with 0.030, 0.060, 0.090 and 0.120% $H_2O_2$ were incubated at 25, 37 and 42ºC.

**Growth curves and post-stationary phase survival.** Cells were grown at 37ºC until stationary phage (18-36 hr) and 500 µl of this culture was added into 50 ml of BHI, TSB or

BHI supplemented with 7% NaCl (1:100 dilution). Then, the culture was incubated at different temperatures (37 ºC in TSB, 4ºC, 10ºC, RT and 37ºC in BHI; 37ºC; and 37ºC and 10ºC in BHI with NaCl) and optical density at wavelength of 600 nm ($OD_{600}$) was monitored with a spectrophotometer (SmartSpec 3000; Bio-Rad, Hercules, CA). In some experiments (growth in BHI at 4, 10 and 37ºC and in BHI with 7% NaCl at 10 and 37ºC), viable cell counting was conducted after stationary phase.

**Disk diffusion assays.** Overnight cultures were diluted by 100-fold (ca. $1x10^7$ cfu/ml) and evenly spread on BHI agar plates with cotton swaps as described by Kallipolitis et al. (11). Sterile 6 mm diameter paper disks (Becton, Dickinson and Company) were placed in the middle of the plate and 10 µl of the following reagents were spotted onto the paper disks: benzalkonium chloride (BC; ACROS organics) (1 mg/ml), 1-hexadecylpyridinium chloride monohydrate (CPC; ACROS organics) (0.005%), $H_2O_2$ (0.5%), polymyxin B (10 units/µl), nisin (generously provided by Dr. T. Klaenhammer) (10 mg/ml), and SDS (Fisher Scientific) (10%). The plates were incubated 48 hr at 37, 25, 42ºC and for 6 days at 10ºC, and zones of inhibition were measured.

**Minimal inhibitory concentrations (MIC).** The MICs were determined as described by Ryan et al. (21). The overnight cultures diluted by $10^4$-fold (ca. $1x10^5$ cfu/ml) was mixed with 2-fold serial dilutions in a 96-well microtiter plate (Corning Incorporated, Lowell, MA) of BC(0.01 mg/ml), CPC (0.0005%), $H_2O_2$ (0.5%), nisin (0.1 mg/ml), polymyxin B (10 units/µl), and SDS (1%). The microtiter plates were incubated at 25, 37 and 42ºC for 48 hr and at 10ºC for 6-7 days. Growth of the cells was visually inspected at 24 and 48 hr during the incubation at 25, 37 and 42ºC and at 6-7 days during the incubation at 10ºC. MIC was

defined as the minimal concentration at which no visible growth was observed. MIC determinations were conducted three times in duplicate.

**Lethality curves.** To assess the death rates following acid stress, an overnight culture (1 ml) was washed with BHI and the cell pellet was mixed with 1 ml of BHI acidified to pH 2.4 with HCl. The resuspended culture was incubated at RT or 37ºC in duplicate for 2 hours and CFUs/ml were compared between prior to and following acid treatment. To assess death rate following exposure to nisin and SDS, 1 ml of the overnight culture was centrifuged at 13000 rpm for 2 min and the supernatant was removed. The cell pellet was mixed with BHI supplemented with 1% SDS or 0.01 mg/ml of nisin, and incubated at RT. At 5, 20 and 40 min after the addition of nisin or SDS, 25 µl of the culture was serially diluted in phosphate buffered saline (PBS; EMD Chemicals, Gibbstown, NJ) solution and spread onto BHI agar plates with a spiral-plater (Autoplate 4000; Spiral Biotech, Inc., Bethesda, MD). Colonies were enumerated with Q-count (Q Count™; Spiral Biotech, Inc., Norwood, MA) after overnight incubation at 37ºC. The ratio of cfu/ml at each time point was compared with the cfu/ml of the overnight culture.

**Phage infection assays.** The deletion mutant SL5-L16-5-5 and parental strain F2365 were grown in BHI overnight at 25, 37 and 42ºC; for 7 days at 8ºC; and for approximately 27days at 4ºC. These cultures were infected with broad-range *Listeria* phage 20422-1 as described in the previous chapter. The efficiency of plaquing (EOP) was defined as the ratio of the plaque forming units (pfu)/ml for the deletion mutant to the pfu/ml for the parental strain. Phage infection experiments were conducted at least three times for each incubation temperature.

**Statistical analysis.** Survival and phage infection assay data were analyzed with the statistical program SAS (SAS Institute Inc., Cary, NC) employing mixed analysis of variance (ANOVA) model.

## RESULTS

***In silico* analysis.** The putative lipoprotein gene that was characterized in this study (*LMOf2365_0329*) was located downstream of the Sau3AI-like RM system in *L. monocytogenes* F2365 (ECI), and was transcribed in the same direction as the DNA binding protein and methyltransferase genes in this RM system. Upstream and downstream of the putative lipoprotein gene in F2365, we identified a gene encoding a protein of unknown function (*LMOf2365_0328*) and a gene encoding threonine aldolase (*LMOf2365_0330*), respectively, which are transcribed in the same direction to the lipoprotein gene (Fig. 1A). One putative Rho-independent terminator was predicted between the DNA binding protein gene and betwee*LMOf2365_0328*, and another was anticipated between the lipoprotein and threonine aldolase genes, suggesting that the putative lipoprotein was in a different transcriptional unit from the RM cassette genes upstream, or from the downstream threonine aldolase (Fig. 1A). Two promoters were in the intergenic region upstream of the lipoprotein gene (Fig. 1B). Any attempt to identify protein motifs or conserved domains of this putative lipoprotein either resulted in no hits (NCBI Conserved Domain Search) or revealed as top hits protein families of unknown function (Pfam search and KEGG database) other than lipoprotein families.

**Deletion and complementation of the lipoprotein gene.** To identify the functions of the putative lipoprotein gene, an in-frame deletion of this gene was constructed via shuttle vector and conjugation. The deletion of the target gene was supported by the absence of the signal in the Southern blot with a probe based on the lipoprotein gene (Fig. 2). Also, the sequencing results of the region around the target gene in the deletion mutant (Fig. 1C) revealed that the 555 bp of the lipoprotein gene was mutated into the 45 bp gene and that no change in sequence occurred except for the newly created BamHI site in the middle of the mutated gene.

The mutant was genetically complemented by cloning the intact gene (with its putative promoter region) in the integration vector pPL2 (Fig. 3A) and conjugating this recombinant plasmid into the deletion mutant SL5-L16-5-5. Since pPL2 possessed the integrase gene and *attPP'* sites originated from the phage PSA (13), this recombinant plasmid was integrated into the PSA attachment site within $tRNA^{Arg}$ in this strain. This was confirmed by PCR with primers NC16/PL95 (Fig. 3B) that is expected to generate a 499-bp amplicon from strains that harbor a lysogenic phage or integration vector at this site (13). The presence of the lipoprotein gene in the complemented mutant was confirmed with PCR employing the primers internal to the target gene (Fig. 3C). To control for potential effects of the plasmid integration, we also constructed an integration mutant with pPL2 integrated in $tRNA^{Arg}$ of the deletion mutant SL5-L16-5-5 (Fig. 3B and 3C).

**Swarming.** Although considerable variation was observed between individual experiments and no distinct difference in motility was observed under microscope between the deletion mutant and the parental strain grown both in agar and broth media, the deletion

mutant often exhibited a moderate reduction in swarming at 25ºC on the BHI soft agar plates in comparison with its parental strain (Fig. 4A). However, swarming was not restored in the genetically complemented mutant Com1S1, for which swarming was similar or larger to that observed for the mutant harboring the empty vector (Fig. 4A). In tubes with BHI soft agar, no distinct difference was observed (Fig. 4C and 4D).

**Morphology, hemolysis and growth on agar plates and in broth.** No difference was observed in cell morphology between the deletion mutant and its parental strain following growth in various media and at different temperatures (Fig. 5). Also, the deletion mutant and its parental strain exhibited comparable hemolysis on blood agar plates (Fig. 6 A). Growth was similar between the deletion mutant and its parental strain when the bacteria were grown on BHI agar plates at different temperatures (4, 25, 37 and 42ºC) aerobically (Fig. 6B) and at 37ºC under the microaerobic condition (Fig. 6C) or in the presence of copper (Fig. 6D). When spotted onto agar plates in the presence of ethanol, NaCl or $H_2O_2$ and grown at different temperatures, no difference in growth was identified between the deletion mutant and the wild type (Fig. 7A-C).

No difference in growth in liquid cultures was observed between the deletion mutant and its parental strain when the cells were grown in BHI at 37, RT, 10 and 4ºC (Fig. 8Ai, Bi, Ci and Di, respectively)), in TSB at 37ºC (Fig. 8E) and in BHI containing 7% NaCl at 10 and 37ºC (Fig. 9Ai and Bi). Post-stationary phase survival was also compared between the deletion mutant and its parental strain, and no distinct difference was identified (Fig. 8Aii, Bii, Cii and Dii, and Fig. 9Aii and Bii), in agreement with the growth curve data.

**Susceptibility to toxic compounds.** The deletion mutant and parental strain were subjected to the disk diffusion assays employing paper disks containing various toxic compounds. After incubation at 37ºC, noticeable halos were observed in both strains around the disks with BC, $H_2O_2$, nisin, and SDS (Fig. 10A, C, D and E, respectively)) whereas the disk diffusion assays employing polymyxin B often resulted in no distinct halos (Fig. 10D). Around the disks with CPC (Fig. 10(B)), a no-growth zone with irregular shape was observed. No significant differences in inhibition zones were noted between mutant and parental strain with any of the toxic agents spotted onto the disks (Fig. 10). Similar results were observed at different temperatures (10, 25 and 42ºC). Meanwhile, in the MIC determination, the deletion mutant revealed2-fold lower MIC for nisin at 10 and 37ºC (24 hr) and 2-fold higher MIC for polymyxin B at 42ºC (48 hr) compared with the parental strain (Table 3).

Assessments of survival following acid treatment (pH 2.4) revealed no significant difference between the deletion mutant and the parental strain either at 37ºC or at RT (p value=0.5541 and 0.3827, respectively) (Fig. 11A). Similar results were obtained following treatment with 0.01 mg/ml of nisin (p value=0.3778, 0.2646 and 0.0719 for 5, 20 and 40 min exposures, respectively) (Fig. 11B). However, when exposed to a lethal level of SDS (1%) for 20 and 40 min, the deletion mutant exhibited a smaller reduction in viability compared with the parental strain (p=0.0004 and 0.0364 for 20 and 40 min exposures, respectively) (Fig. 11C). The difference in survivability between the deletion mutant and the parental strain amounted to $10^{0.55}$ (3.5-fold) for 20 min exposure and $10^{0.32}$ (2.1-fold) for 40 min exposure.

**Phage infectivity.** The deletion mutant and parental strain grown at different temperatures (4, 8, 25, 37 and 42ºC) were both highly susceptible to *Listeria* phage 20422-1. No significant differences in EOP were detected between the deletion mutant and parental strain regardless of the growth temperatures (p=0.3020, 0.5825, 0.7049, 0.4365 and 0.8603 for cultures grown at 4, 8, 25, 37 and 42ºC, respectively) (Fig. 12).

## DISCUSSION

In this study, we aimed to characterize the function of a putative lipoprotein gene for which our previous investigation indicated over-representation among clinical and environmental/food isolates but under-representation among animal isolates (Chapter 3). This gene did not seem to belong to essential genes since some isolates were missing this gene and a deletion mutant was readily constructed. However, the prevalence data for this gene led us to speculation that it might play an important part in the life cycles of *L. monocytogenes* in the environments and/or in the infectious process. In the transcriptomic analysis of serotype 1/2a *L. monocytogenes* EGDe by Toledo-Arana et al., the gene homologous to the lipoprotein gene (*lmo0303*) constituted one operon with its upstream gene *lmo0302*, which is missing in the strain studied here (26). Compared with exponential growth at 37ºC in the nutrient medium (BHI), the transcription of *lmo0302* was decreased when EGDe was grown in blood or intestine whereas *lmo0303* exhibited lower transcription in the cells grown in blood, suggesting that the lipoprotein gene might be more likely to be involved in the survival in the environment (26). Unfortunately, repeated attempts to obtain clues with various bioinformatics failed to provide clues for the function of this lipoprotein gene.

Our experiments to study phenotypes attributable to the lipoprotein gene were conducted by comparing the wild type strain and the deletion mutant under conditions expected to be important for environmental survival of *L. monocytogenes.* For instance, growth at low temperatures has been considered as a key trait of *L. monocytogenes* that facilitates dissemination of this microorganism through food distribution system (27). Following the same rationale, our experiment conditions also included salt (27), nisin (10, 18), and compounds used as disinfectants such as CPC, SDS (21), BC (8, 20, 21), ethanol (11) and $H_2O_2$ (20). We included polymyxin B in our experiments since susceptibility to polymyxin B was observed in a mutant deficient in diacylglyceryl transferase gene (*lgt*) that is involved in the lipoprotein maturation (15). The deletion mutant was also compared with its parental strain in regard to other bacteriological attributes including motility, cell morphology, hemolysis, growth under microaerobic conditions, phage infectivity, and resistance to hydrochloric acid and copper.

In most of these experimental conditions, the phenotype of the lipoprotein deletion mutant was comparable to that of the parental strain. This is consistent with the findings by Reglier-Poupet et al. on the functions of the mutant missing the signal peptidase gene (*lsp*), another gene involved in the lipoprotein maturation process (19). They reported that the deletion mutant was identical to the wild-type strain with regards to cell morphology, hemolysis, and viability in BHI at different temperatures (4, 30, 37 and 42ºC) (19). However, unlike this previous study that reported no effect by *lsp* on motility, our deletion mutant exhibited a moderately reduced motility at 25ºC compared with the parental strain. A previous study by Lemon et al. suggested that flagellar motility might be imperative in

attaching to the surface and establishing biofilms (14). Thus, a speculation could be drawn that the lipoprotein gene might also contribute to biofilm formation, thus facilitating inhabitation of *L. monocytogenes* in food-processing facilities. Also, in comparison with the ECI parental strain, the deletion mutant revealed two-fold decrease in MIC for nisin when cells were grown at 10 and 37ºC, suggesting that the lipoprotein gene might be involved in resistance to nisin at these temperatures. Given that nisin effectively inhibits *L. monocytogenes* growth and has been approved as a food additive (10, 18), this role in nisin resistance might result in the widespread distribution of the lipoprotein gene among *L. monocytogenes* strains of clinical and environment/food origins. Meanwhile, the deletion mutant revealed two-fold increase in MIC for polymyxin B and a significantly higher survival following exposure to lethal level of SDS in comparison with the parental strain. These phenotypes suggest that the lipoprotein gene might render a cell more vulnerable to polymyxin B and SDS. Such apparent disadvantages do not seem to be consistent with the epidemiological data for this gene; however, these characteristics might be a relatively small price that *L. monocytogenes* should pay in maintaining this particular gene in exchange for the advantages that this gene confers with regards to the aforementioned resistance to nisin and motility as well as under some, unidentified conditions.

It should be noted that the lipoprotein deletion mutant exhibited no or limited phenotypic differences from the parental strain under the experimental conditions used in this study. To advance our understanding of the function of this gene, experimental conditions should be broadened to include additional conditions, for instance, biofilm formation (14) and in minimal medium (1). Also, the scope of investigation could be furthered into

pathogenicity. Studies on the lipoprotein maturation process reported that the deletion of the genes involved in this process (*lst* and *lgt)* resulted in decreased pathogenicity. The *lst* deletion mutant exhibited reduced phagosomal escape (19) and the *lgt* deletion mutant revealed impaired invasion and intracellular survival (15). Both mutations resulted in virulence attenuation in mice (15, 19). These findings suggest that many lipoproteins, presumably including our target lipoprotein, might contribute to pathogenicity.

In this study, the swarming defect was not restored in the complemented mutant. This suggests that reduced motility might not be due to the targeted lipoprotein gene or that the gene was not adequately expressed in the complemented. The complemented mutant c could be verified by assessing transcription of the lipoprotein gene. Alternatively, a new complementation mutant could be constructed with an independent promoter upstream to the lipoprotein gene.

In summary, our data revealed that the lipoprotein gene *LMOf2365_0329* harbored by a high proportion of clinical and food/environmental isolates of *L. monocytogenes* might be involved in motility and nisin resistance but might increase susceptibility to polymyxin B and SDS. Although these findings provided clues on why this gene was maintained in many *L. monocytogenes* strains including major epidemic-associated clonal groups, further studies are warranted to characterize the role of this gene. This functional characterization will further our knowledge on a member of the most abundant group of *Listeria* surface proteins and will be instrumental in drawing inferences on the functions of its divergent homologs that were predominantly presented among animal isolates.

**Table 5.1 Bacterial strains used in this study**

| Species and strains | Genotype and features | Reference |
|---|---|---|
| <*E. coli*> | | |
| SM10 | Conjugation donor; $F^-$ *thi-1 thr-1 leuB6 recA tonA21 lacY1 supE44* ($Mu_C^+$) $\lambda^-$ [RP4-2(Tc::Mu)] $Km^r$ $Tra^+$ | (13) |
| *E. coli* harboring pCON-1 | Shuttle vector, pCON-1/ *E. coli* | (2) |
| *E. coli* containing pPL2 | Integration vector, pPL2/ *E. coli* | (13) |
| <*L. monocytogenes*> | | |
| F2365 | Wild type | (17) |
| SL5-L16-5-5 | F2365Δ*LMOf2365_0329* | This study |
| PPL2S1 | F2365Δ*LMOf2365_0329*, $tRNA^{Arg}$::pPL2 | This study |
| Com1S1 | F2365Δ*LMOf2365_0329*, $tRNA^{Arg}$::*LMOf2365_0329* integration vector, pPL2C | This study |

**Table 5.2 Primers employed in this study**

| Name | Sequence (5' to 3') | Target | Reference |
|---|---|---|---|
| LPUF | GTAGAATTCGTAGAAGCCATAGAAGCAAG<br>EcoRI | *LMOf2365_0328* | This study |
| LPUR | TTAGGATCCGAAAATAATACTTACTAAC<br>BamHI | *LMOf2365_0329* | This study |
| LPDF2 | GGCGGATCCAAAAATTAAATTATTAGAAGC<br>BamHI | *LMOf2365_0329* | This study |
| LPDU | GTCAAGCTTTTCCAAGTAATATGGGCGTG<br>HindIII | *LMOf2365_0330* | This study |
| LP_Com_F | TTTGTCGACCTGTCTTTGTCCAACATGTC<br>SalI | *LMOf2365_0328* | This study |
| LP_Com_R | TAAGCGGCCGCCCTAGAACTCAAGCCTATTG<br>NotI | Intergenic region between *LMOf2365_0329* and *LMOf2365_0330* | This study |
| LPseqF | TGATGCTGAAGAAATAATG | *LMOf2365-0328* | This study |
| LPseqR | AATACTCCGAAATCTCCTC | *LMOf2365_0330* | This study |
| LPF | GTTGTTGATAGTAAGTATAC | *LMOf2365_0329* | This study |
| LPR | TCTGCTTTGACATCTATGTG | *LMOf2365_0329* | This study |
| NC16 | GTCAAAACATACGCTCTTATC | PSA *int* in pPL2 | (13) |
| PL95 | ACATAATCAGTCCAAAGTAGATGC | Upstream to $tRNA^{Arg}$ | (13) |

**Table 5.3 MIC of the wild type strain (F2365) and deletion mutant (SL5-L16-5-5) for toxic agents**

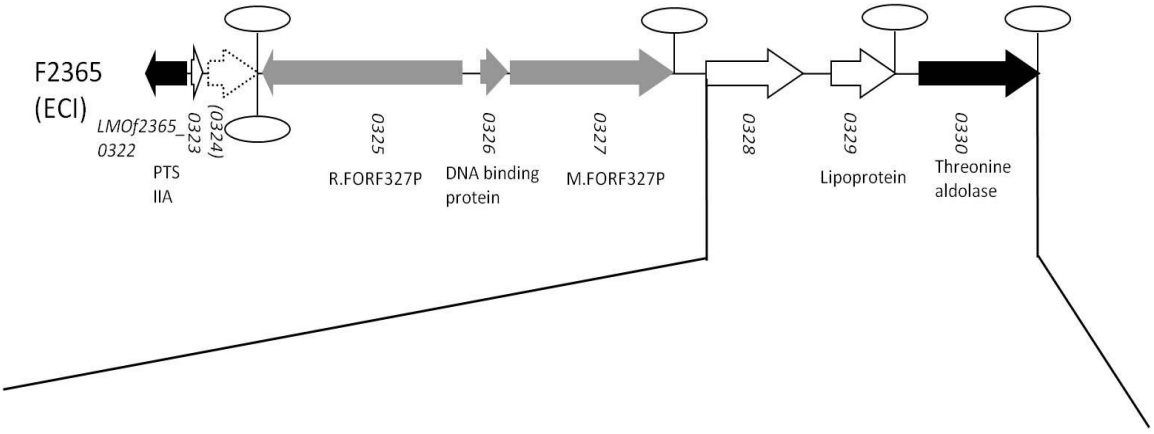
| Stress | 10ºC | | 25ºC | | 37ºC | | 42ºC | |
|---|---|---|---|---|---|---|---|---|
| | F2354 | SL5-L16-5-5 | F2354 | SL5-L16-5-5 | F2354 | SL5-L16-5-5 | F2354 | SL5-L16-5-5 |
| BC (µg/µl) | $1.25*10^{-3}$[a] | $1.25*10^{-3}$ | $1.25*10^{-3}$ / $1.25*10^{-3}$[b] | $1.25*10^{-3}$ / $1.25*10^{-3}$ | $1.25*10^{-3}$ / $1.25*10^{-3}$ | $1.25*10^{-3}$ / $1.25*10^{-3}$ | $6.25*10^{-4}$ / $1.25*10^{-3}$ | $6.25*10^{-4}$ / $1.25*10^{-3}$ |
| CPC (%) | $6.25*10^{-5}$ | $6.25*10^{-5}$ | $6.25*10^{-5}$ / $6.25*10^{-5}$ | $6.25*10^{-5}$ / $6.25*10^{-5}$ | $6.25*10^{-5}$ / $1.25*10^{-4}$ | $6.25*10^{-5}$ / $1.25*10^{-4}$ | $3.13*10^{-5}$ or $6.25*10^{-5}$ / $6.25*10^{-5}$ | $3.13*10^{-5}$ or $6.25*10^{-5}$ / $1.25*10^{-4}$, $3.13*10^{-5}$ or $6.25*10^{-5}$ |
| $H_2O_2$ (%) | $7.81*10^{-3}$ | $7.81*10^{-3}$ | $3.91*10^{-3}$/ $7.81*10^{-3}$ | $3.91*10^{-3}$/ $7.81*10^{-3}$ | $7.81*10^{-3}$/ $1.56*10^{-2}$ | $7.81*10^{-3}$/ $1.56*10^{-2}$ | $7.81*10^{-3}$/ $7.81*10^{-3}$ | $7.81*10^{-3}$/ $7.81*10^{-3}$ |
| Nisin (µg/µl) | $6.25*10^{-3}$ | $3.13*10^{-3}$ | $3.13*10^{-3}$/ $3.13*10^{-3}$ or $6.25*10^{-3}$ | $3.13*10^{-3}$/ $3.13*10^{-3}$ or $6.25*10^{-3}$ | $6.25*10^{-3}$/ $6.25*10^{-3}$ | $3.13*10^{-3}$/ $3.13*10^{-3}$ or $6.25*10^{-3}$ | $1.56*10^{-3}$ or $3.13*10^{-3}$/ $3.13*10^{-3}$ | $1.56*10^{-3}$/ $3.13*10^{-3}$ |
| Polymyxin B (units/µl) | $3.13*10^{-1}$ | $3.13*10^{-1}$ | $3.13*10^{-1}$/ $6.25*10^{-1}$ | $3.13*10^{-1}$/ $6.25*10^{-1}$ | $3.13*10^{-1}$/ $6.25*10^{-1}$ | $3.13*10^{-1}$/ $6.25*10^{-1}$ | $3.91*10^{-2}$ / $7.81*10^{-2}$ | $3.91*10^{-2}$/ $1.56*10^{-1}$ |
| SDS (%) | $1.56*10^{-2}$ | $1.56*10^{-2}$ | $3.13*10^{-2}$/ $6.25*10^{-2}$ | $3.13*10^{-2}$/ $6.25*10^{-2}$ | $1.56*10^{-2}$/ $3.13*10^{-2}$ or $6.25*10^{-2}$ | $1.56*10^{-2}$/ $3.13*10^{-2}$ | $1.56*10^{-2}$/ $1.56*10^{-2}$ | $1.56*10^{-2}$/ $1.56*10^{-2}$ |

[a] The MIC represents the mode (the most frequently identified MIC level) of three independent experiments conducted in duplicate. When several levels were observed at the same frequency, all of these were presented in the table.

[b] At temperatures 25, 37 and 42ºC, the first MIC was obtained after 24 hr incubation whereas the second MIC after the slash mark signifies the MIC after 48 hr incubation.

**Figure 5.1 Genomic region harboring the putative lipoprotein gene in *L. monocytogenes.*** (A) Genetic region that harbors the ECI Sau3AI-like RM system and the putative lipoprotein gene. RM system genes and conserved flanking genes were colored in grey and black, respectively whereas other genes were marked in white. Stem-loop symbols represent putative Rho-dependent terminators. (B) DNA sequence of the region encompassing the lipoprotein gene (*LMOf2365_0329*, marked in red) and its flanking genes (*LMOf2365_0328* and *LMOf2365_0330*, colored in blue). Start codons were marked in grey and stop codons were surrounded by box. Sequences on which primers were designed were underlined. Predicted promoters are italicized and highlighted in turquoise blue. Putative transcription start sites, -10 and -35 sites were colored in pink. The potential Rho-dependent terminator was boldfaced and the reverse complementary sequences within the terminator were underlined. (C) Sequencing result of the deletion mutant with the primer LPseqF. The BamHI site that was generated due to genetic manipulation was marked in bold and underlined.

**A.**




**B.**

*LMOf2365_0328*

TTGTATAATTTTAGAGTTATTTCTGCTAATGAGGACATCTTTATAAAAACCAAATATAGATTTATCTCAG

AGTTAGAAGATTTTAAAGAATTTGCACAAACTCTAACTGATGATGCTGGAAAAGCAAAAGTTAAAAGTGG

AAAGGCTAGTTCATATACTAGATACCTTATACGACTAATAATTTTTTATTGTGAAACTCACAGAGACAAT

CTAGAGGAATTATCATCATTTGAAGCAGTGAAAAAACTTGAAAAAATAAGATTAGACGAAAGTTTTAAAG

CATTCAATCAGGAATCTAACCGTTTTTATAGCGCGACTTTAAGCTGTTATTTGGCATATGTAACATTTAA

LPseqF

AAATTCTGATGCTGAAGAAATAATGGATAACCAAATTATTAGTACTTTGAATAATTTTGAAAATCCAGCA

CAAGTAGTAAAGGAAAATAAATCAACTTATTTAGTTAGGGTGCCTAAAAGTAGAGCTAGTAAAAGAAAAA

LPU

ATAATAATGTTTATTCATATCCAAGAAACTATGTAGAAGCCATAGAAGCAAGAAGAAGATGTGGATGGGT

ATGTGAATTTAATCAAGAGCATAAGACTTTTATTAACATTAAGGATGGTAAGCCTTATGTGGAAGGACAT

LP_Com_F

CATCTTATACCTATGGCAGTACAGGATTATTATGATAATACAATTGATTTTGCAGATAATATAATCTGTC

TTTGTCCAACATGTCATAGAAAAATTCATTATGCTGTTCAAAGTGAGAAAAAAGAAATGTTAATAAAAAT

TTTTAATGAAAGGGAAAAATTTTATTTACAACATGGTATAGAGATTGATAGAAAAACACTTTTAAGTTTT

-35 *Predicted promoter 1* -10 +1

TATGGTATTTTTTAACTCTATTGTGAAAATTTCGAAAAATCTTAAATAATAAATTGTAAAATTAATTGTC

*Predicted promoter 2* -10

ATGGTTAGTGAAAAAAACAACAGAATTAGACTAGTTTTGGACACATTTTTTATAGGTGATTAAGTTCTAG

+1

GATGAGTATTTTGAAGAGAGAACGCAAGCAATGGAGTACATCGAGGTGGAACGGGATTAACCGTTAATTA

*LMOf2365_0329*

LPU

TATACAAAAAAAGGAGAGCAACAAGATGAGCAAAAAGTTAGTAAGTATTATTTTCTTAGGGATTTGTCTT

LPF

TTGTTAGTTGGTTGTGGTAAAGAGGTTGTTGATAGTAAGTATACCGGAAAATGGAAGTTGGAATCGTCGG

GGATTGGTATGAGCGATGGATCAAATTACAAGGAAACGGATGCAGCGGGTAAAGTTAGTCTTGACTTGGA

CATTGATGCAGAGGGAAATGTAAAAGAATTGTATATTGATGATGCAGTTGCTATAACTAATTCTTATAAA

TTAAAGAAAAAAGGTGACGATGAATATCAATATGATGGTACATTGATTAGCAAAAAAGTATACAGTTATG

AAACTGAAAGTGATAAAGAGAATGTAGAAACAATGCTAAAATTTCTTAAATCCAAGAATAATGTCAAAAT

AACCAAAAATGAACAAAAAGGGGATGAATATCACATAGATGTCAAAGCAGATGAAAAAGATTTCGTATTC LPR

TCTTTAGAAATGCAAAGTGGAAAATTAATTTTCAAAAAAGTCGATACTAAAGAGAAAGTTTTGATTAAGG

AAACCTACAAGAAAAATTAAATTATTAGAAGCGAGAAGCGGAAGTCCGTTGTAACCACCTATTTTTTAAA LPDF2 -35

*Predicted promoter* -10 +1 **Predicted terminator**

***ATTTATTTGTGAAGACTATTCTTTCGAGGGTCCTCGG*****GAGGATAGTTTTTTTGTTACCGT**AGTTAAATTT

ACAATAGGCTTGAGTTCTAGGTATAATATAGATAAATTTAGAAATTATATCGTAAAAGGAGTGAACCAAA LP_Com_R

*LMOf2365_0330*

ATGATAAACACACTAAAAACCAGCTACCAAAAAACACCATATAAACTAGGCGGGAACGGTCCGCGCAATG

TAGGTGTTCTGACAGAAGCACTTCAAAATATAGATGATAACCTAGAAAGCGATATTTACGGAAATGGCGC

AGTTATTGAAAACTTTGAAACAAAAATCGCGAAAATTCTCGGGAAACAATCTGCGGTATTTTTTCCAAGT

GGGACGATGGCTCAGCAAATTGCGTTAAGGATTTGGGCTGACCGGAAAGAGAATCGGTGCGTGGCATATC

ATCCACTCTCTCATTTGGAGATTCACGAGCAAGATGGGTTAAAAGAATTGCAACAAATCACGCCCATATT LPDU

ACTTGGAACAGCTAATCGACTTTTGACAATTGACGATATTAAAAGCCTTCGCGAGCCAGTTTCTAGTGTG

CTTATCGAACTTCCACAACGAGAAATTGGTGGGCAGCTGCCTGCTTTTGAAGAGCTTGAGGAGATTTCGG LPseqR

AGTATTGCCATGAACAGGGTATTTCGTTGCATCTTGACGGGGCACGGTTATGGGAAATCACGCCGTTTTA

TCAGAAGTCTGCGGAAGAAATTTGTGCGCTGTTTGATAGTGTGTATGTGTCGTTTTACAAAGGAATTGGC

GGGATTGCTGGGGCGATTCTGGCTGGGAATGATGATTTTGTGCAAGAGGCGAAAATCTGGAAACGACGGT

ATGGCGGAGATTTGATTAGTCTGTATCCGTATATTTTGTCCGCGGATTACTATTTTGAAAAACGGATTGG

GAAAATGGCGGAATATTTTGAAGCGGCGAAAGGATTGGCTGAGCGATTTAATTCGTGTTCGGGCGTGAAG

ACTGTGCCGGAAGTGCCGGTTTCTAATATGTTTCATGTTTATTTTGAGAAGTCGGCCGATGAGGTTGGGG

CGATTTTGACGAAGATTCAGGATGAAACTGGAGTTGGAATTTCGGGGTATTTGCAGGAGAAATCGGCGGA

TGTTTGCGCGTTTGAGGTTTCGGTTGGGGATGCATTTGCGGAGATACCGGTGGAGATTTTGGATGATTTA

TTTAAGAAACTTGTAAATTCAGTAAAATAG

**C.**

*LMOf2365_0328*

GTACTTTGATAATTTTGAAATCCAGCACAAGTAGTAAAGGAAAATAAATCAACTTATTTAGTTAGGGTGC

CTAAAAGTAGAGCTAGTAAAAGAAAAAATAATAATGTTTATTCATATCCAAGAAACTATGTAGAAGCCAT

AGAAGCAAGAAGAAGATGTGGATGGGTATGTGAATTTAATCAAGAGCATAAGACTTTTATTAACATTAAG

GATGGTAAGCCTTATGTGGAAGGACATCATCTTATACCTATGGCAGTACAGGATTATTATGATAATACAA

TTGATTTTGCAGATAATATAATCTGTCTTTGTCCAACATGTCATAGAAAAATTCATTATGCTGTTCAAAG

TGAGAAAAAAGAAATGTTAATAAAAATTTTTAATGAAAGGGAAAAATTTTATTTACAACATGGTATAGAG

ATTGATAGAAAAACACTTTTAAGTTTTTATGGTATTTTTTAACTCTATTGTGAAAATTTCGAAAAATCTT

AAATAATAAATTGTAAAATTAATTGTCATGGTTAGTGAAAAAAACAACAGAATTAGACTAGTTTTGGACA

CATTTTTTATAGGTGATTAAGTTCTAGGATGAGTATTTTGAAGAGAGAACGCAAGCAATGGAGTACATCG

*LMOf2365_0329* remnant

AGGTGGAACGGGATTAACCGTTAATTATATACAAAAAAAGGAGAGCAACAAGATGAGCAAAAAGTTAGTA

**BamHI**

AGTATTATTTTC**GGATCC**AAAAATTAAATTATTAGAAGCGAGAAGCGGAAGTCCGTTGTAACCACCTATT

TTTTAAAATTTATTTGTGAAGACTATTCTTTCGAGGGTCCTCGGGAGGATAGTTTTTTTGTTACCGTAGT

TAAATTTACAATAGGCTTGAGTTCTAGGTACAATATAGATAAATTTAGAAATTATATCGTAAAAGGAGTG

*LMOf2365_0330*

AACCAAAATGATAAACACACTAAAAACCAGCTACCAAAAAACACCATATAAACTAGGCGGGAACGGTCCG

CGCAATGTAGTGTTCTGACAGAAGCACTTCAAAATATAGATGATACCTAGAAGCGATATTTACGAAATGG

CGCAGTTATGAAACTTTGAACAAAATCGCGAAATCTCGGGAACAATCTGCGTATTTTTCCAAGTGGACGA

TGCTCAGCAATGCGTAGATTGGGCTGACGAAGAGAATCGGTGCGTGCATATCATCCACTCTCTCATTGGA

GATCACGGAGCCAAG

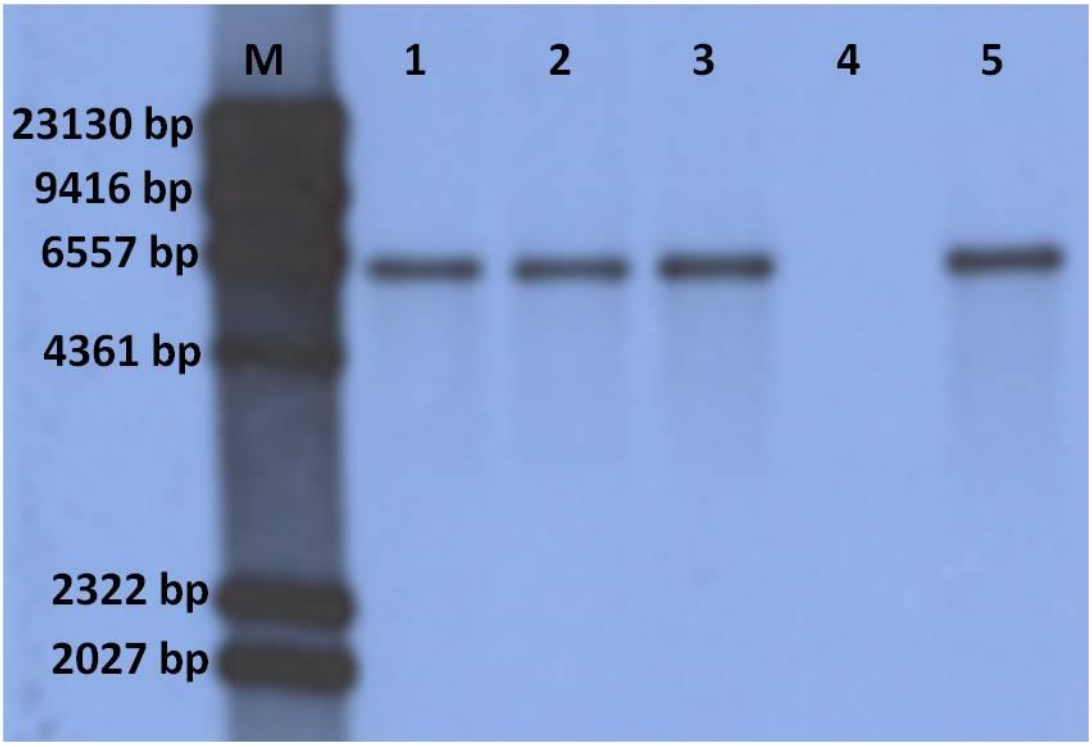
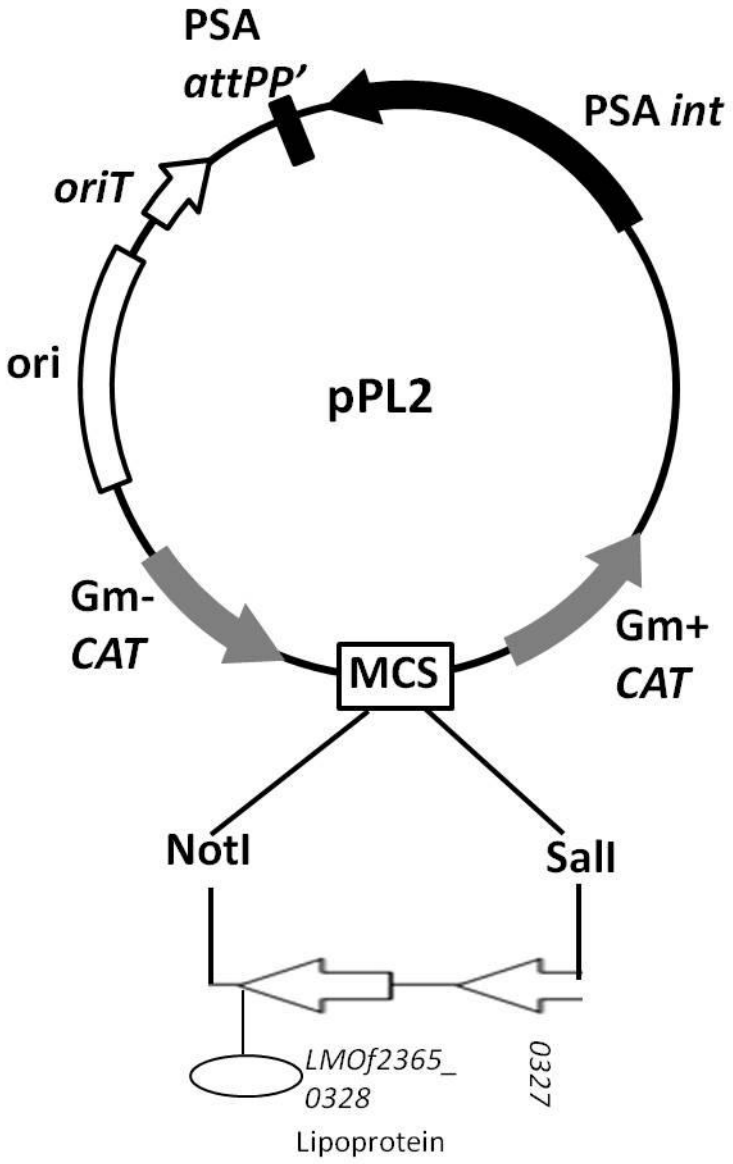



**Figure 5.2 Southern hybridization of the genomic DNA digested with EcoRI with the DNA probe (LP) amplified with primers LPF/LPR based on the lipoprotein gene.** The letter M represents the DIG-labeled DNA marker that was run in parallel with the genomic DNA. The size of the bands from the marker was exhibited next to the corresponding band. Lanes 1 and 3 were loaded with the genomic DNA from three different colonies which maintained the lipoprotein gene after all the deletion mutant construction procedures (SL5-L16-5-1, SL5-L16-5-1-2 and SL5-L16-5-10, respectively). Lanes 4 and 5 correspond to the deletion mutant (SL5-L16-5-5) and the parental strain (F2365).

**Figure 5.3 Construction and confirmation of the complementation mutant.** (A) Schematic diagram of the cloning of the lipoprotein gene into the plasmid pPL2 (12) for complementation mutant construction. (B and C) Confirmation of the complementation mutants and empty vector mutants via PCR using primers NC16/PL95 (B) and LPF/LPR (C). Letter C signifies the complementation mutants in which the pPL2 containing the lipoprotein gene was integrated into the chromosome of the deletion mutant SL5-L16-5-5 whereas letter E represents the empty vector mutants in which the pPL2 was incorporated into the chromosome of the deletion mutant. Letters F, S, N and M represent the parental strain (F2365), deletion mutant (SL5-L16-5-5), no-DNA control, and molecular marker, respectively. In Figure B, lanes 1 to 5 were loaded with PCR amplicons from strains Com1S1, Com1S2, PPL2S1, PPL2S2, and PPL2S4, respectively. In Figure C, lanes 1 to 5 contained PCR products from strains PPL2S1, PPL2S2, PPL2S4, Com1S1 and Com1S2, respectively.

**A.**


PSA
attPP'
PSA int
oriT
ori
pPL2
Gm-
CAT
MCS
Gm+
CAT
NotI
SalI
LMOf2365_
0328
0327
Lipoprotein

**B. NC16/PL95** **C. LPF/LPR**

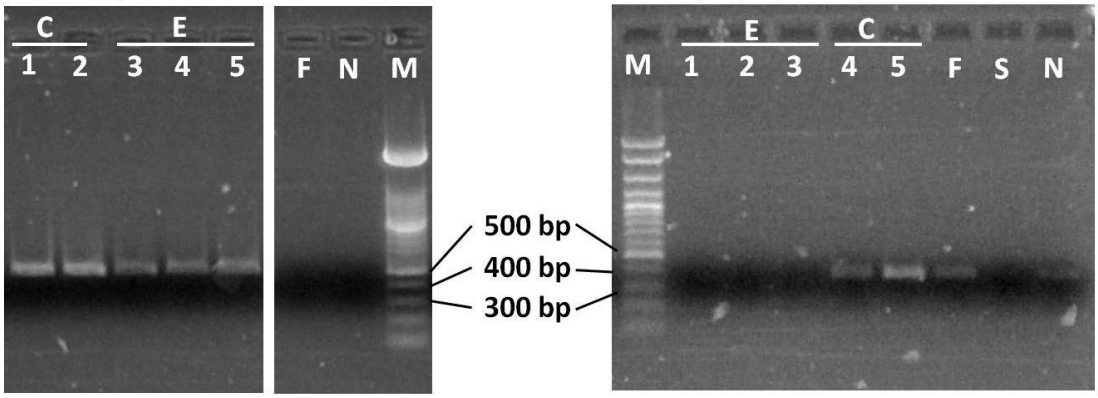
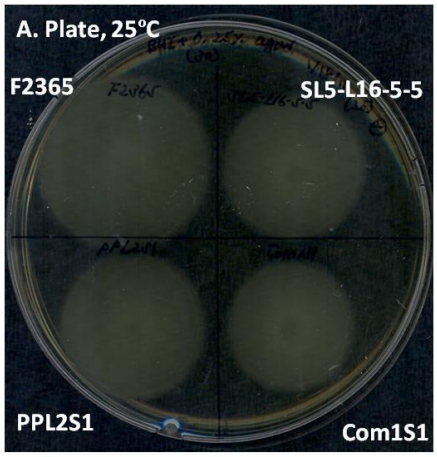
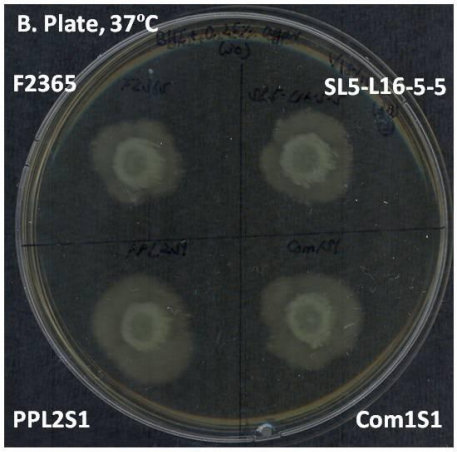
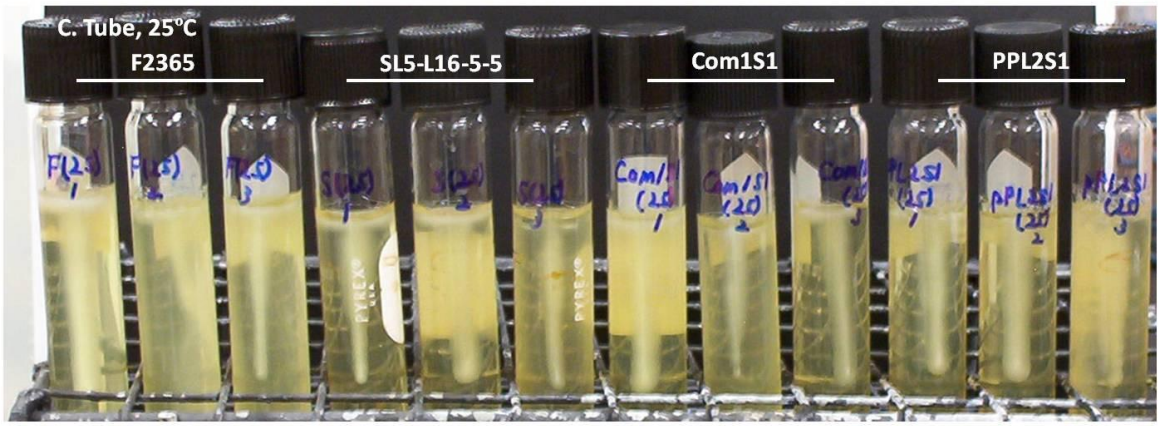
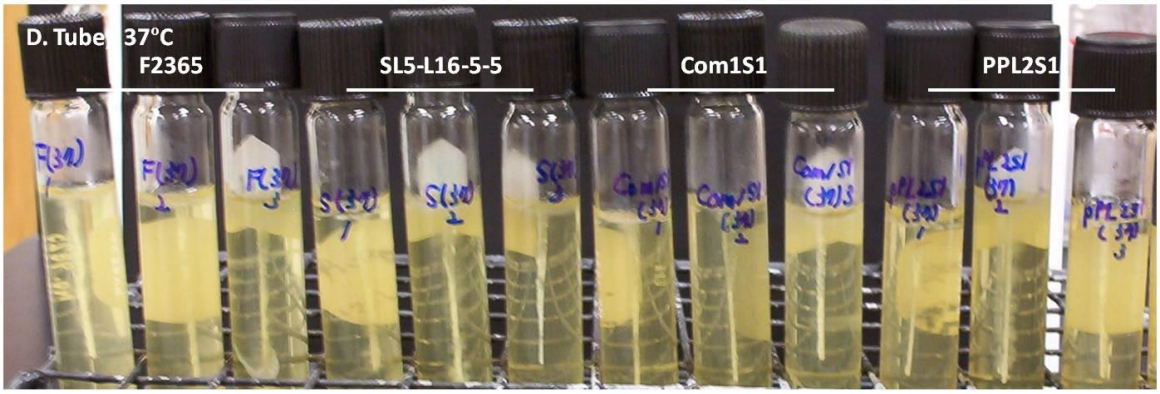


**Figure 5.4 Results of motility tests.** (A and B) The motility tests were conducted by spotting single colony suspensions on the 0.25% agar BHI plates and incubating for 48 hr at 25ºC (A) and 37ºC (B). (C and D) The cells were stabbed into the test tubes filled with 0.4% BHI agar and incubated for 48 hr at 25ºC (C) and 37ºC (C).


A. Plate, 25°C
F2365
SL5-L16-5-5
PPL2S1
Com1S1


B. Plate, 37°C
F2365
SL5-L16-5-5
PPL2S1
Com1S1


C. Tube, 25°C
F2365
SL5-L16-5-5
Com1S1
PPL2S1


D. Tube, 37°C
F2365
SL5-L16-5-5
Com1S1
PPL2S1

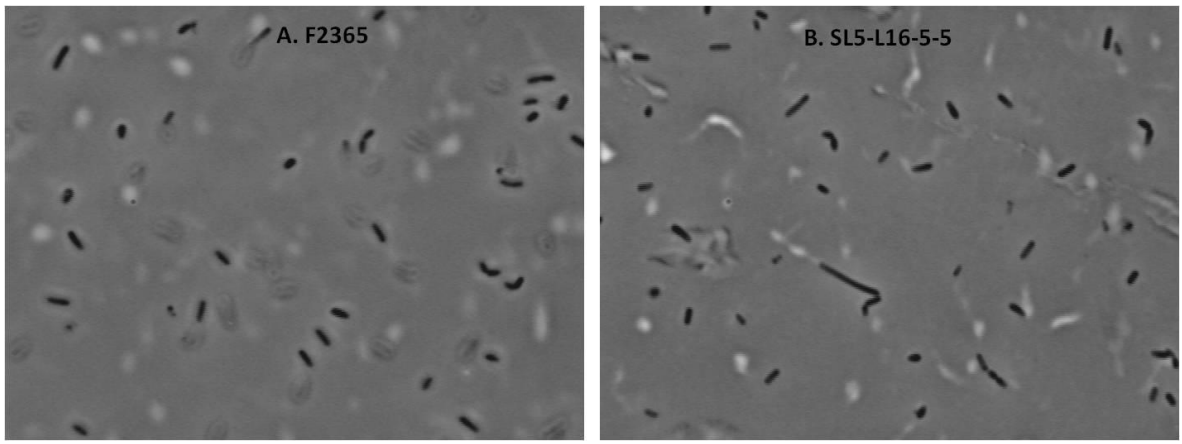



**Figure 5.5 Microscopic observation of the parental strain F2365 (A) and the lipoprotein deletion mutant SL5-L16-5-5 (B) after 24 hr incubation in BHI at 37ºC.** Similar morphology was observed when cells were grown at different temperatures (4ºC, RT, 30ºC) and/or in different media (TSB, TSBYE, TSAYE and blood agar).

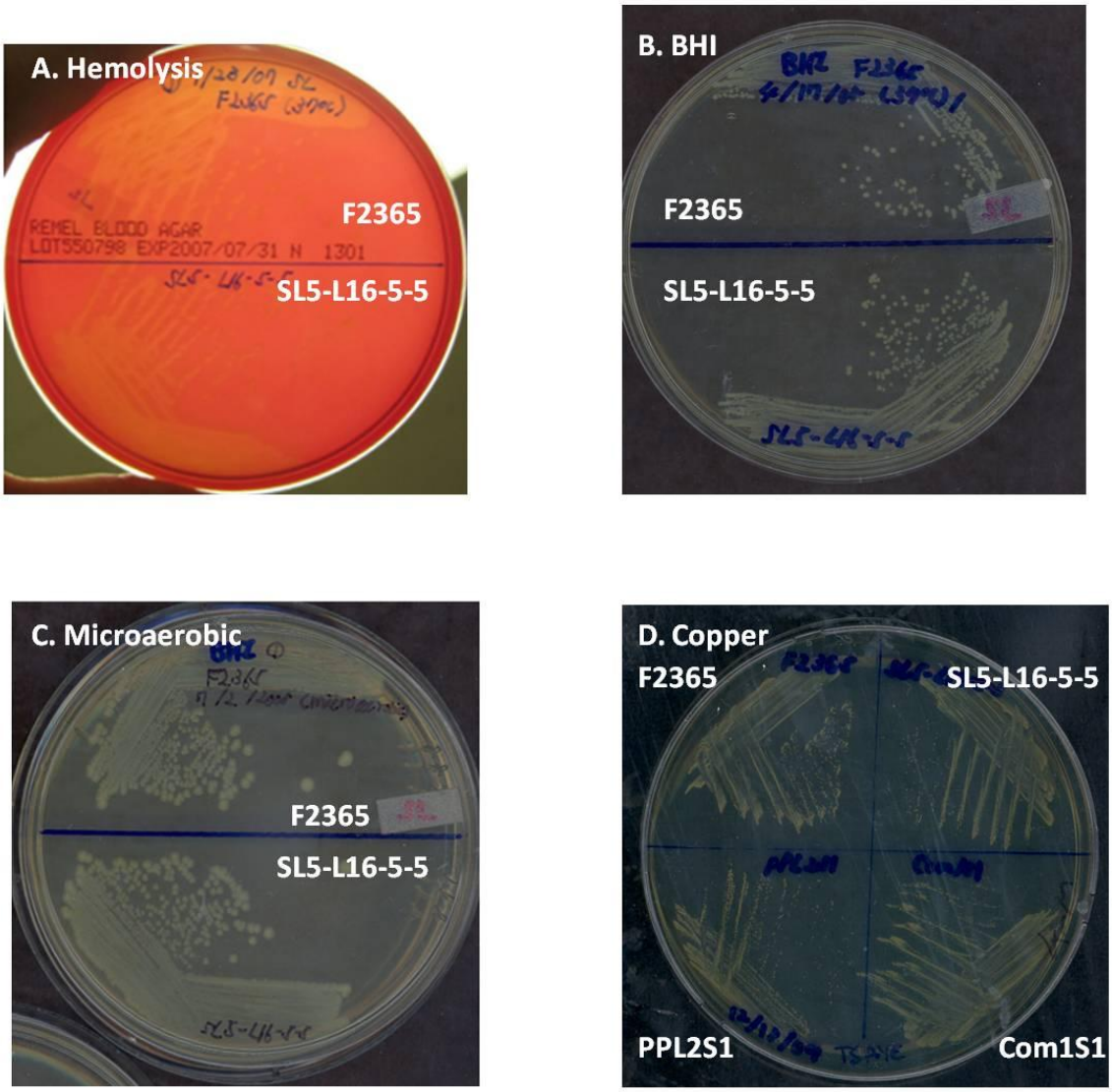



**Figure 5.6 Hemolysis (A) and growth on the agar plates (B-D).** The cells were grown overnight at 37ºC under appropriate conditions: on the blood agar plate (A); on the BHI agar plate (B); on the BHI agar plate in the microaerobic container (C); and on the TSAYE plate supplemented with 4µM of copper (D). On the BHI agar plates, the parental strain and deletion mutant were grown at different temperatures (4, 25 and 42ºC) and exhibited similar results to Figure B.

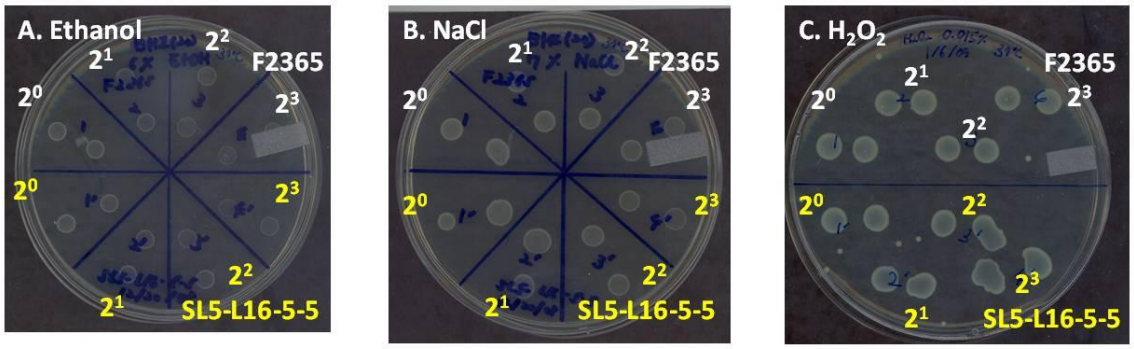
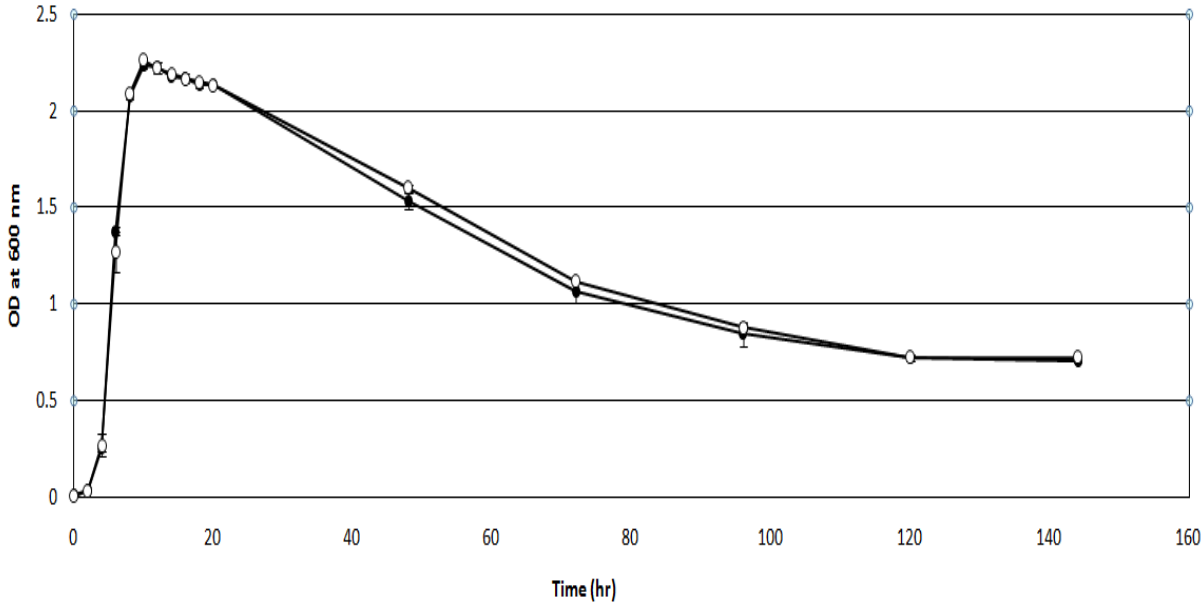



**Figure 5.7 Growth of the spotted cultures on the agar plates.** Cells were grown for 24 hr at 37ºC on the BHI agar plates supplemented with 6% ethanol (A), 7% NaCl (B), and 0.015% $H_2O_2$ (C). The number represents the dilution factor of the overnight culture. The parental strain (F2365) and deletion mutant (SL5-L16-5-5) were marked in white and yellow, respectively. Similar results were observed in the plates containing different concentrations of the stress agents. Similar results were observed at other incubation temperatures (4, 25 and 42ºC).

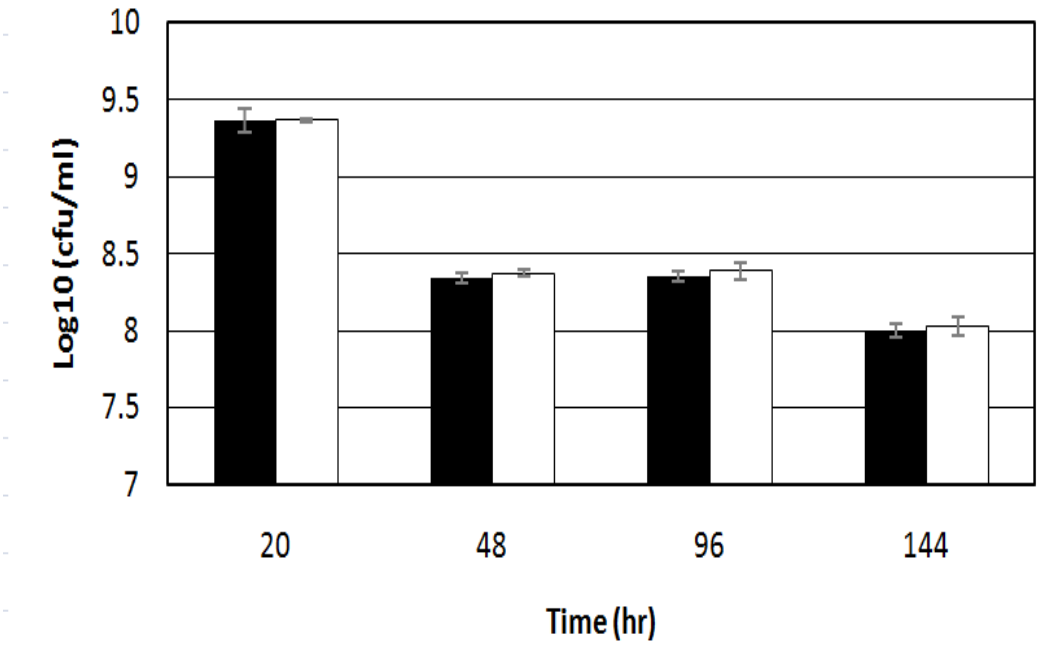
**Figure 5.8 Growth curves and post-stationary survival in BHI and TSB.** In the growth curve, the wild type F2365 and deletion mutant SL5-L16-5-5 were marked with black and open circles, respectively. In the figures on survivability, black bar signifies the wild type whereas white bar corresponds to the deletion mutant. Each individual figure represents growth curve (Ai) and post-stationary phase viability (Aii) in BHI at 37ºC; growth curve in BHI at RT (B); growth curve (Ci) and post-stationary phase survival (Cii) in BHI at 10ºC; growth curve (Di) and survivability after stationary phase (Dii) in BHI at 4ºC; and growth in TSB at 37ºC (E). The data at 37ºC, RT and 10ºC consist of the average obtained from one representative experiment conducted in duplicate at each temperature whereas experiments at 4ºC were performed twice in duplicate. Error bars were determined by the maximum and minimal observations.

**A. BHI, 37°C**

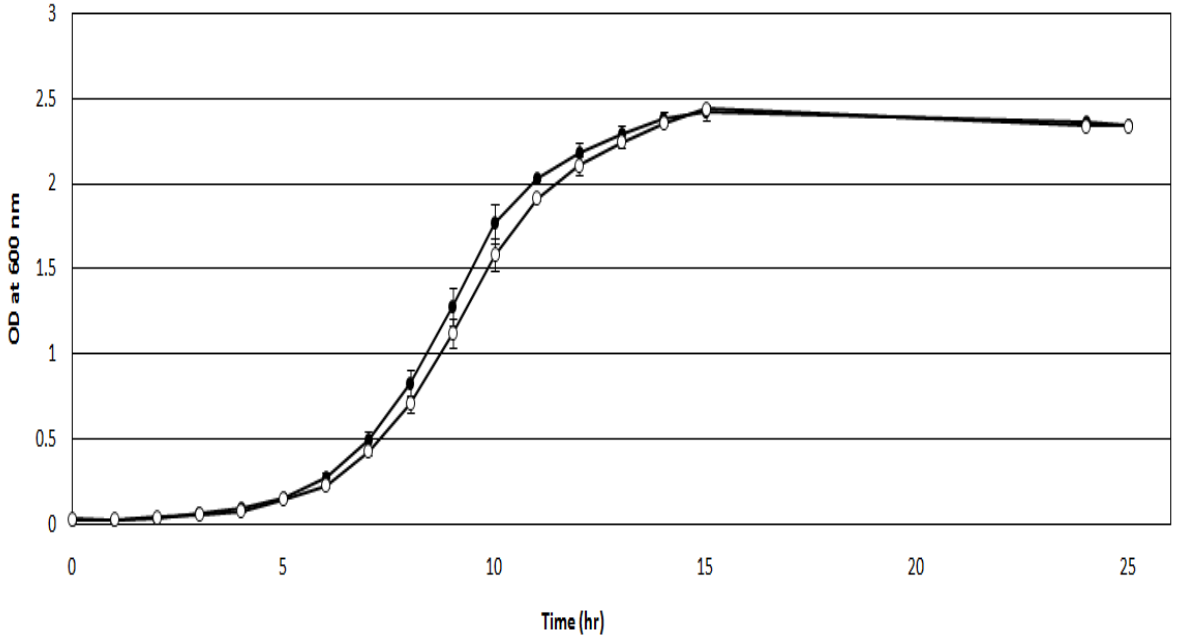
**i**


OD at 600 nm
2.5
2
1.5
1
0.5
0
0
20
40
60
80
100
120
140
160
Time (hr)

**ii**

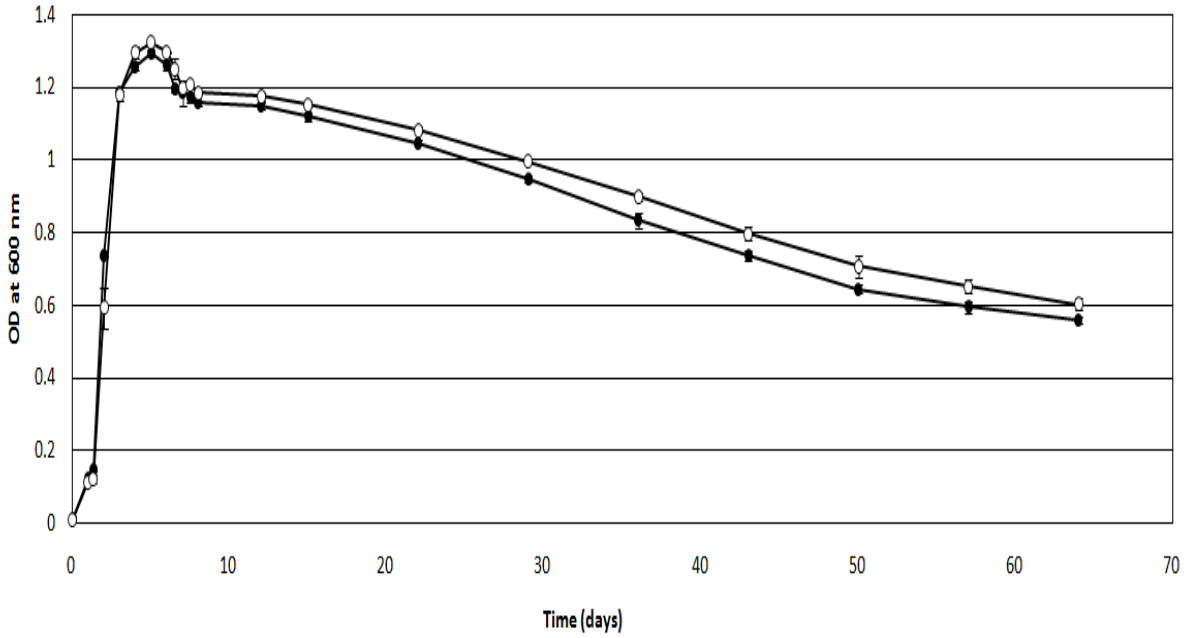

Log10 (cfu/ml)
10
9.5
9
8.5
8
7.5
7
20
48
96
144
Time (hr)

**B. BHI, RT**

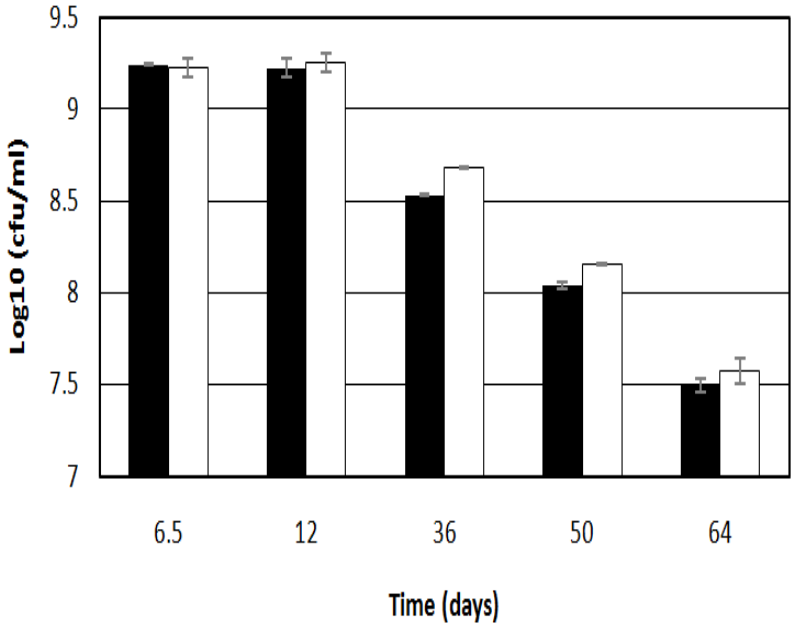


**C. BHI, 10°C**

**i**

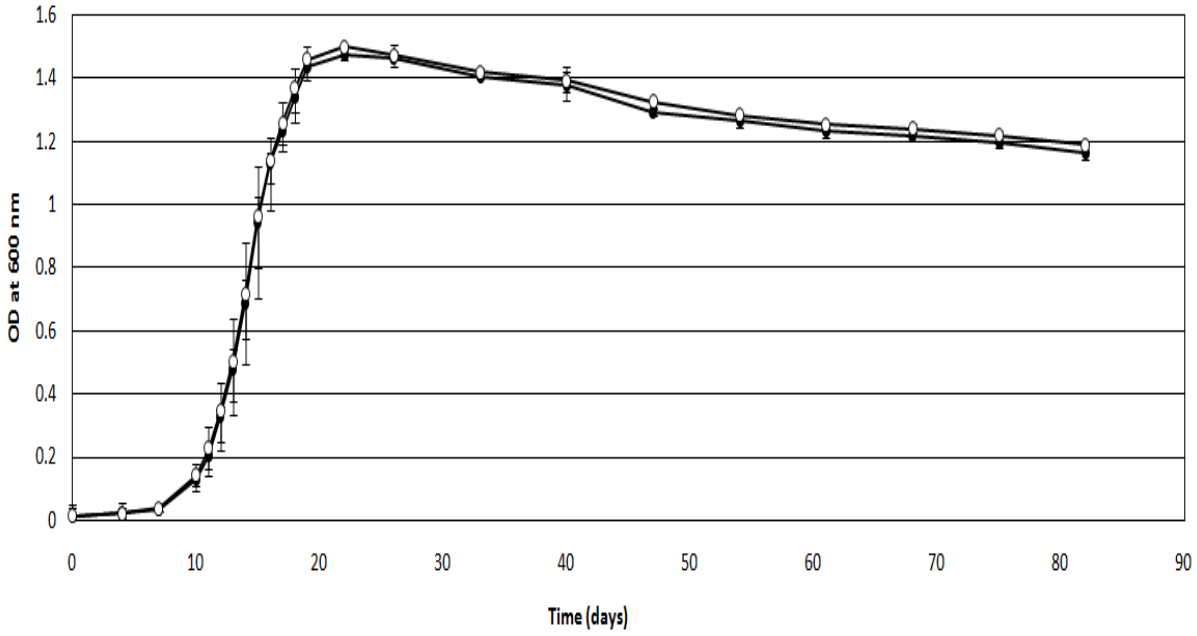

OD at 600 nm
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
10
20
30
40
50
60
70
Time (days)

**ii**

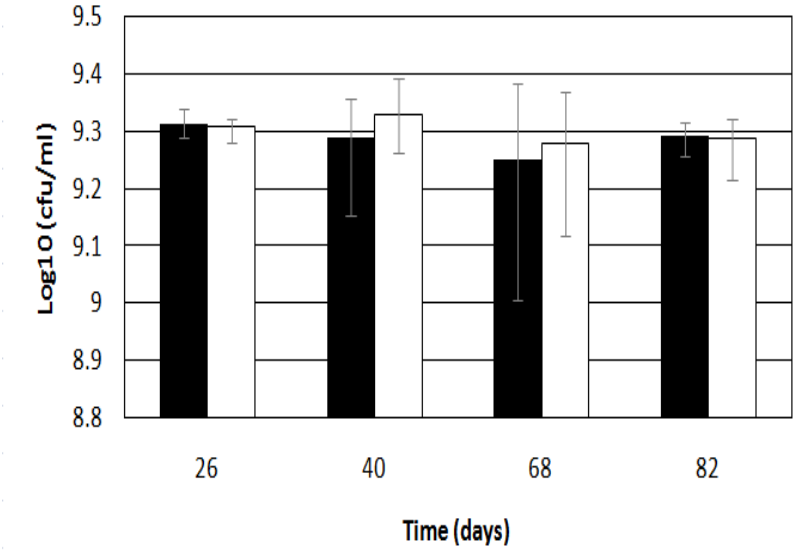

Log10 (cfu/ml)
9.5
9
8.5
8
7.5
7
6.5
12
36
50
64
Time (days)

**D. BHI, 4°C**

**i**


OD at 600 nm
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
10
20
30
40
50
60
70
80
90
Time (days)

**ii**


Log10 (cfu/ml)
9.5
9.4
9.3
9.2
9.1
9
8.9
8.8
26
40
68
82
Time (days)

**E. TSB, 37°C**

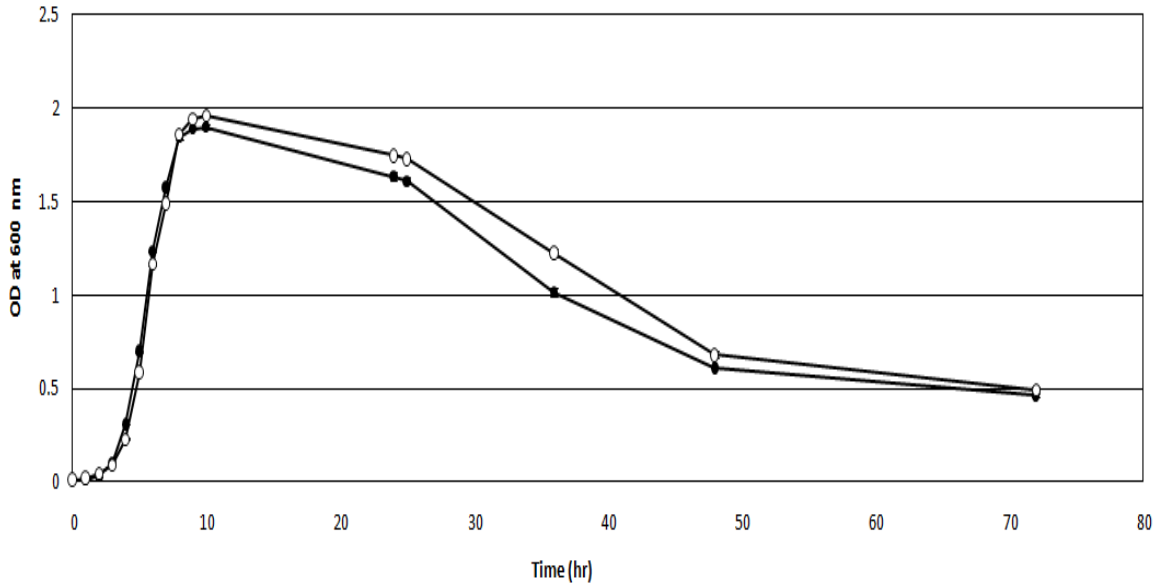
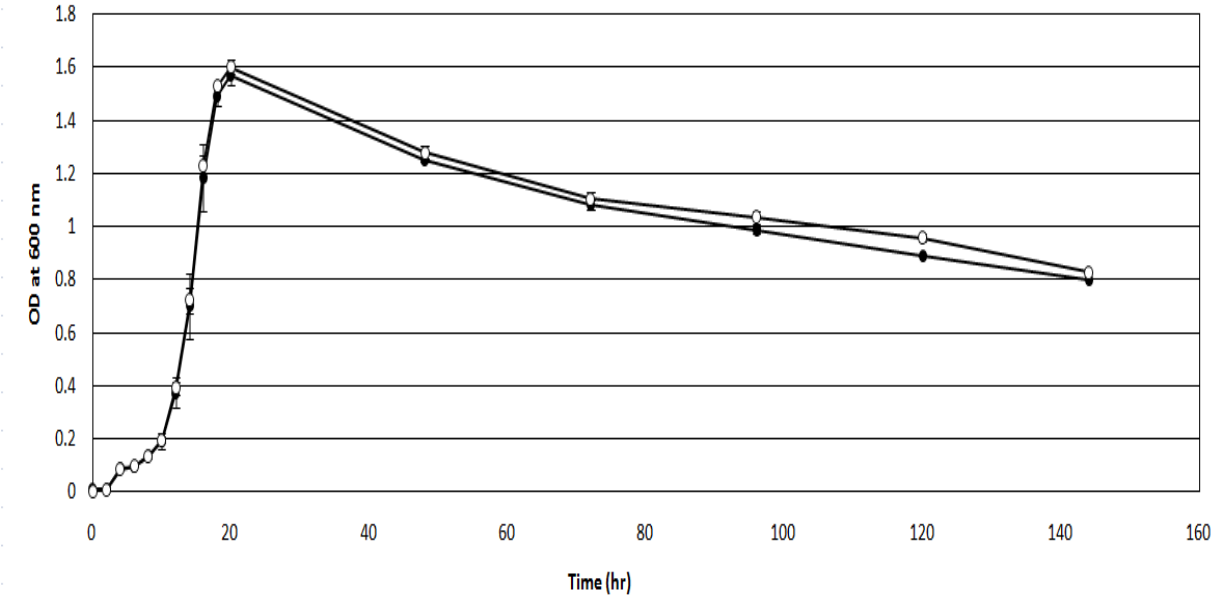


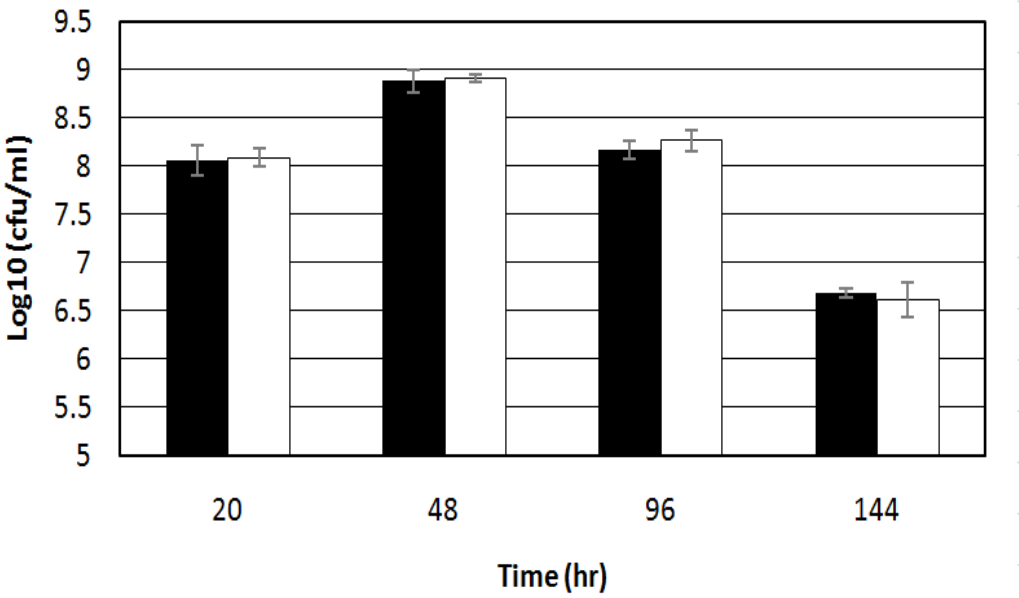
**Figure 5.9 Growth curves and post-stationary phase viability in BHI supplemented with 7% NaCl.** Figures A and B represent growth curve (i) and survivability (ii) at 37 and 10ºC, respectively. The parental strain F2365 and deletion mutant SL5-L16-5-5 were represented by black and open circles in the growth curve, respectively. In the graphs on survivability, black bar symbolizes F2365 and white bar corresponds to the deletion mutant. Each figure is based on the average of one representative experiment conducted in duplicate and error bars were calculated with the maximal and minimal measurements.

**A. BHI+7% NaCl, 37ºC**

**i**

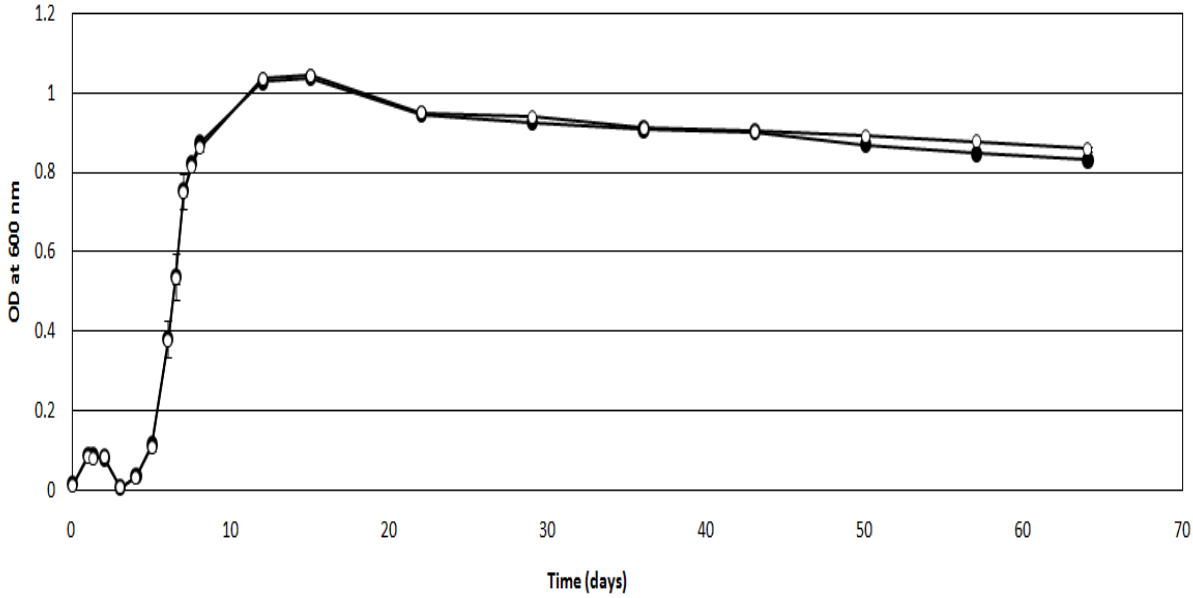



**ii**

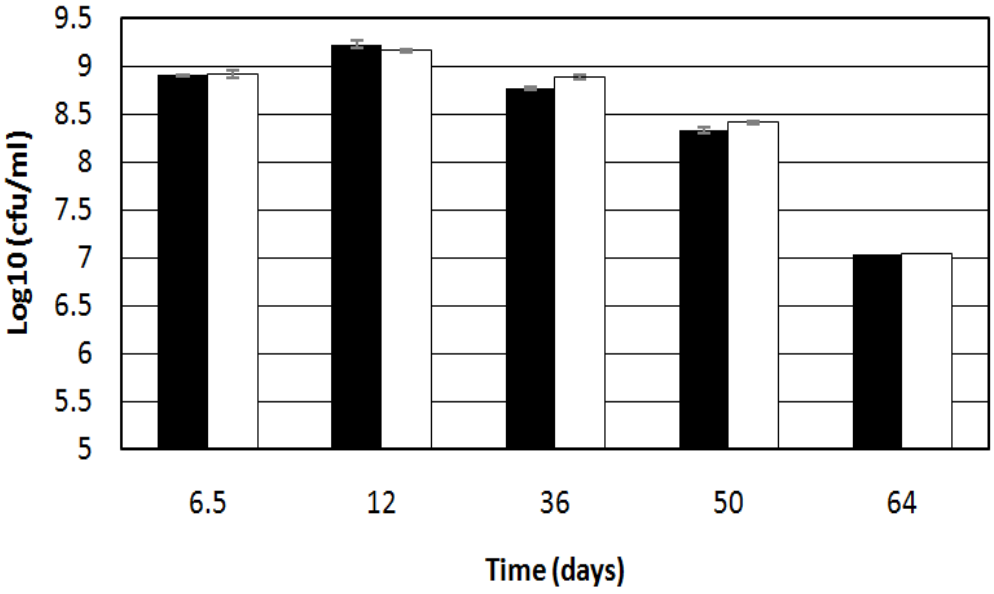


**B. BHI+7% NaCl, 10ºC**

**i**




**ii**



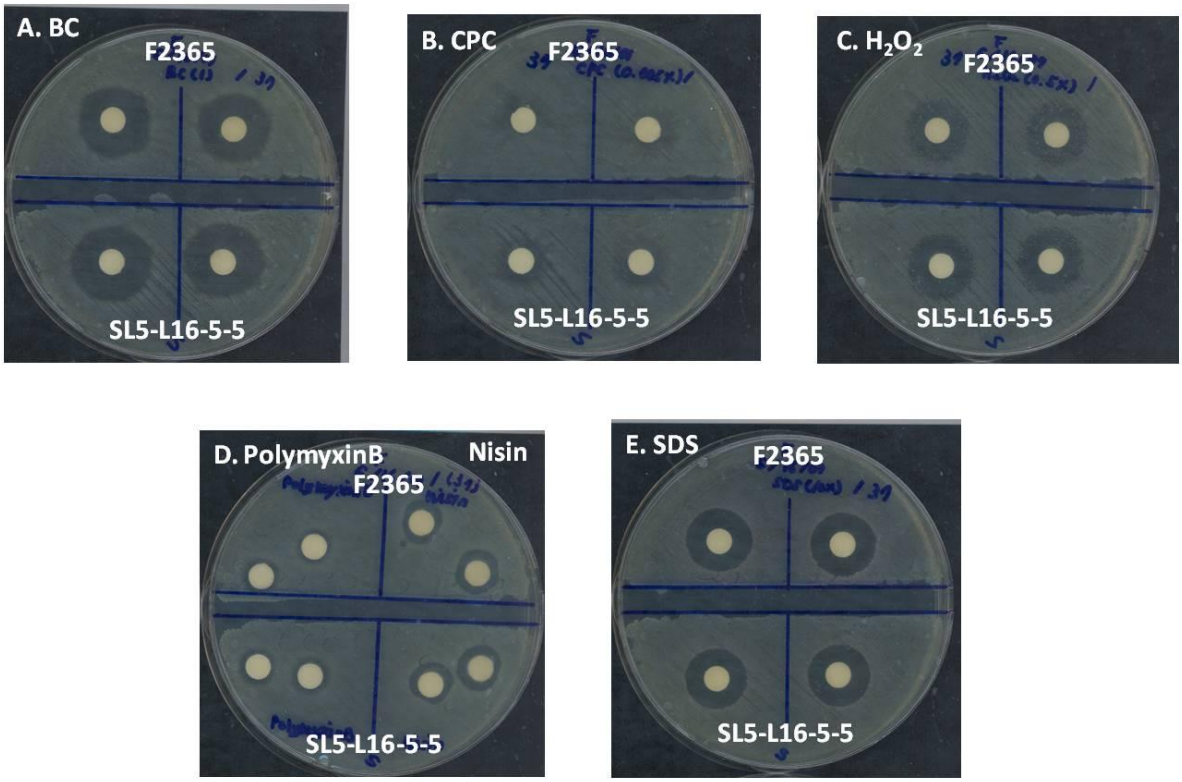
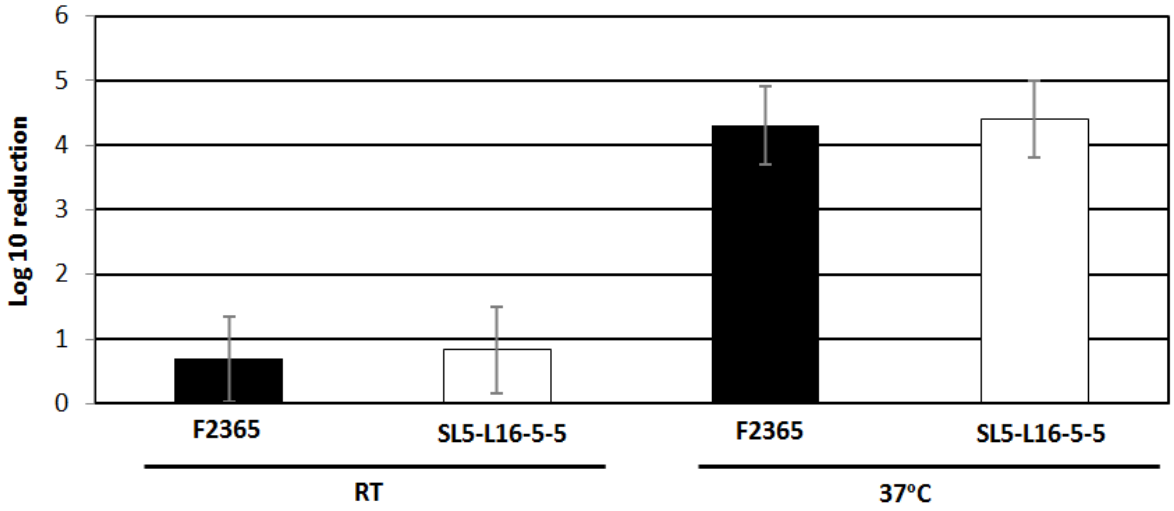


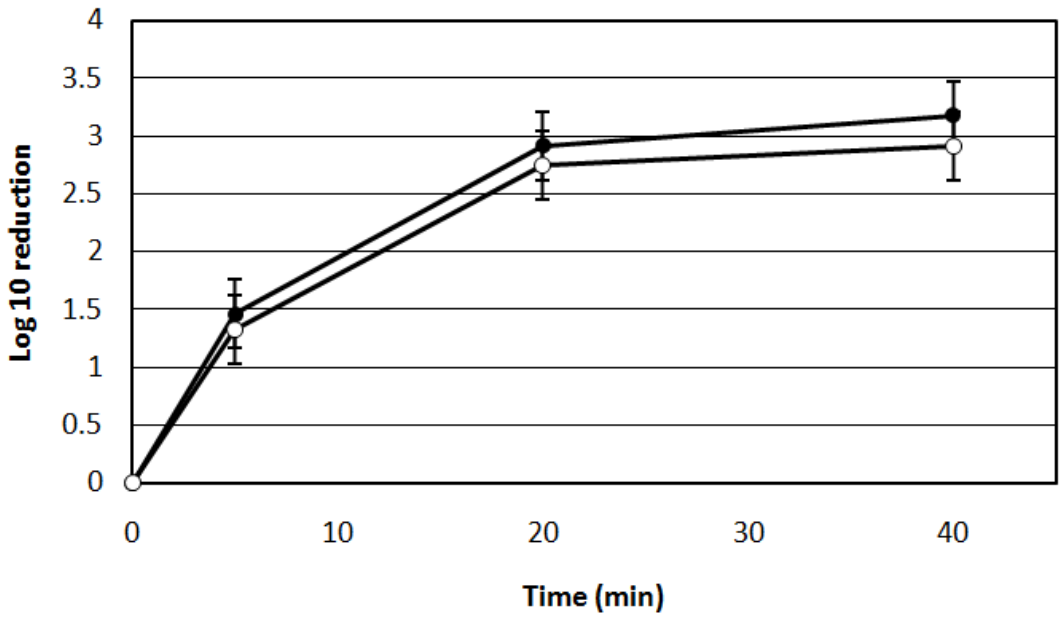

**Figure 5.10 Results of disk diffusion assays after 24 hr incubation at 37ºC.** Similar results were observed at different incubation temperatures (10, 25 and 42ºC).

**Figure 5.11 Killing curves in the lethal level of acid (A), nisin (B) and SDS (C).** Figure A represents the average of at least four independent experiments conducted in duplicate. In this figure, the wild type F2365 and the deletion mutant SL5-L16-5-5 were symbolized with black and white bars. Data of Figures B and C were obtained from four independent experiments performed in duplicate. In Figures B and C, black and open circles signify F2365 and SL5-L16-5-5. In all figures, error bars are based on the standard error calculated from statistical analysis. Asterisks mark the time points that revealed significance difference between the wild type and deletion mutant.

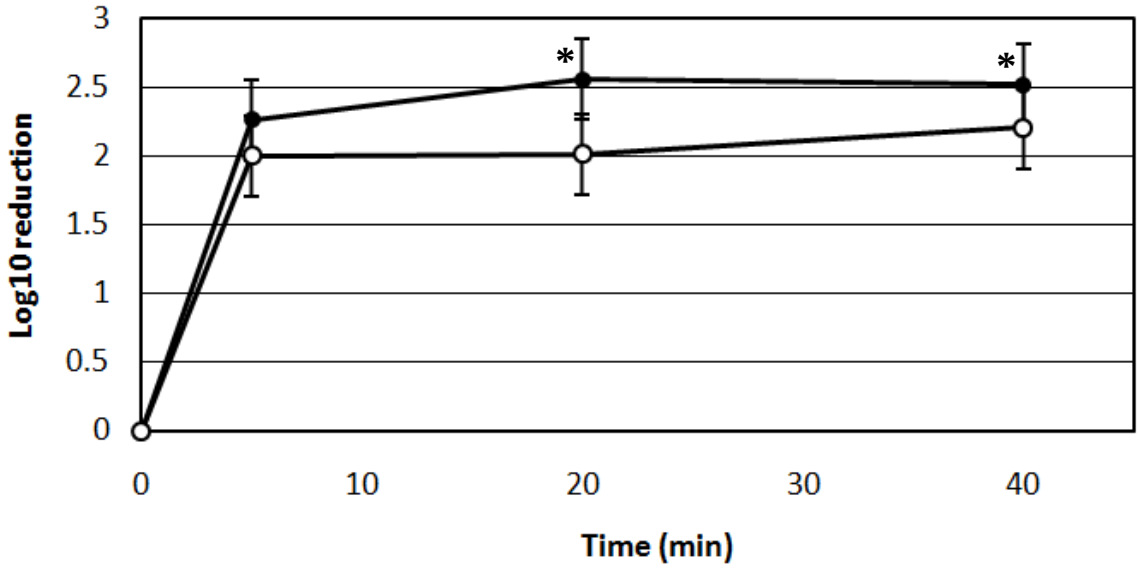
**A. BHI+HCl (pH 2.4)**




**B. BHI+Nisin**



**C. BHI+SDS**



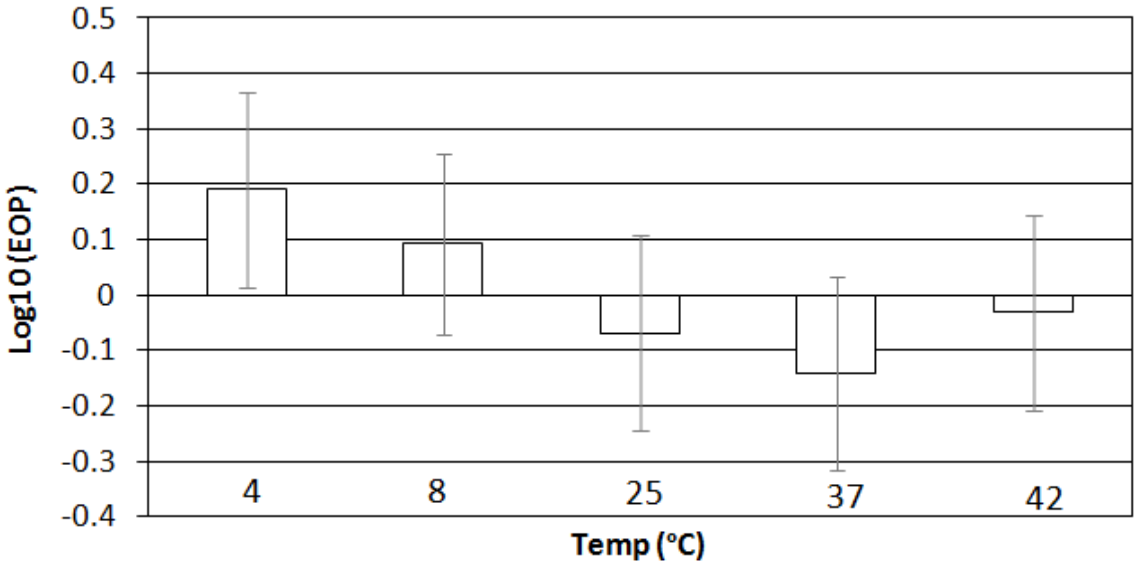



**Figure 5.12 Phage infectivity assays.** The vertical axis measures the relative infectivity to the deletion mutant SL5-L16-5-5 in comparison with the parental strain (F2365) whose log10 (EOP) is always zero. This figure represents the average of at least three independent experiments at one temperature. Error bars signify standard errors calculated from the statistical analysis.

## REFERENCES


1. **Baumgartner, M., U. Karst, B. Gerstel, M. Loessner, J. Wehland, and L. Jansch.** 2007. Inactivation of Lgt allows systematic characterization of lipoproteins from *Listeria monocytogenes*. J. Bacteriol. **189:**313-324.
2. **Behari, J., and P. Youngman.** 1998. Regulation of *hly* expression in *Listeria monocytogenes* by carbon sources and pH occurs through separate mechanisms mediated by PrfA. Infect. Immun. **66:**3635-3642.
3. **Bierne, H., and P. Cossart.** 2007. *Listeria monocytogenes* surface proteins: from genome predictions to function. Microbiol. Mol. Biol. Rev. **71:**377-397.
4. **Cheng, Y., R. M. Siletzky, and S. Kathariou.** 2008. Genomic Divisions/Lineages, Epidemic Clones, and Population Structure, p. 337-358. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.
5. **Doumith, M., C. Cazalet, N. Simoes, L. Frangeul, C. Jacquet, F. Kunst, P. Martin, P. Cossart, P. Glaser, and C. Buchrieser.** 2004. New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. Infect. Immun. **72:**1072-1083.
6. **Drevets, D. A., and M. S. Bronze.** 2008. *Listeria monocytogenes*: epidemiology, human disease, and mechanisms of brain invasion. FEMS Immunol. Med. Microbiol. **53:**151-165.
7. **Eifert, J. D., P. A. Curtis, M. C. Bazaco, R. J. Meinersmann, M. E. Berrang, S. Kernodle, C. Stam, L. A. Jaykus, and S. Kathariou.** 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog. Dis. **2:**192-200.

8. **Elhanafi, D., V. Dutta, and S. Kathariou.** 2010. Genetic characterization of plasmid-associated benzalkonium chloride resistance determinants in a *Listeria monocytogenes* strain from the 1998-1999 outbreak. Appl. Environ. Microbiol. **76:**8231-8238.

9. **Gasanov, U., D. Hughes, and P. M. Hansbro.** 2005. Methods for the isolation and identification of *Listeria* spp. and *Listeria monocytogenes*: a review. FEMS Microbiol. Rev. **29:**851-875.

10. **Harris, L. J., H. P. Fleming, and T. R. Klaenhammer.** 1991. Sensitivity and resistance of *Listeria monocytogenes* ATCC 19115, Scott A, and UAL500 to Nisin. Journal of Food Protection. **54:**836-840.

11. **Kallipolitis, B. H., H. Ingmer, C. G. Gahan, C. Hill, and L. Sogaard-Andersen.** 2003. CesRK, a two-component signal transduction system in *Listeria monocytogenes*, responds to the presence of cell wall-acting antibiotics and affects beta-lactam resistance. Antimicrob. Agents Chemother. **47:**3421-3429.

12. **Kingsford, C. L., K. Ayanbule, and S. L. Salzberg.** 2007. Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake. Genome Biol. **8:**R22.

13. **Lauer, P., M. Y. Chow, M. J. Loessner, D. A. Portnoy, and R. Calendar.** 2002. Construction, characterization, and use of two *Listeria monocytogenes* site-specific phage integration vectors. J. Bacteriol. **184:**4177-4186.

14. **Lemon, K. P., D. E. Higgins, and R. Kolter.** 2007. Flagellar motility is critical for *Listeria monocytogenes* biofilm formation. J. Bacteriol. **189:**4418-4424.

15. **Machata, S., S. Tchatalbachev, W. Mohamed, L. Jansch, T. Hain, and T. Chakraborty.** 2008. Lipoproteins of *Listeria monocytogenes* are critical for virulence and TLR2-mediated immune activation. J. Immunol. **181:**2028-2035.

16. **Marchler-Bauer, A., S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng, and S. H. Bryant.** 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. **39:**D225-9.

17. **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:**2386-2395.

18. **Palmer, M. E., M. Wiedmann, and K. J. Boor.** 2009. sigma(B) and sigma(L) contribute to *Listeria monocytogenes* 10403S response to the antimicrobial peptides SdpC and nisin. Foodborne Pathog. Dis. **6:**1057-1065.

19. **Reglier-Poupet, H., C. Frehel, I. Dubail, J. L. Beretti, P. Berche, A. Charbit, and C. Raynaud.** 2003. Maturation of lipoproteins by type II signal peptidase is required for phagosomal escape of *Listeria monocytogenes*. J. Biol. Chem. **278:**49469-49477.

20. **Rodrigues, D., N. Cerca, P. Teixeira, R. Oliveira, H. Ceri, and J. Azeredo.** 2011. *Listeria monocytogenes* and Salmonella *enterica Enteritidis* Biofilms Susceptibility to Different Disinfectants and Stress-Response and Virulence Gene Expression of Surviving Cells. Microb. Drug Resist.

21. **Ryan, E. M., C. G. Gahan, and C. Hill.** 2008. A significant role for Sigma B in the detergent stress response of *Listeria monocytogenes*. Lett. Appl. Microbiol. **46:**148-154.

22. **Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin.** 2011. Foodborne illness acquired in the United States-major pathogens. Emerg. Infect. Dis. **17:**7-15.

23. **Stack, H. M., C. Hill, and C. G. M. Gahan.** 2008. Stress Responses, p. 61-96. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

24. **Sutcliffe, I. C., and R. R. Russell.** 1995. Lipoproteins of gram-positive bacteria. J. Bacteriol. **177:**1123-1128.

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

26. **Toledo-Arana, A., O. Dussurget, G. Nikitas, N. Sesto, H. Guet-Revillet, D. Balestrino, E. Loh, J. Gripenland, T. Tiensuu, K. Vaitkevicius, M. Barthelemy, M. Vergassola, M. A. Nahori, G. Soubigou, B. Regnault, J. Y. Coppee, M. Lecuit, J.**

**Johansson, and P. Cossart.** 2009. The *Listeria* transcriptional landscape from saprophytism to virulence. Nature. **459:**950-956.

27. **Wagner, M., and J. MacLauchlin.** 2008. Biology, p. 3-26. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

28. **Waibel, A. H., T. Hanazawa, G. E. Hinton, K. Shikano, and K. J. Lang.** 1989. Phoneme Recognition Using Time-Delay Neural Networks. IEEE Transactions on Acoustic, Speech, and Signal Processing. **37:**328-339.

29. **Yildirim, S., D. Elhanafi, W. Lin, A. D. Hitchins, R. M. Siletzky, and S. Kathariou.** 2010. Conservation of genomic localization and sequence content of Sau3AI-like restriction-modification gene cassettes among *Listeria monocytogenes* epidemic clone I and selected strains of serotype 1/2a. Appl. Environ. Microbiol. **76:**5577-5584.

## Chapter 6 Restriction-Modification Systems in *Listeria* Species and Their Phages


### ABSTRACT

*Listeria monocytogenes* is the only member of genus *Listeria* that causes a human disease (listeriosis). Bacteriophages influence ecology and evolution of this bacterium and have attracted attention as a bio-control agent. Restriction-modification (RM) systems hinder bacteriophage infection and horizontal gene transfer. Although one Immigration Control Region (ICR) of *Listeria* species was characterized that harbors various RM systems, *Listeria* RM systems remains poorly understood. In this study, we compared all RM systems and solitary methyltransferases of *Listeria* species and phages available in a public database (REBASE) and analyzed their flanking regions. Our data revealed that *Listeria* species possessed one to six RM systems and solitary methyltransferases, and that several were shared between different strains and species. Phage RM systems and solitary methyltransferases displayed homology with chromosomal RM systems often harbored by the strains within the host range. An additional ICR was identified wherein a family of Type I RM systems and a bacteriophage integrase gene were located. Both findings suggest that phage-mediated transfer might be imperative in disseminating RM systems in this genus. Further functional and population studies on RM systems will increase our understanding on their biological roles affecting the ecology and evolution of *Listeria* genus, particularly *L. monocytogenes*.

## INTRODUCTION

*Listeria monocytogenes* is an opportunistic bacterium which inhabits a variety of environmental niches and is capable of multiplying intracellularly, causing listeriosis in humans and animals (52, 55). Listeriosis is manifested with mild flu-like symptoms and gastroenteritis in healthy adults; however, this disease can be highly invasive for at-risk groups such as people at extreme ages, with immunocompromised conditions or during pregnancy, resulting in bacteremia, meningitis and abortions (47, 52, 55). *L. monocytogenes* is mainly transmitted via foods (29, 47, 52) such as ready-to-eat (RTE) meats, unpasteurized milk and frankfurters (18, 47, 52, 55); and the source of contamination has been often within food processing facilities (18, 52, 55). Previous studies suggested that bacteriophages might play an important role in the ecology and evolution of *L. monocytogenes* within food processing facilities (21, 35, 51). In 2006, a mixture of *Listeria* phages was allowed by FDA to be used as biocontrol agent against *L. monocytogenes* (50).

The population of *L. monocytogenes* consists of three distinct lineages that include different serotypes and have demonstrated different epidemiological attributes and degree of homogeneity (4, 19). Lineage I (serotypes 4b, 1/2b and 3b) is over-represented among clinical isolates and noticeably clonal in comparison with the other lineages (6, 36, 38). Lineage II (serotypes 1/2a, 3a, 1/2c and 3c) often represents isolates from foods and environments; and is comprised of genetically diverse strains (6, 36). Lineage III (serotypes 4a, 4c and certain serotype 4b isolates) is associated with animal listeriosis cases and contains heterogeneous isolates (39). A subset of lineage III has recently been designated into an independent population lineage (lineage IV) (53). Many proposed that a recent bottleneck

event might be responsible for the marked clonality of lineage I (6, 31, 54); however, some suspected that lineage I isolates might be equipped with certain molecular mechanisms that reduce genetic diversity (36).

Restriction-modification (RM) systems are genetic systems that identify and degrade non-self DNA whereas protecting the host genomic DNA (49). Classical RM systems consist of the restriction endonuclease that cleaves the foreign DNA lacking proper methylation; and methyltransferase that attaches methyl groups to the host DNA, thus enabling the cellular DNA to escape from restriction (17, 49). RM systems are divided into Types I to IV (17, 49) based on the enzyme organization, nature of the substrate, and recognition and cleavage sites (32, 49). Type I RM systems consist of the methyltransferase (HsdM), restriction endonuclease (HsdR) and specificity (HsdS) subunits that form enzyme complexes (32, 49). Restriction by these systems occurs at a location ~1000 bp distant from the recognition site (32, 49). Type I RM systems are sub-divided into families whose members share HsdM and HsdR but differ in HsdS (32). Type II RM systems are comprised of the methyltransferase and restriction endonuclease; and cut within or in proximity of the recognition sites (32, 49). Type III RM systems share similarities with Type I RM systems in that its two components, the Mod and Res subunits, form oligomers capable of methylation and restriction (32, 44, 49) and that the cleavage site differs from the recognition sequence by 25-27 bp (44, 49). Type IV RM systems refer to restriction endonucleases whose substrate is methylated DNA (40, 49).

RM systems provide a major barrier against bacteriophage infection (17) and horizontal gene transfer (48). In *L. monocytogenes*, these cellular systems are expected to

influence the evolution mediated by phage infection and lower the potency of the phage mixtures as a control measure. Also, RM systems are one of the molecular mechanisms that might hinder the horizontal gene transfer within lineage I isolates, culminating in the current clonality of this population lineage (36). Previous studies on *L. monocytogenes* RM systems focused on a putative Sau3AI-like RM system that conferred resistance to Sau3AI to a major epidemic-associated clonal group of *L. monocytogenes* (Epidemic Clone I (ECI)) (56). Recent work in our laboratory revealed that the genomic region in *L. monocytogenes* wherein this Sau3AI-like RM system resides is abundant in RM systems of different types such as one family of Type I RM systems, two additional Type II systems, and Type IV proteins (Mrr and McrB). This region that we called "Region 85" displayed considerable diversity among *L. monocytogenes* isolates and other *Listeria* species, and was considered the first Immigration Control Region (ICR) in this species.

Taken together, *L. monocytogenes* RM systems have not been sufficiently studied in spite of the importance both in control and evolution of this microorganism; and our current knowledge is mostly limited to a particular genomic region. Hence, we systematically analyzed the RM systems of *L. monocytogenes* available on a public database as well as those in non-pathogenic *Listeria* species and phages.

## MATERIALS AND METHODS

**Collection and comparison of the RM systems in *Listeria* species and phages.** The DNA and protein sequences of the RM systems examined in this study were retrieved from REBASE (http://rebase.neb.com/rebase/rebase.html) (41) with the keyword of *Listeria* as the

organism name on Jun 22, 2010. Methyltransferases without cognate restriction endonucleases were regarded as solitary methyltransferases and included in the analysis. All these RM systems are listed in Table 1.

When a *Listeria* strain or phage was sequenced at the genomic level, the RM systems from this strain or phage were assumed to represent all the RM systems in the genome. Redundant or non-*Listeria* RM systems were manually eliminated based on the information on the REBASE and literature search. Types of RM systems, functions, methylation types and putative or verified recognition sites were obtained from REBASE and GC contents of the genes were calculated with Oligonucleotide Properties Calculator (http://www.basic.northwestern.edu/biotools/oligocalc.html) (20). Genes encoding the RM systems followed the information presented by REBASE but their names were confirmed and corrected according to the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) database and the relevant literature.

For comparison, components of RM systems were classified into three criteria of methyltransferases, restriction endonucleases and specificity subunits, and each division was compared both at DNA and protein levels with ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) (5). In addition to the RM systems collected from REBASE, we added novel RM systems that we identified in Probe 85 region in the previous study: two Type II methyltransferases (M.LmoJ2P and M.LmoJ3P); two Type II restriction endonucleases (R.LmoJ2P and R.LmoJ3P); and one HsdS (S.LmoJ3976P). The phylograms obtained from these comparisons are exhibited in Figure 1 and the homologous RM systems were listed in Table 1. The DNA sequences encompassing the homologous RM

systems and three kb flanking regions upstream and downstream were compared with the online Artemis Comparison Tool (WebACT, http://www.webact.org/WebACT/home) (3).

**Analysis of the flanking regions of the RM systems.** For the RM systems encoded in the sequenced genomes, three kb DNA sequences flanking the RM systems both in the upstream and downstream were retrieved from NCBI database. When multiple RM systems were present in one genomic region, flanking regions referred to three kb DNA sequences upstream to the gene at the 5' end and downstream to the 3'-proximal gene. These flanking regions were compared with ClustalW and the resulting phylogram is displayed in Figure 2.

For a putative ICR, the syntenic genomic regions were compared with the WebACT and major homologous regions were visualized in Figure 3A. Phylogeny of the genes in Figure 3A was determined by comparing DNA sequences with ClustalW and compared with the phylogeny of the concatenated DNA sequence of the seven housekeeping genes targeted by Ragon et al. for multilocus sequence typing (MLST) (38), which represents the phylogenetic relationship of the entire genome (Fig. 4). Protein homologs were identified via BLASTp at NCBI homepage (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&BLAST_PROGRAMS=blastp&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome) (1). Rho-independent terminators were searched for by means of available genome annotation and manual inspection. Inverted repeats (Fig. 4B, C and D) were identified with einverted program provided by European Bioinformatics Institute (http://www.ebi.ac.uk/soaplab/) (43).

## RESULTS

**Number of RM systems in *Listeria* species and phages.** The search against REBASE with the keyword of *Listeria* as the organism name revealed 103 proteins comprising RM systems in *Listeria* species and phages. Since four were found to be redundant and another four were harbored by another microorganism than *Listeria*, 95 proteins were further analyzed (Table 1). A majority of these RM system proteins (81 proteins) were encoded by genes of the GC content lower than the *L. monocytogenes* genome average of 38% (33) (Table 1).

Analysis with 11 *L. monocytogenes* sequenced strains revealed that serotype 1/2a strains harbored one (10403S) to six (08-5578) RM systems and solitary methyltransferases and serotype 4b strains bore one (F2365) to two (H7858 and others) RM systems and solitary methyltransferases. Serotype 4a strain HCC23 possessed one RM systems and two solitary methyltransferases (Table 2). Of the sequenced *L. monocytogenes* strains, strains 08-5578 and 08-5923 were implicated in the same outbreak that occurred in Canada in 2008; thus, both strains are closely related except for a few differences (11). Analysis with other *Listeria* species genomes revealed that these strains contained two (*L. welshimeri* SLCC5334) to four (*L. innocua* CLIP11262 and *L. seeligeri* SLCC3954) RM systems and solitary methyltransferases (Table 2). A majority of plasmid- and phage-encoded RM system proteins were solitary methyltransferases (Table 2).

**Comparison of *Listeria* RM system proteins.** RM system proteins whose sequences were available in REBASE were compared both at DNA and protein levels (Fig. 1). These analyses revealed that among Type I RM systems, *L. seeligeri* SLCC3954 harbored two

unique Type I RM systems whereas all the other Type I RM systems possessed similar HsdM and HsdR but distinct HsdS (G1 in Fig. 1), suggesting that these Type I RM systems could be assigned into one family. Interestingly, these Type I RM systems except for those in *L. monocytogenes* ATCC19115 possessed two *hsdS* genes whose products will be termed HsdS1 (encoded by the upstream gene) and HsdS2 (encoded by the downstream gene) in this study (Table 1). Comparison of these HsdS subunits suggested that HsdS1 and HsdS2 are generally clustered into two separate groups (Fig. 1C). Exceptions were found in the HsdS subunits in *L. innocua* CLIP11252 (S1.LinCORF522P and S2.LinCOR522P) that were related to each other; and HsdS2 in *L. monocytogenes* CLIP51849 (S2.LmoORF544P) that was distinct from other HsdS subunits (Fig. 1C). The Type I RM system in ATCC19115 contained the specificity subunit that shared similarity with S1 proteins (Fig. 1C). Meanwhile, the novel HsdS identified in Probe 85 region (S.LmoJ3976P) was distantly related to S.Lse3954ORF279P, which comprised the Type I RM system of the potentially same family (Fig. 1C).

Type II RM systems and solitary methyltransferases, which all belonged to Type II RM systems, were characterized with marked diversity in our analysis (Fig. 1). In fact, the novel Type II RM systems identified in Region 85 (LmoJ2P/M.LmoJ2P and LmoJ3P/M.LmoJ3P) revealed close relationship with neither of the other RM system proteins examined here (Fig. 1A and B). Within Type II RM systems and solitary methyltransferases, we identified nine homologous (G2-G10 in Fig. 1). The group G2 consisted of the Sau3AI-like RM systems that previous studies confirmed its presence (33, 56) (Table 1). The group G3 was composed of M.Lmo19115ORF1P/Lmo19115ORF1P (*L.*

*monocytogenes* ATCC19115) and M.LmoORF679P/LmoORF679P (*L. monocytogenes* CLIP81459) (Table 1). An attempt to investigate the flanking region of these RM systems was hampered by the lack of the DNA sequence of the ATCC9115 RM system (Fig. 2A). The group G4 included M.LmoAP/LmoAP (*L. monocytogenes*) and M.LweSORF291P/LweSORF291P (*L. welshimeri* SLCC5334) (Table 1). The upstream flanking region sequences were available for both RM systems and shared similarity (Fig. 2B).

The group G5 consisted of paralogous solitary methyltransferases in *L. monocytogenes* HCC23 (M.Lmo23ORF1402P and M.Lmo23ORF2958P) and the genes encoding these proteins were located inside the inversion of ca. 3 kb (Fig. 2C). A pair of the paralogous solitary methyltransferases (M.LinCORF1262P and M.LinCORF1737P in *L. innocua* CLIP11262 (serotype 6a)) were also identified in the group G6 and an inversion was also identified between the DNA sequences encompassing the genes encoding these proteins. Another solitary methyltransferase member in G6 (M.LmoB054ORFAP) originated from the phage B054, which infects serotypes 5 and 6 (Table 1) (7). The flanking regions revealed similarity among all these solitary methyltransferases, suggesting that the genetic regions in M.LinCORF1262P and M.LinCORF1737P might originate from phages (Fig. 2D).

The group G7 is comprised of the solitary methyltransferase in *L. innocua* CLIP11262 plasmid pLI100 (M.LinCORF29P) and the Type III RM system in *L. monocytogenes* H7858 plasmid pLM80 (M.LmoHORF56P/LmoHORF56P) (Table 1). Comparison of the genetic regions where these proteins reside revealed two inversions involving the methyltransferase genes and transposase genes, respectively (Fig 2E). In

addition, the transposase gene was duplicated in pLI100, resulting in mirror repeats (Fig. 2E). Downstream to the gene encoding M. LmoHORF56P, we found a gene encoding a Type III methyltransferase gene, which might have not been identified by the REBASE (Fig2. E).

The group G8 consisted of M. LinCORF88P (*L. innocua* CLIP11262), M. LmoEORF2316P (*L. monocytogenes* EGDe (1/2a)), M. LmoF6ORF2379P/LmoF6ORF2379P (*L. monocytogenes* F6854 (1/2a)), and M. LmoPSAORF38P/LmoPSAORF38P (phage PSA infecting serotype 4) (Table 1) (7). Between the phage PSA and *L. innocua* CLIP11262, some parts of the flanking regions as well as the methyltransferase genes revealed similarity (Fig. 2F). Meanwhile, multiple inversions were observed between *L. innocua* CLIP11262 and *L. monocytogenes* strains (Fig. 2F). Upstream to the gene encoding M. LmoEORF2316P, we identified a gene homologous to the LmoF6ORF2379P, suggesting that this gene might be a restriction endonuclease gene which has not been identified by REBASE (Fig. 2F).

The group G9 included two phage-encoded solitary methyltransferases M.LmoA006ORFAP in A006 (host range of serotype ½ (7)) and M.LmoA118I in A118 (host range of serotype ½ (7)) along with the chromosome-encoded M.Lmo85578ORF677P in serotype 1/2a *L. monocytogenes* 08-5578 (Table 1). As revealed by the previous study that A006 and A118 shared "early genes" involved in the cellular functions required for the early state of phage infection, e.g., DNA replication, recombination and modification (7), the flanking regions of the methyltransferase genes were highly similar between A006 and A118 (Fig. 2G). However, the homologous DNA sequence was limited to the solitary

methyltransferase gene and some parts of the downstream flanking region between *L. monocytogenes* 08-5578 and the phages (Fig. 2G).

The last Type II homologous RM system group G10 included LinCORF21P/S.LinCORF21P (*L. innocua* CLIP11262 plasmid pLI100) and LmoDORFAP/S.LmoDORFAP (*L. monocytogenes* DRDC8 plasmid pCT100) that are comprised of a protein capable of both restriction and methylation and a specificity protein (Table 1). The main difference between these RM systems was found near the start of a restriction/methyltransferase gene (Fig. 2H). Comparison of the genetic regions revealed complicated genetic changes such as insertion/deletion and inversions (Fig. 2H).

Type III RM systems were identified in *L. monocytogenes* strains involved in 2008 Canadian outbreak (08-5578 and 08-5923) and plasmid pLM80 (Table 1). Type IV RM systems in *Listeria* species were comprised of one Mrr protein in *L. seeligeri* and multiple McrB proteins that were identified in *L. seeligeri* and *L. monocytogenes* (Table 1). The McrB proteins were highly homologous and formed Group 11 (Fig. 1).

**Identification of ICRs.** The availability of the genome-wide information provided an opportunity to explore the genetic locations wherein RM systems often reside. The comparison of the DNA sequences flanking RM systems within sequenced *Listeria* genomes revealed two such genetic locations as depicted in Figure 2. One location was inhabited by one Type I RM system, Mrr and McrB in *L. seeligeri* SLCC3954; the Sau3AI-like RM system in *L. monocytogenes* F2365; and McrB in other *L. monocytogenes* (Fig. 3). In other words, this locus corresponded to Region 85, the previously characterized ICR that is flanked

by a PTA IIA gene in the upstream and a threonine aldolase gene in the downstream in *Listeria* species.

In the other location lies the aforementioned family of Type I RM systems that contained two specificity subunits (Fig. 3 and 4A). This region was flanked in the upstream by an EmrB/QacA family multidrug-resistance gene encoding a Major Facilitator Superfamily (MFS, pfam07690) protein, and a transcription regulator gene coding for a ROK family protein (pfam00480); and in the downstream by a 6-phospho-beta-glucosidase gene (Fig. 4A). At the end of each of these genes, a putative Rho-independent terminator was identified. BLASTp searches with the protein products of these three genes in 08-5578 identified homologs in many *Listeria* species including *L. seeligeri, L. marthii,* and *L. ivanovii*. In fact, even the most distant *Listeria* species (*L. grayi*) (42) harbored the homologs of the multidrug-resistance and 6-phospho-beta-glucosidase proteins. However, the GC contents of these genes ranged between 40-42%, which is higher than the genome average, suggesting that these genes might be introduced via horizontal gene transfer. Indeed, our BLASTp results revealed that among microorganisms other than *Listeria* species, various *Bacillus* species contained homologs to the proteins encoded in these genes. In the meantime, the DNA homologs of the transcription regulator gene exhibited relationships consistent to those of the concatenated MLST target genes (Fig. 5).

Between these flanking genes laid the Type I RM system and a bacteriophage integrase gene was located between two *hsdS* genes (Fig. 4A). The DNA sequence comparison suggested that two short DNA sequence, one within each *hsdS* gene, might be conserved in spite of complicated homology patterns between *hsdS* genes and that these

conserved sequences might form mirror repeats (Fig. 3A). The integrase gene between the *hsdS* genes was generally conserved among different Type I RM systems in Figure 3A. However, an extensive inversion was observed in the integrase gene of *L. monocytogenes* CLIP81459 in comparison with the others (Fig. 4A). Two inverted repeats spanning up to 130 bp (Fig. 4D) were identified in the regions that flank the integrase gene and constitute the 3' end of each *hsdS* gene (Fig. 4B and C). Putative Rho-independent terminators were identified between these inverted repeats and the integrase gene in most cases; however, no Rho-independent terminator was discovered between the integrase gene and *hsdS2* in CLIP81459 (Fig. 4B and C).

All Type I RM systems and integrase genes revealed the GC contents that were lower than the genome average in contrast to the flanking genes (Fig. 4A). The GC contents of *hsdR* and *hsdM* genes amounted to 34-37% and those for *hsdS* and integrase genes further decreased to 31-33%. BLASTp for HsdR, HsdM and the integrase protein in 08-5578 revealed that among non-*Listeria* microorganisms, multiple strains of *Enterococcus faecalis* possessed homologs of these proteins with high percent identity. For instance, *E. faecalis* TX2134 possessed homologs to HsdR (ZP_07555598.1), HsdM (ZP_08555597.1) and integrase (ZP_07555595.1) that exhibited 84, 86 and 82% identities to the corresponding proteins encoded in 08-5578. BLASTp results for HsdS proteins in the strains in Figure 3A showed that these proteins shared similarity with various microorganisms and that half of the protein often displayed homologs distinct from those for the other half, which is indicative of the mosaic characteristics of these HsdS proteins (32). Relationships of the Type I RM

system genes and the integrase genes differed from those of the concatenated MLST target genes (Fig. 5).

## DISCUSSION

**Quantity of RM systems in *Listeria* species.** In this study, RM systems in *Listeria* species and phages were searched for in a public database and analyzed via *in silico* tools, revealing intriguing features of these genetic systems in *Listeria* species and phages. The first finding is that *Listeria* species harbored one to six RM systems. Previous studies indicated that RM systems deviate in number from one bacterial species to another. In one extreme, RM systems were nearly or completely absent in *Rickettsia prowazekii*, *Treponema pallidum*, and *Chlamydia* and *Buchnera* (23). On the other hand, *Helicobacter pylori* genomes contained almost two dozens of RM systems (24). Extraordinarily many RM systems were also identified in *Neisseria gonorrhoeae*, *Haemophilus influenzae*, *N. meningitidis* and *Xylella fastidiosa* (22, 23). Since *Neisseria*, *Helicobacter* and *Haemophilus* species displayed active natural competence (16, 28) and *Xylella fastidiosa* genome contained a DNA transformation transporter (30), RM systems have been proposed to be important in maintaining the integrity of the genome. Although the number of RM systems in *Listeria* species was much less than the aforementioned RM-abundant species, similar rationale might be applied for *Listeria* RM systems. For instance, previous studies suggested that frequent recombination events contributed to the evolution of *L. monocytogenes*; thus, the RM systems in *L. monocytogenes* might result from the need to defend the genome from such

recombination events. In fact, the genome of *L. monocytogenes* encoded putative DNA transformation genes (12).

Marked clonality of lineage I of *L. monocytogenes* led us to the speculation that lineage I strains might possess fewer RM systems than lineage II strains. In our analysis, serotype 1/2a strains (lineage II) generally seemed to bear more RM systems than serotype 4b (lineage I). This finding might reflect additional pressure for lineage II strains to maintain their genomic stability in agreement with the previous studies revealing that lineage II experienced more recombination events in comparison with lineage I (36). However, it should be emphasized that only limited number of *L. monocytogenes* strains was included in our analysis; and further study is warranted to examine difference in the number of RM systems among different population lineages of *L. monocytogenes*.

**Diversity and genetic locations of different types of RM systems of *Listeria* species.** In our previous study, we identified one family of Type I RM systems, two Type IV RM systems (Mrr and McrB) and three different Type II RM systems that were exclusively harbored in Region 85. In this study, we attempted to address (1) how diverse RM systems are in the chromosome of *Listeria* species outside Probe 85 region; (2) whether RM systems are shared among *Listeria* chromosomes, plasmids and phages; and (3) whether there is another location rich in RM systems; and.

In our analysis, *Listeria* species harbored two additional families of Type I RM systems as well as the one in Region 85. One of these newly identified families abode in a genetic location distant from Region 85 which will be discussed later in more detail. For Type IV RM systems, no additional Type IV RM systems were discovered outside Region

85, suggesting that all Type IV RM systems in *Listeria* species might inhabit Region 85. In contrast, a noticeable variety of Type II RM systems was identified in other multiple genetic locations than Region 85 and those regions often revealed complicated genetic changes such as inversions and insertions/deletions.

In our study, plasmid-encoded RM systems were shared only between plasmids; however, *Listeria* phage-encoded RM systems displayed similarity with the chromosomal RM systems of *Listeria* species, suggesting horizontal gene transfer between *Listeria* species and phages. It is worth mentioning that homologs of all phage solitary methyltransferases were identified within the host range of the phages. These RM systems might have been acquired by phages from their host as a strategy to counteract the action of the cellular restriction endonucleases. In this context, absence of restriction endonuclease genes might be due to a lesser need for phages to have restriction endonucleases although this phenomenon might still result from the failure of bioinformatics tools to identify cognate restriction endonucleases. Conversely, phages might have acted as the donor of the homologous RM systems on the chromosome of *Listeria* species.

**Evolution of the second ICR.** Comparison of flanking regions suggested two ICRs in *Listeria* species: previously characterized Region 85; and a genetic region between a transcription regulator gene and a 6-phospho-beta glucosidase gene that harbored Type I RM systems of the same family possessing two *hsdS* genes. Upstream to the transcription regulator gene flanking the second ICR, a multi-drug resistance gene was located in various *Listeria* species that encoded a putative MFS protein similar to QacA/EmrA. MFS proteins have been involved in multi-drug resistance of Gram-positive pathogens as exemplified in

NorA in *Streptococcus aureus* and PmrA in *S. pneumonia* (37). Some MFS proteins were also reported to pump out biocides such as benzalkonium chloride and cetyltrimethylammonium bromide, which were extruded by QacA and QacB in *S. aureus* (34). Based on these previous findings, one might infer that this multidrug-resistance gene product might contribute to the ability for *Listeria* species to colonize various environmental niches by increasing tolerance to antimicrobial agents or biocides. Drug transporter genes are frequently subjected to the control by a transcription regulator encoded by a nearby gene (13). Previous identified regulators of drug exporter proteins have been classified into the AraC, MarR, MerR and TetR families (13). Although a ROK family protein was encoded by the putative transcription regulator gene next to the multi-drug resistance gene of interest, its sheer genetic location would suffice to consider this ROK family protein as a transcriptional regulator of the multi-drug resistance gene. Regardless of the presence of the Type I RM systems, two Rho-independent terminators were predicted at the end of the multi-drug resistance gene and transcription regulator gene, suggesting that these genes might be transcribed into separate mRNAs without influencing the transcription level of the Type I RM system genes.

Clues for the evolution that occurred in this new ICR were found in the bacteriophage integrase gene separating the *hsdS* genes and the inverted repeats in adjacent to this integrase gene. Furuta and colleagues reported in their genetic context analysis of RM systems published in 2010 that both elements are often associated with RM systems (10). In addition to this information, other lines of evidence such as GC content, comparison with other proteins and phylogenetic relationships allowed us to envision the following genetic events

that might have occurred in this location: (1) horizontal transfer through which an ancestor of *Listeria* species acquired the above-mentioned three flanking genes from *Bacillus* species; (2) vertical inheritance of the flanking genes in the evolutionary course of *Listeria* species; (3) integration of a bacteriophage that is capable of infecting both *E. faecalis* and *Listeria* species and possesses the Type I RM system; (4) loss of the phage DNA except for the Type I RM system and integrase gene; and (5) inversion of the integrase gene in certain strains and diversification of *hsdS* genes. In fact, the intergenic region between the transcription regulator gene and the 6-phospho-beta glucosidase gene might provide a site that phages prefer to integrate in a similar manner that three different *Listeria* phages A006, PSA and B025 are present as prophages within $tRNA^{Arg}$ gene (7, 26).

In conclusion, our systematic analysis of currently available RM systems and solitary methyltransferases in *Listeria* species and phages revealed that one to six RM systems and solitary methyltransferases were present in *Listeria* species. Types I and IV RM systems featured relatively limited diversity and tended to be confined in certain genetic locations whereas Type II RM systems considerably varied and exhibited evidence of complicated horizontal transfers between *Listeria* species, plasmids and phages. Another main finding of this study is the identification of an additional ICR which harbored a family of Type I RM systems presumably via phage integration.

All these observations were acquired through the *in silico* analysis of the RM systems in *Listeria* species and phages available in the public database; hence, those included in this analysis tended to be biased towards the RM systems vigorously examined in the previous studies. In spite of this limitation, our study presents intriguing information on the diversity

of RM systems and involvement of phages in transferring these genetic systems. Further efforts to delve into *Listeria* RM systems in particular in terms of prevalence and functional characterization will enhance our understanding on the evolution and ecology of *Listeria* species and on the efficacy of phage solutions as a control measure for *L. monocytogenes.*

**Table 6.1 List of RM systems in Listeria and phages retrieved from REBASE**

| Type | Classification | Strain/ Phages/ Plasmids | Serotype/ Specificity/ Host | Homologous RM | Name | Function | Methylation[a] | Specificity | Gene name | GC content (%) | DNA sequence | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | *L. welshimeri* | SLCC5334 | 6b | G1 | M.LweSORF478P | M | N6A | | *lwe0478* | 34 | NC_008555.1 | (14) |
| | | | | G1 | LweSORF478P | R | | | *lwe0477* | 36 | NC_008555.1 | (14) |
| | | | | G1 | S1.LweSORF478P | S | | | *lwe0479* | 31 | NC_008555.1 | (14) |
| | | | | G1 | S2.LweSORF478P | S | | | *lwe0481* | 31 | NC_008555.1 | (14) |
| I | *L. innocua* | CLIP11262 | 6a | G1 | M.LinCORF522P | M | N6A | | *lin0522* | 35 | NC_003212.1 | (12) |
| | | | | G1 | LinCORF522P | R | | | *lin0521* | 37 | NC_003212.1 | (12) |
| | | | | G1 | S1.LinCORF522P | S | | | *lin0523* | 33 | NC_003212.1 | (12) |
| | | | | G1 | S2.LinCORF522P | S | | | *lin0525* | 32 | NC_003212.1 | (12) |
| I | *L. monocytogenes* | CLIP81459 | 4b | G1 | M.LmoORF544P | M | N6A | | *Lm4b_00544* | 35 | NC_012488.1 | (9) |
| | | | | G1 | LmoORF544P | R | | | *Lm4b_00543* | 37 | NC_012488.1 | (9) |
| | | | | G1 | S1.LmoORF544P | S | | | *Lm4b_00545* | 31 | NC_012488.1 | (9) |
| | | | | G1 | S2.LmoORF544P | S | | | *Lm4b_00547* | 32 | NC_012488.1 | (9) |
| I | *L. monocytogenes* | ATCC 19115 | 4b | G1 | Lmo19115ORF3P | R | | | | 36 | AX416569 | |
| I | *L. monocytogenes* | ATCC 19115 | 4b | G1 | M.Lmo19115ORF2P | M | N6A | | | 35 | AX416623.1 | |
| | | | | G1 | Lmo19115ORF2P | R | | | | 38 | AX416623.1 | |
| | | | | G1 | S.Lmo19115ORF2P | S | | | | 31 | AX416623.1 | |
| I | *L. monocytogenes* | (20)08-5923 | 1/2a | G1 | M.Lmo85923ORF544P | M | N6A | | *LM5923_0544* | 35 | NC_013768.1 | (11) |
| | | | | G1 | Lmo85923ORF544P | R | | | *LM5923_0543* | 36 | NC_013768.1 | (11) |
| | | | | G1 | S1.Lmo85923ORF544P | S | | | *LM5923_0545* | 31 | NC_013768.1 | (11) |

**Table 6.1 Continued**

| Type | Classification | Strain/ Phages/ Plasmids | Serotype/ Specificity/ Host | Homologous RM | Name | Function | Methylation[a] | Specificity | Gene name | GC content (%) | DNA sequence | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | G1 | S2.Lmo85923ORF544P | S | | | *LM5923_0547* | 33 | NC_013768.1 | (11) |
| I | *L. monocytogenes* | (20)08-5578 | 1/2a | G1 | M.Lmo85578ORF545P | M | N6A | | *LM5578_0545* | 35 | NC_013766.1 | (11) |
| | | | | G1 | Lmo85578ORF545P | R | | | *LM5578_0544* | 36 | NC_013766.1 | (11) |
| | | | | G1 | S1.Lmo85578ORF545P | S | | | *LM5578_0546* | 31 | NC_013766.1 | (11) |
| | | | | G1 | S2.Lmo85578ORF545P | S | | | *LM5578_0548* | 33 | NC_013766.1 | (11) |
| I | *Listeria seeligeri* | SLCC3954 | 1/2b | | M.Lse3954ORF1765P | M | N6A | | *lse_1765* | 33 | NC_013891.1 | (46) |
| | | | | | Lse3954ORF1765P | R | | | *lse_1763* | 33 | NC_013891.1 | (46) |
| | | | | | S.Lse3954ORF1765P | S | | | *lse_1764* | 30 | NC_013891.1 | (46) |
| I | *Listeria seeligeri* | SLCC3954 | 1/2b | | M.Lse3954ORF279P | M | N6A | | *lse_0279* | 35 | NC_013891.1 | (46) |
| | | | | | Lse3954ORF279P | R | | | *lse_0281* | 36 | NC_013891.1 | (46) |
| | | | | | S.Lse3954ORF279P | S | | | *lse_0280* | 35 | NC_013891.1 | (46) |
| II | *L. monocytogenes* | F2365 | 4b | G2 | M.LmoFORF327P | M | C5 | GATC | *LMOf2365_0327* | 33 | NC_002973.6 | (33) |
| | | | | G2 | LmoFORF327P | R | | GATC | *LMOf2365_0325* | 30 | NC_002973.6 | (33) |
| II ($S^b$) | *L. monocytogenes* | F4565 | 4b | G2 | M.LmoF4565I*[c] | M | C5 | GATC | | | AJ410373 | (15) |
| II | *L. monocytogenes* | FSL N3-165 | 1/2a | G2 | M.LmoN3ORF1850P | M | C5 | GATC | *LMIG_01850* | 33 | NZ_AARQ02000010.1 | (2) |
| | | | | G2 | LmoN3ORF1850P | R | | GATC | *LMIG_01848* | 30 | NZ_AARQ02000010.1 | (2) |
| II | *L. monocytogenes* | A7 | 1/2a | G2 | M.LmoA7ORFAP | M | C5 | GATC | | 33 | DQ003209 | (56) |
| | | | | G2 | LmoA7ORFAP | R | | GATC | | 30 | DQ003209 | (56) |
| II | *L. monocytogenes* | CLIP81459 | 4b | G3 | M.LmoORF679P | M | C5 | ACGT | *Lm4b_00679* | 33 | NC_012488.1 | (9) |
| | | | | G3 | LmoORF679P | R | | ACGT | *Lm4b_00680* | 30 | NC_012488.1 | (9) |
| II | *L. monocytogenes* | ATCC 19115 | 4b | G3 | M.Lmo19115ORF1P | M | C5 | ACGT | | 33 | AX416366.1 | |

**Table 6.1 Continued**

| Type | Classification | Strain/ Phages/ Plasmids | Serotype/ Specificity/ Host | Homologous RM | Name | Function | Methylation[a] | Specificity | Gene name | GC content (%) | DNA sequence | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | G3 | Lmo19115ORF1P | R | | | | 30 | AX416366.1 | |
| II | *L. welshimeri* | SLCC5334 | 6b | G4 | M.LweSORF291P | M | C5 | GTCG AC | *lwe0291* | 33 | NC_008555.1 | (14) |
| | | | | G4 | LweSORF291P | R | | GTCG AC | *lwe0292* | 29 | NC_008555.1 | (14) |
| II | *L. monocytogenes* | | | G4 | M.LmoAP | M | C5 | GTCG AC | *lcmA* | 33 | AJ302030.1 | |
| | | | | G4 | LmoAP | R | | GTCG AC | *Hsp* | 30 | AJ302030.1 | |
| II (S) | *L. monocytogenes* | HCC23 | 4a | G5 | M.Lmo23ORF2958P | M | C5 | | *LMHCC_2958* | 39 | NC_011660.1 | (45) |
| II (S) | *L. monocytogenes* | HCC23 | 4a | G5 | M.Lmo23ORF1402P | M | C5 | | *LMHCC_1402* | 43 | NC_011660.1 | (45) |
| II (S) | *L. innocua* | CLIP11262 | 6a | G6 | M.LinCORF1737P | M | N4C or N6A | | *lin1737* | 36 | NC_003212.1 | (12) |
| II (S) | *L. innocua* | CLIP11262 | 6a | G6 | M.LinCORF1262P | M | N4C or N6A | | *lin1262* | 36 | NC_003212.1 | (12) |
| II (S) | *Phage* | B054 | 5, 6 | G6 | M.LmoB054ORFAP | M | N4C or N6A | | *LiPB054_gp77* | 35 | NC_009813.1 | (7) |
| II (S) | *Plasmid* | pLI100 | *L. innocua* CLIP11262 | G7 | M.LinCORF29P | M | N4C | | *pli0029* | 33 | NC_003383.1 | (12) |
| II (S) | *L. innocua* | CLIP11262 | 6a | G8 | M.LinCORF88P | M | N4C or N6A | | *lin0088* | 37 | NC_003212.1 | (12) |
| II | *L. monocytogenes* | F6854 | 1/2a | G8 | M.LmoF6ORF2379P | M | N4C or N6A | | *LMOf6854_2379* | 37 | NZ_AADQ01000014.1 | (33) |
| | | | | G8 | LmoF6ORF2379P | R | | | *LMOf6854_2378* | 38 | NZ_AADQ01000014.1 | (33) |
| II (S) | *L. monocytogenes* | EGDe | 1/2a | G8 | M.LmoEORF2316P | M | N4C or N6A | | *lmo2316* | 37 | NC_003210.1 | (12) |
| II | *Phage* | PSA | 4 | G8 | M.LmoPSAORF38P | M | N4C or N6A | | *ORF34* | 37 | AJ312240.2 | (57) |
| | | | | G8 | LmoPSAORF38P | R | | | *ORF35* | 39 | AJ312240.2 | (57) |

**Table 6.1 Continued**

| Type | Classification | Strain/ Phages/ Plasmids | Serotype/ Specificity/ Host | Homologous RM | Name | Function | Methylation[a] | Specificity | Gene name | GC content (%) | DNA sequence | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| II (S) | *L. monocytogenes* | (20)08-5578 | 1/2a | G9 | M.Lmo85578ORF677P | M | N4C or N6A | | *LM5578_0677* | 38 | NC_013766.1 | (11) |
| II (S) | *Phage* | A118 | ½ | G9 | M.LmoA118I | M | N4C or N6A | | *A118054* | 37 | NC_003216.1 | (27) |
| II (S) | *Phage* | A006 | ½ | G9 | M.LmoA006ORFAP | M | N4C or N6A | | *LiPA006_gp44* | 37 | NC_009815.1 | (7) |
| II | *Plasmid* | pCT100 | *L. monocytogenes* DRDC8 | G10 | RM.LmoDORFAP | RM | N6A | | *pCT0012* | 34 | U15554.3 | (25) |
| | | | | G10 | S.LmoDORFAP | S | | | *pCT0013* | 31 | U15554.3 | (25) |
| II | *Plasmid* | pLI100 | *L. innocua* CLIP11262 | G10 | LinCORF21P | RM | N6A | | *pli0021* | 35 | NC_003383.1 | (12) |
| | | | | G10 | S.LinCORF21P | S | | | *pli0022* | 31 | NC_003383.1 | (12) |
| II | *L. monocytogenes* | HCC23 | 4a | | M.Lmo23ORF2320P | M | C5 | GGCC | *LMHCC_2320* | 31 | NC_011660.1 | (45) |
| | | | | | Lmo23ORF2320P | R | | GGCC | *LMHCC_2321* | 30 | NC_011660.1 | (45) |
| II | *L. monocytogenes* | H7858 | 4b | | M.LmoHORF2754P | M | N4C or N6A | GTATCC | *LMOh7858_2754* | 32 | NZ_AADR01000037.1 | (33) |
| | | | | | LmoHORF2754P | R | | GTATCC | *LMOh7858_2753* | 32 | NZ_AADR01000037.1 | (33) |
| II | *L. monocytogenes* | FSL N1-017 | 4b (1/2b or 3b)[d] | | M.LmoN1ORF1065P | M | N4C or N6A | | *LMHG_01065* | 44 | NZ_AARP04000008.1* | (2) |
| | | | | | LmoN1ORF1065P | R | | | *LMHG_01066* | 41 | NZ_AARP04000008.1* | (2) |
| II (S) | *L. monocytogenes* | FSL N1-017 | 4b (1/2b or 3b) | | LmoN1ORF1083P | RM | N6A | | *LMHG_01083* | 37 | NZ_AARP04000008.1* | (2) |
| II (S) | *L. monocytogenes* | EGDe | 1/2a | | M.LmoEORF1119P | M | N4C or N6A | | *lmo1119* | 29 | NC_003210.1 | (12) |
| II (S) | *L. monocytogenes* | EGDe | 1/2a | | M.LmoEORF470P | M | N4C or N6A | CTS[e]AG | *lmo0470* | 35 | NC_003210.1 | (12) |
| II (S) | *L. monocytogenes* | (20)08-5923 | 1/2a | | M.Lmo85923ORF1853P | M | N4C or N6A | | *LM5923_1853* | 45 | NC_013768.1 | (11) |

**Table 6.1 Continued**

| Type | Classification | Strain/ Phages/ Plasmids | Serotype/ Specificity/ Host | Homologous RM | Name | Function | Methylation[a] | Specificity | Gene name | GC content (%) | DNA sequence | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| II (S) | *L. monocytogenes* | (20)08-5923 | 1/2a | | M.Lmo85923ORF1805P | M | N4C or N6A | GTGC AG | *LM5923_1805* | 28 | NC_013768.1 | (11) |
| II (S) | *L. monocytogenes* | (20)08-5578 | 1/2a | | M.Lmo85578ORF1901P | M | N4C or N6A | | *LM5578_1901* | 45 | NC_013766.1 | (11) |
| II (S) | *L. monocytogenes* | (20)08-5578 | 1/2a | | M.Lmo85578ORF1853P | M | N4C or N6A | GTGC AG | *LM5578_1853* | 28 | NC_013766.1 | (11) |
| II (S) | *L. welshimeri* | RFL131 | | | LweI* | R | | GCAT C (5/9) | | | | |
| II (S) | *Listeria murrayi* | 60 | | | Lmu60I* | R | | CC^T NAGG | | | | |
| II (S) | *Plasmid* | pLM80 | *L. monocytogenes* H7858 | | M.LmoHORF40P | M | N4C or N6A | | *LMOh7858_pLM80_0040* | 39 | NZ_AADR01000006.1 | (33) |
| III | *Plasmid* | pLM80 | *L. monocytogenes* H7858 | G7 | M.LmoHORF56P | M | N6A | | *LMOh7858_pLM80_0056* | 34 | NZ_AADR01000010.1 | (33) |
| | | | | | LmoHORF56P | R | | | *LMOh7858_pLM80_0058* | 34 | NZ_AADR01000010.1 | (33) |
| III | *L. monocytogenes* | (20)08-5923 | 1/2a | | M.Lmo85923ORF1134P | M | N6A | | *LM5923_1134* | 36 | NC_013768.1 | (11) |
| | | | | | Lmo85923ORF1134P | R | | | *LM5923_1135* | 35 | NC_013768.1 | (11) |
| III | *L. monocytogenes* | (20)08-5578 | 1/2a | | M.Lmo85578ORF1180P | M | N6A | | *LM5578_1180* | 36 | NC_013766.1 | (11) |
| | | | | | Lmo85578ORF1180P | R | | | *LM5578_1181* | 35 | NC_013766.1 | (11) |
| IV | *Listeria seeligeri* | SLCC3954 | 1/2b | G11 | Lse3954McrBP | R | | | *lse_0288* | 31 | NC_013891.1 | (46) |
| IV | *L. monocytogenes* | H7858 | 4b | G11 | LmoHMcrBP | R | | | *LMOh7858_0337* | 31 | NZ_AADR01000001.1 | (33) |
| IV | *L. monocytogenes* | FSL N3-165 | 1/2a | G11 | LmoN3McrB2P | R | | | Intergenic region between *LMIG_01852* and *LMIG_01853* | 32 | NZ_AARQ02000010.1 | (2) |

**Table 6.1 Continued**

| Type | Classification | Strain/ Phages/ Plasmids | Serotype/ Specificity/ Host | Homologous RM | Name | Function | Methylation[a] | Specificity | Gene name | GC content (%) | DNA sequence | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IV | *L. monocytogenes* | F6854 | 1/2a | G11 | LmoF6McrBP | R | | | *LMOf6854_0319* | 31 | NZ_AADQ01000021.1 | (33) |
| IV | *L. monocytogenes* | EGDe | 1/2a | G11 | LmoEMcrBP | R | | | *lmo0304* | 31 | NC_003210.1 | (12) |
| IV | *L. monocytogenes* | 10403S | 1/2a | G11 | Lmo10403McrBP | R | | | *LMRG_00001* | 30 | AARZ01000001.1 | (2) |
| IV | *L. monocytogenes* | (20)08-5923 | 1/2a | G11 | Lmo85923McrBP | R | | | *LM5923_0349* | 31 | NC_013768.1 | (11) |
| IV | *L. monocytogenes* | (20)08-5578 | 1/2a | G11 | Lmo85578McrBP | R | | | *LM5578_0350* | 31 | NC_013766.1 | (11) |
| IV | *Listeria seeligeri* | SLCC3954 | 1/2b | | Lse3954MrrP | R | | | *lse_0282* | 34 | NC_013891.1 | (46) |

[a] N6A, methylation of nitrogen at position 6 of adenine; N4C, methylation of nitroge at position 4 of cytosine; and C5, methylation of carbon at position 5 of cytosine.

[b] II (S), solitary TypeII methyltransferase.

[c] Asterisk represents that no sequence information was available for the RM system protein in REBASE.

[d] Serotype within the parentheses was obtained from the multiplex PCR serotyping devised by Doumith et al (8).

[e] S, C or G.

**Table 6.2 Number of the RM systems in *Listeria* and phages retrieved from REBASE.**

| Type | I | II | III | IV | Total |
|---|---|---|---|---|---|
| **Chromosome** | **30 (9)**[a] | **41 (12/ 17)** [b] | **4 (2)**[a] | **9 (9)**[a] | **84 (46)**[c] |
| ***L. monocytogenes*** | **16 (5)** | **32 (10/ 12)** | **4 (2)** | **7 (7)** | **59 (36)** |
| (20)08-5578 (1/2a) | 4 (1) | 3 (0/ 3) | 2 (1) | 1 (1) | 10 (6) |
| (20)08-5923 (1/2a) | 4 (1) | 2 (0/ 2) | 2 (1) | 1 (1) | 9 (5) |
| 10403S (1/2a) | | | | 1 (1) | 1 (1) |
| EGDe (1/2a) | | 3 (0/ 3) | | 1 (1) | 4 (4) |
| F6854 (1/2a) | | 2 (1/ 0) | | 1 (1) | 3 (2) |
| FSL N3-165 (1/2a) | | 2 (1/ 0) | | 1 (1) | 3 (2) |
| CLIP80459 (4b) | 4 (1) | 2 (1/ 0) | | | 6 (2) |
| F2365 (4b) | | 2 (1/ 0) | | | 2 (1) |
| FSL N1-017 (4b (1/2b or 3b)) | | 3 (1/ 1) | | | 3 (2) |
| H7858 (4b) | | 2 (1/ 0) | | 1 (1) | 3 (2) |
| HCC23 (4a) | | 4 (1/ 2) | | | 4 (3) |
| Others[d] | 4 (2) | 7 (3/ 1) | | | 11 (6) |
| ***L. innocua*** | **4 (1)** | **3 (0/ 3)** | | | **7 (4)** |
| CLIP11262 (6a) | 4 (1) | 3 (0/ 3) | | | 7 (4) |
| ***L. murrayi*** | | **1 (0/ 1)** | | | **1 (1)** |
| ***L. seeligeri*** | **6 (2)** | | | **2 (2)** | **8 (4)** |
| SLCC3954 (1/2b) | 6 (2) | | | 2 (2) | 8 (4) |
| ***L. welshimeri*** | **4 (1)** | **3 (1/ 1)** | | | **7 (3)** |
| SLCC5334 (6b) | 4 (1) | 2 (1) | | | 6 (2) |
| Others | | 1 (1) | | | 1 (1) |
| **Phage** | | **5 (1/ 3)** | | | **5 (4)** |
| A006 | | 1 (0/ 1) | | | 1 (1) |
| A118 | | 1 (0/ 1) | | | 1 (1) |
| B054 | | 1 (0/ 1) | | | 1 (1) |
| PSA | | 2 (1/ 0) | | | 2 (1) |
| **Plasmid** | | **6 (2/ 2)** | **2 (1)** | | **8 (5)** |
| pLI100 | | 3 (1/ 1) | | | 3 (2) |
| pLM80 | | 1 (0/ 1) | 2 (1) | | 3 (2) |
| pCT100 | | 2 (1/0) | | | 2 (1) |
| **Total** | **30 (9)** | **50 (14/ 22)** | **6 (3)** | **9 (9)** | **95 (57)** |

[a] Numbers outside and inside parentheses indicate the number of individual RM system proteins and RM systems, respectively.

[b] The number after the slash mark represent the number of solitary methyltransferases.

**Table 6.2 Continued**

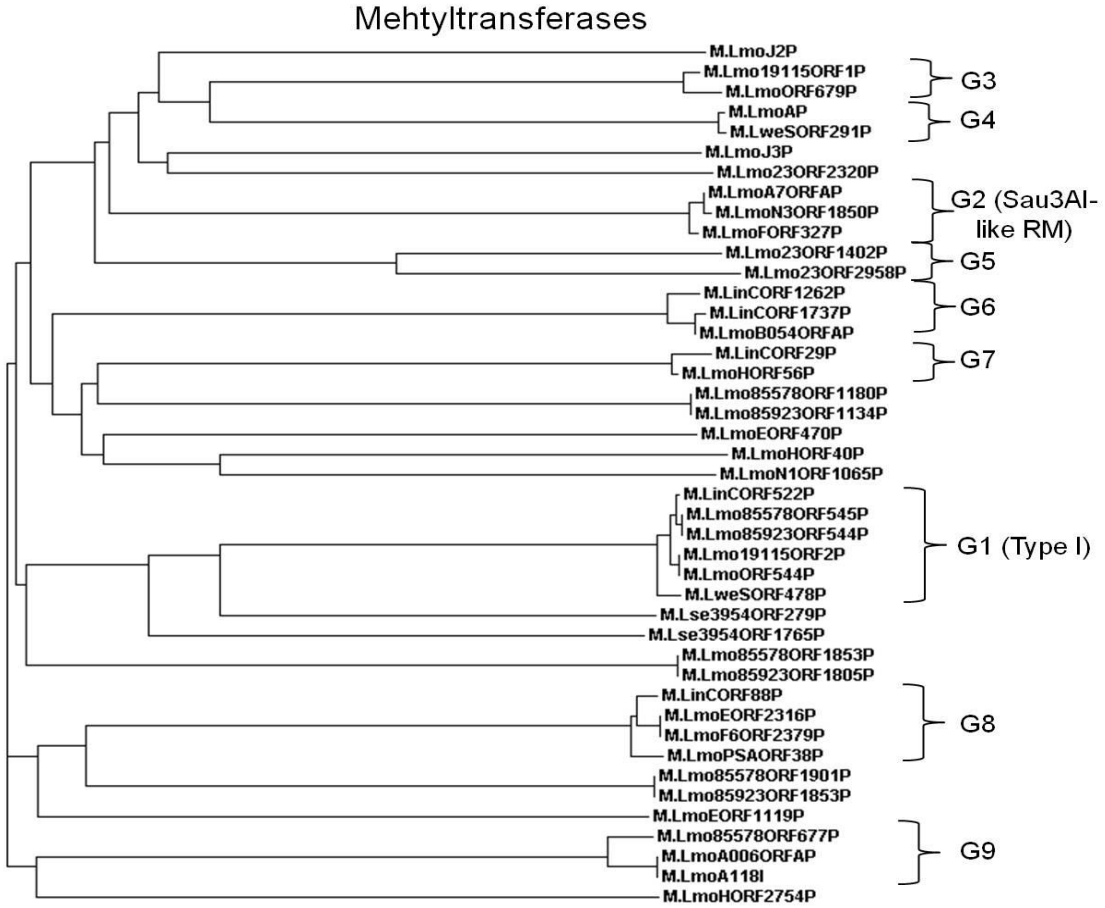
[c] Numbers outside and inside parentheses indicate the number of individual RM system proteins and the number of RM systems and solitary methyltransferases, respectively.

[d] This serotype was obtained from the multiplex PCR serotyping devised by Doumith et al (8).

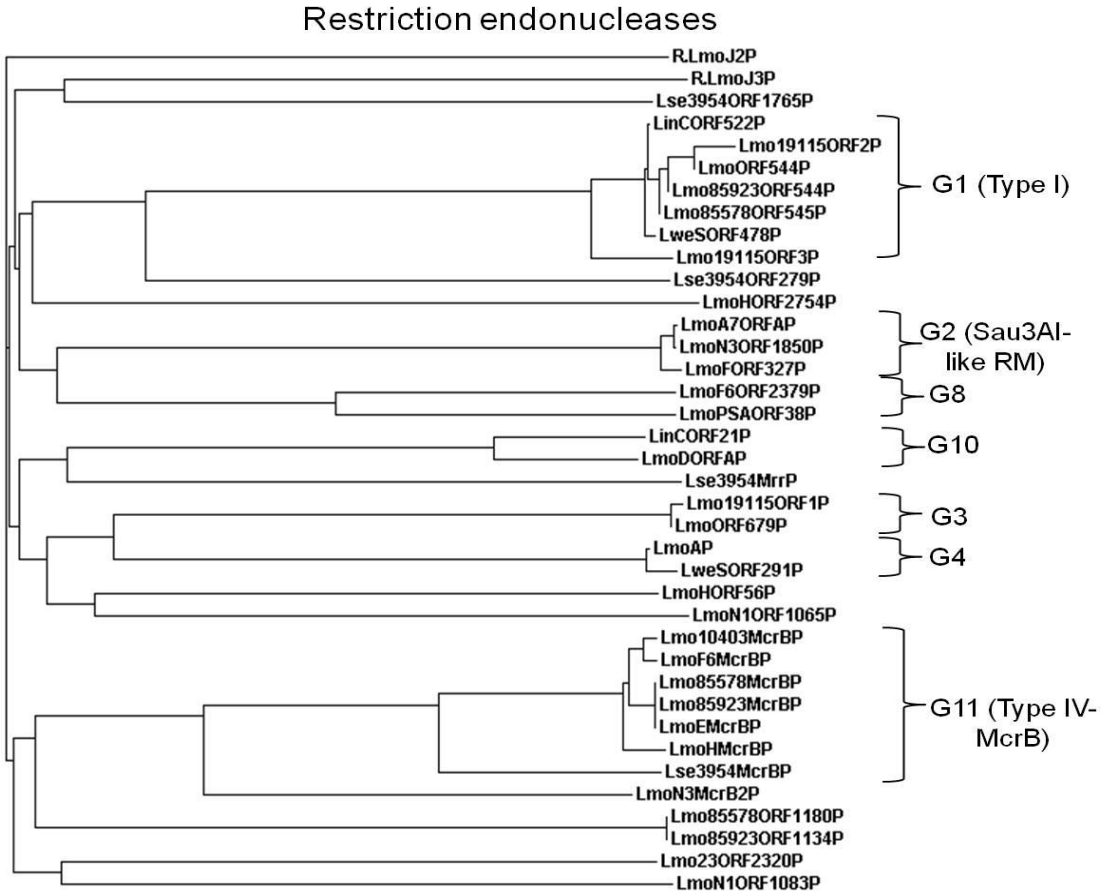
[e] Others category refers to the strains whose genome has not been sequenced.

**Figure 6.1 Comparison of methyltransferases (A), restriction endonucleases (B) and specificity subunits (C) comprising RM systems in *Listeria* and phages.**

**A.**


Mehtyltransferases
M.LmoJ2P
M.Lmo19115ORF1P
M.LmoORF679P
M.LmoAP
M.LweSORF291P
M.LmoJ3P
M.Lmo23ORF2320P
M.LmoA7ORFAP
M.LmoN3ORF1850P
M.LmoFORF327P
M.Lmo23ORF1402P
M.Lmo23ORF2958P
M.LinCORF1262P
M.LinCORF1737P
M.LmoB054ORFAP
M.LinCORF29P
M.LmoHORF56P
M.Lmo85578ORF1180P
M.Lmo85923ORF1134P
M.LmoEORF470P
M.LmoHORF40P
M.LmoN1ORF1065P
M.LinCORF522P
M.Lmo85578ORF545P
M.Lmo85923ORF544P
M.Lmo19115ORF2P
M.LmoORF544P
M.LweSORF478P
M.Lse3954ORF279P
M.Lse3954ORF1765P
M.Lmo85578ORF1853P
M.Lmo85923ORF1805P
M.LinCORF88P
M.LmoEORF2316P
M.LmoF6ORF2379P
M.LmoPSAORF38P
M.Lmo85578ORF1901P
M.Lmo85923ORF1853P
M.LmoEORF1119P
M.Lmo85578ORF677P
M.LmoA006ORFAP
M.LmoA118I
M.LmoHORF2754P
G3
G4
G2 (Sau3AI-like RM)
G5
G6
G7
G1 (Type I)
G8
G9

**B.**


Restriction endonucleases
R.LmoJ2P
R.LmoJ3P
Lse3954ORF1765P
LinCORF522P
Lmo19115ORF2P
LmoORF544P
Lmo85923ORF544P
Lmo85578ORF545P
LweSORF478P
Lmo19115ORF3P
Lse3954ORF279P
LmoHORF2754P
LmoA7ORFAP
LmoN3ORF1850P
LmoFORF327P
LmoF6ORF2379P
LmoPSAORF38P
LinCORF21P
LmoDORFAP
Lse3954MrrP
Lmo19115ORF1P
LmoORF679P
LmoAP
LweSORF291P
LmoHORF56P
LmoN1ORF1065P
Lmo10403McrBP
LmoF6McrBP
Lmo85578McrBP
Lmo85923McrBP
LmoEMcrBP
LmoHMcrBP
Lse3954McrBP
LmoN3McrB2P
Lmo85578ORF1180P
Lmo85923ORF1134P
Lmo23ORF2320P
LmoN1ORF1083P
G1 (Type I)
G2 (Sau3AI-like RM)
G8
G10
G3
G4
G11 (Type IV-McrB)

**C.**

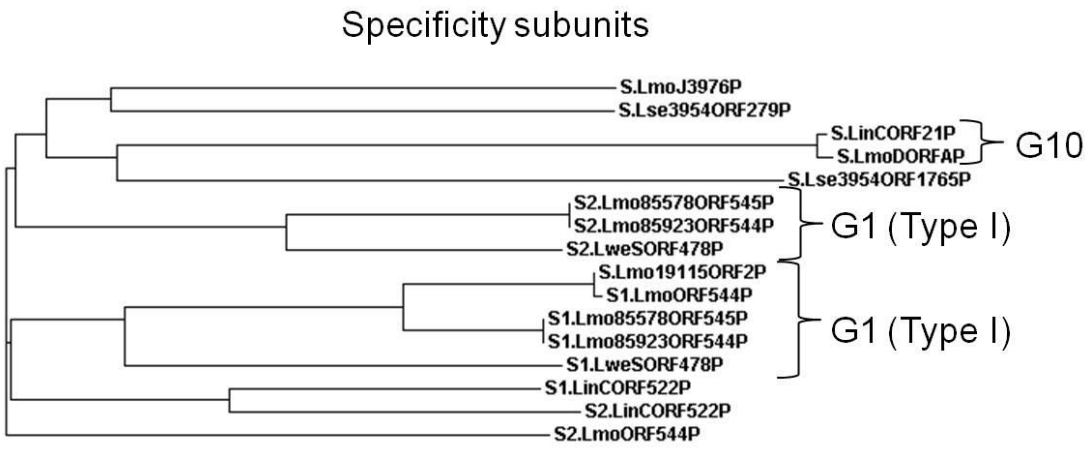
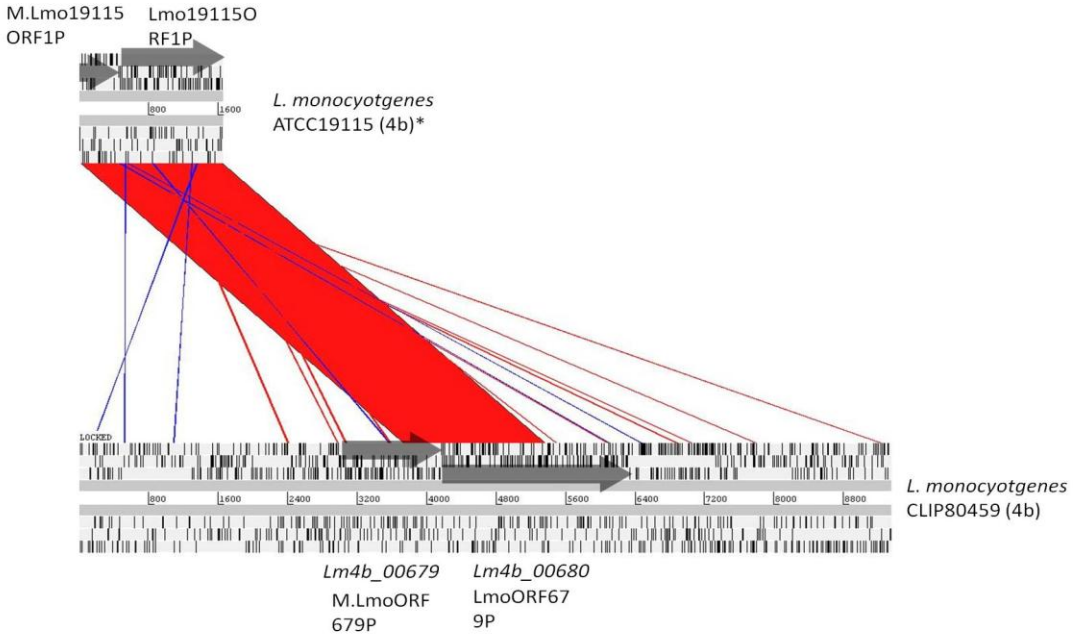
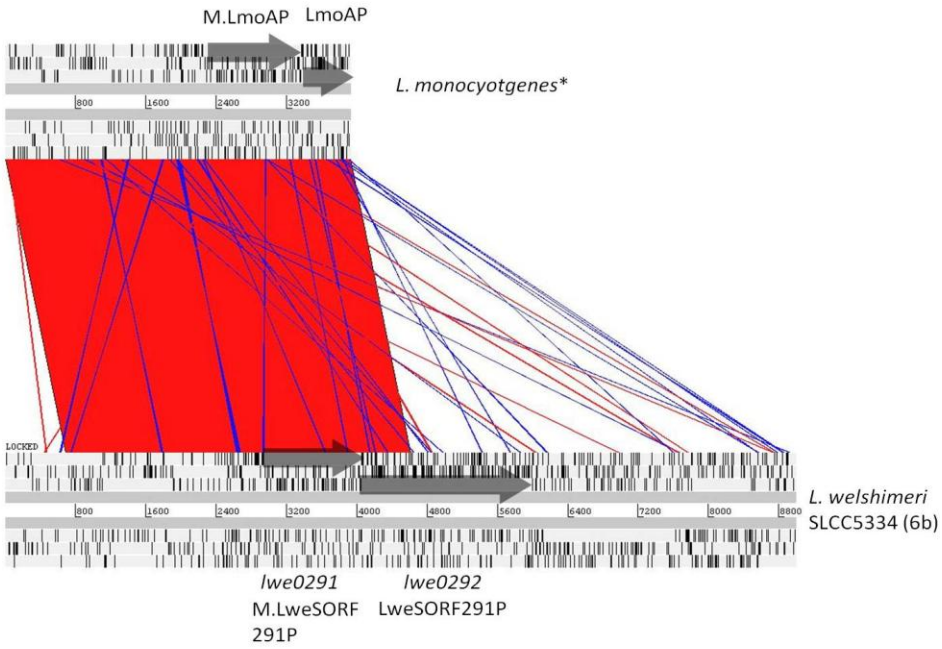

Specificity subunits
S.LmoJ3976P
S.Lse3954ORF279P
S.LinCORF21P
S.LmoDORFAP
G10
S.Lse3954ORF1765P
S2.Lmo85578ORF545P
S2.Lmo85923ORF544P
S2.LweSORF478P
G1 (Type I)
S.Lmo19115ORF2P
S1.LmoORF544P
S1.Lmo85578ORF545P
S1.Lmo85923ORF544P
S1.LweSORF478P
G1 (Type I)
S1.LinCORF522P
S2.LinCORF522P
S2.LmoORF544P

**Figure 6.2 Comparison of the DNA of the homologous RM systems (G3-G10 shown in A-H, respectively) with the WebACT.** The DNA sequence represents the RM system genes and additional 3 kb upstream and downstream. When the DNA fragment was too short to retrieve the 3 kb flanking region, it was marked with asterisk. The genes encoding the RM system proteins were colored in grey and some noteworthy flanking genes were marked in red.

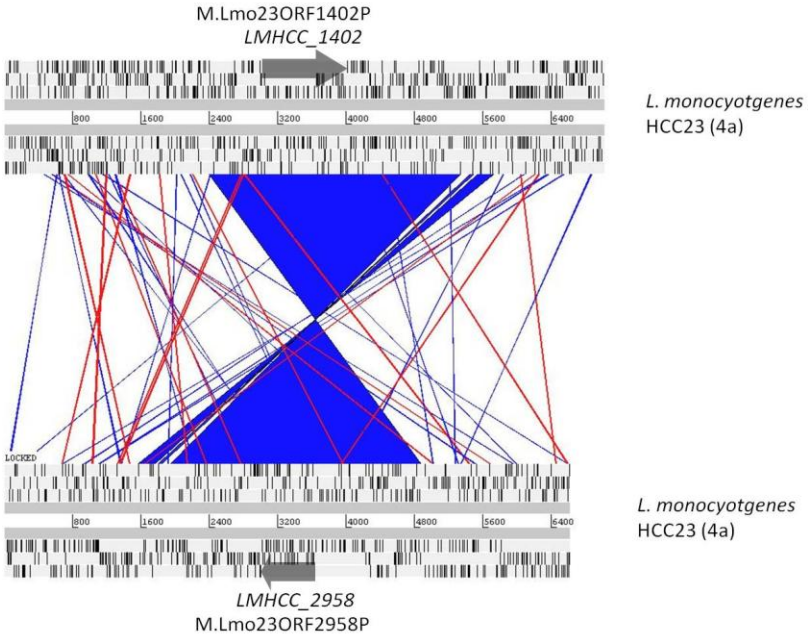
**A. G3**


M.Lmo19115 ORF1P
Lmo19115O RF1P
L. monocytogenes ATCC19115 (4b)*
L. monocytogenes CLIP80459 (4b)
Lm4b_00679 M.LmoORF 679P
Lm4b_00680 LmoORF67 9P

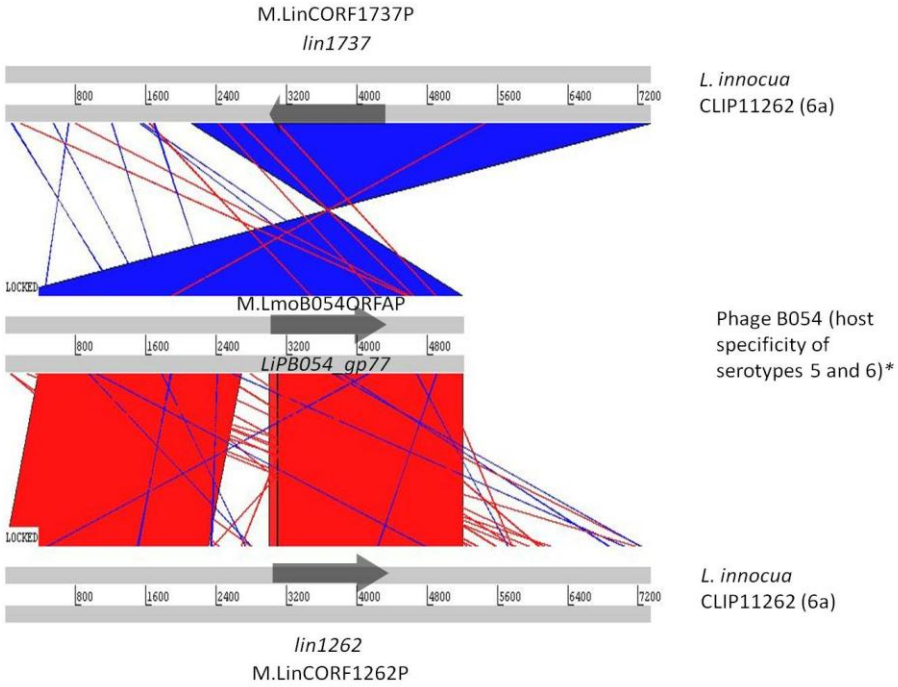
**B. G4**


M.LmoAP
LmoAP
L. monocytogenes*
L. welshimeri SLCC5334 (6b)
lwe0291 M.LweSORF 291P
lwe0292 LweSORF291P

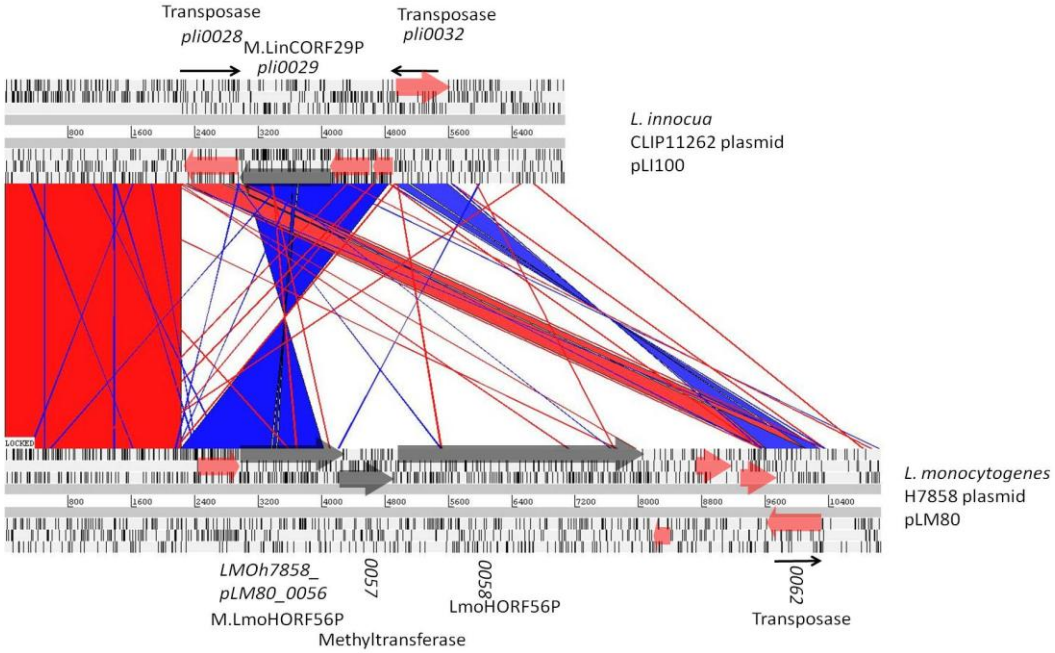
## C. G5


M.Lmo23ORF1402P
LMHCC_1402
L. monocyotgenes
HCC23 (4a)
L. monocyotgenes
HCC23 (4a)
LMHCC_2958
M.Lmo23ORF2958P

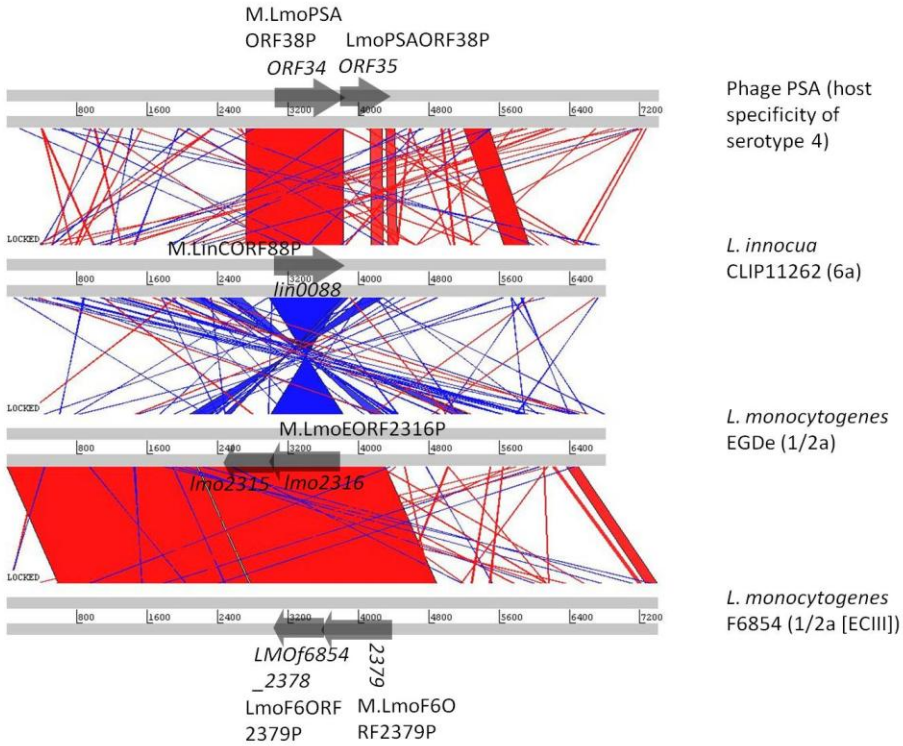
## D. G6


M.LinCORF1737P
lin1737
L. innocua
CLIP11262 (6a)
M.LmoB054ORFAP
LiPB054_gp77
Phage B054 (host specificity of serotypes 5 and 6)*
L. innocua
CLIP11262 (6a)
lin1262
M.LinCORF1262P

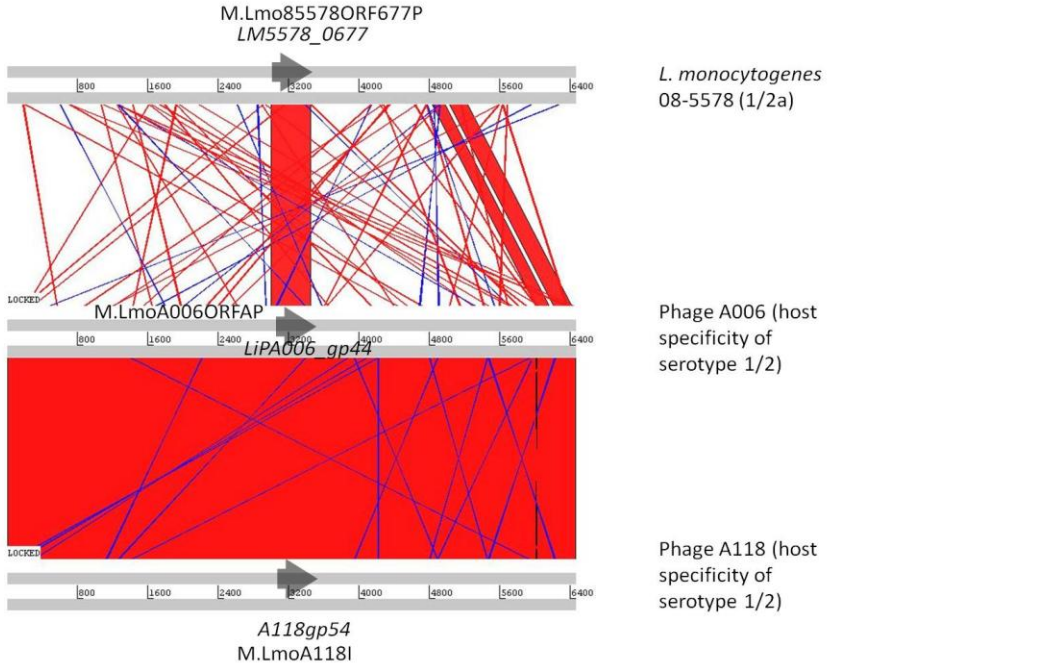
## E. G7


Transposase
pli0028
M.LinCORF29P
pli0029
Transposase
pli0032
L. innocua
CLIP11262 plasmid
pLI100
L. monocytogenes
H7858 plasmid
pLM80
LMOh7858_
pLM80_0056
M.LmoHORF56P
Methyltransferase
0057
0058
LmoHORF56P
0062
Transposase

## F. G8


M.LmoPSA
ORF38P
LmoPSAORF38P
ORF34
ORF35
Phage PSA (host specificity of serotype 4)
M.LinCORF88P
lin0088
L. innocua
CLIP11262 (6a)
M.LmoEORF2316P
lmo2315
lmo2316
L. monocytogenes
EGDe (1/2a)
L. monocytogenes
F6854 (1/2a [ECIII])
LMOf6854
_2378
LmoF6ORF
2379P
2379
M.LmoF6O
RF2379P

## G. G9


M.Lmo85578ORF677P
LM5578_0677
L. monocytogenes
08-5578 (1/2a)
M.LmoA006ORFAP
LiPA006_gp44
Phage A006 (host specificity of serotype 1/2)
Phage A118 (host specificity of serotype 1/2)
A118gp54
M.LmoA118I

## H. G10

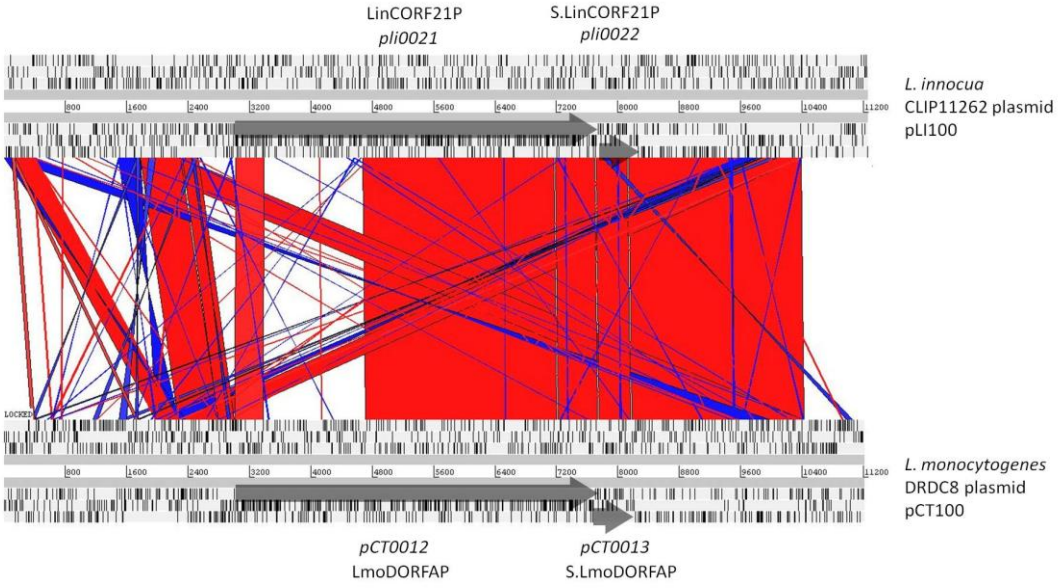

LinCORF21P
pli0021
S.LinCORF21P
pli0022
L. innocua
CLIP11262 plasmid
pLI100
L. monocytogenes
DRDC8 plasmid
pCT100
pCT0012
LmoDORFAP
pCT0013
S.LmoDORFAP

**Figure 6.3 Comparison of the regions flanking the RM systems in *Listeria* and phages whose genome was sequenced.** The Up, Dn and RC represent upstream flanking region, downstream flanking region and reverse complementary, respectively. Flanking regions were obtained from the DNA strand on which the locus tag was numbered from the smallest to the largest. For HCC23, however, the reverse complementary sequence of the flanking DNA sequence was used since its locus tag was number on the opposite strand. Black triangles signify that the flanking region was shorter than 3 kb due to the lack of sequence information.

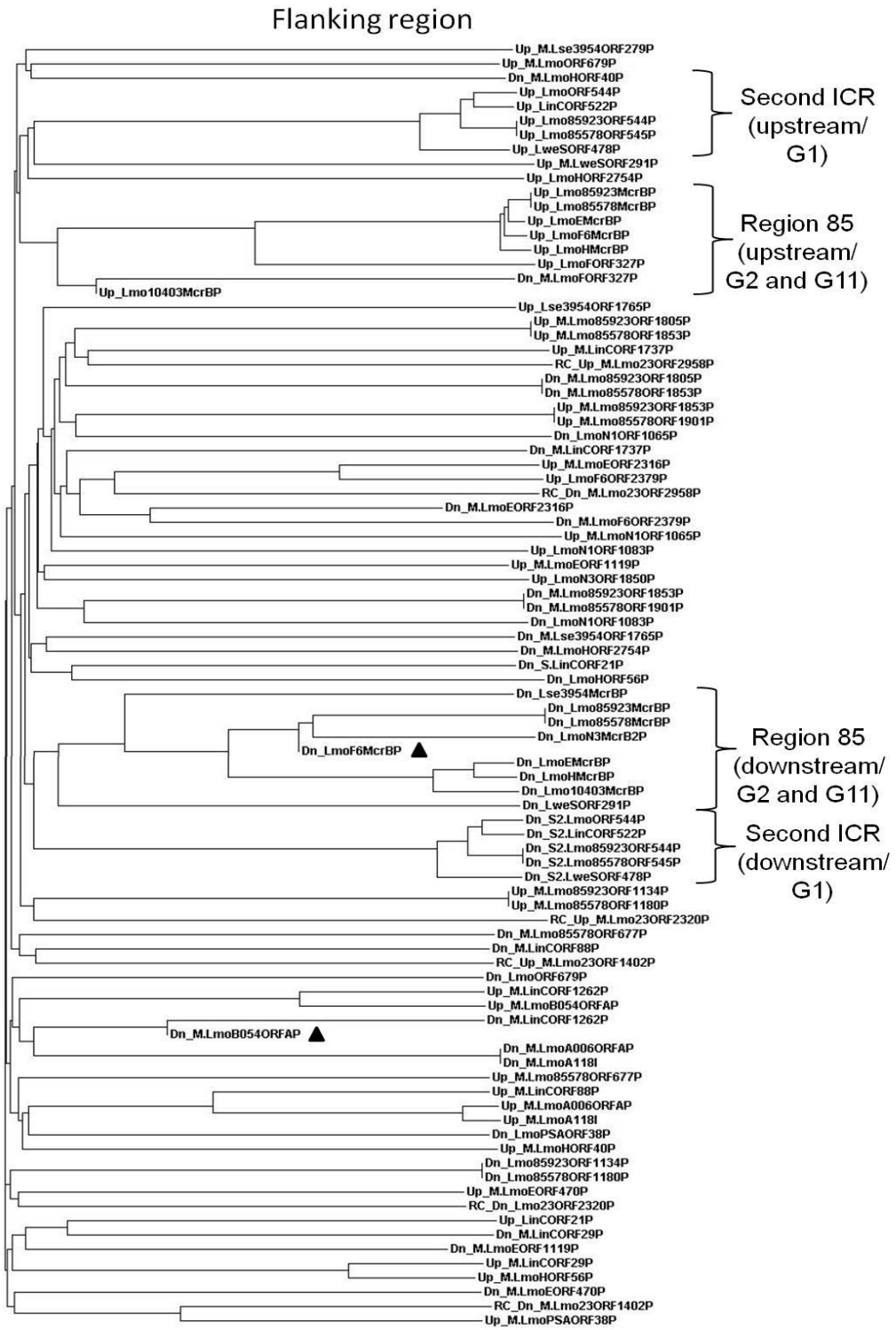
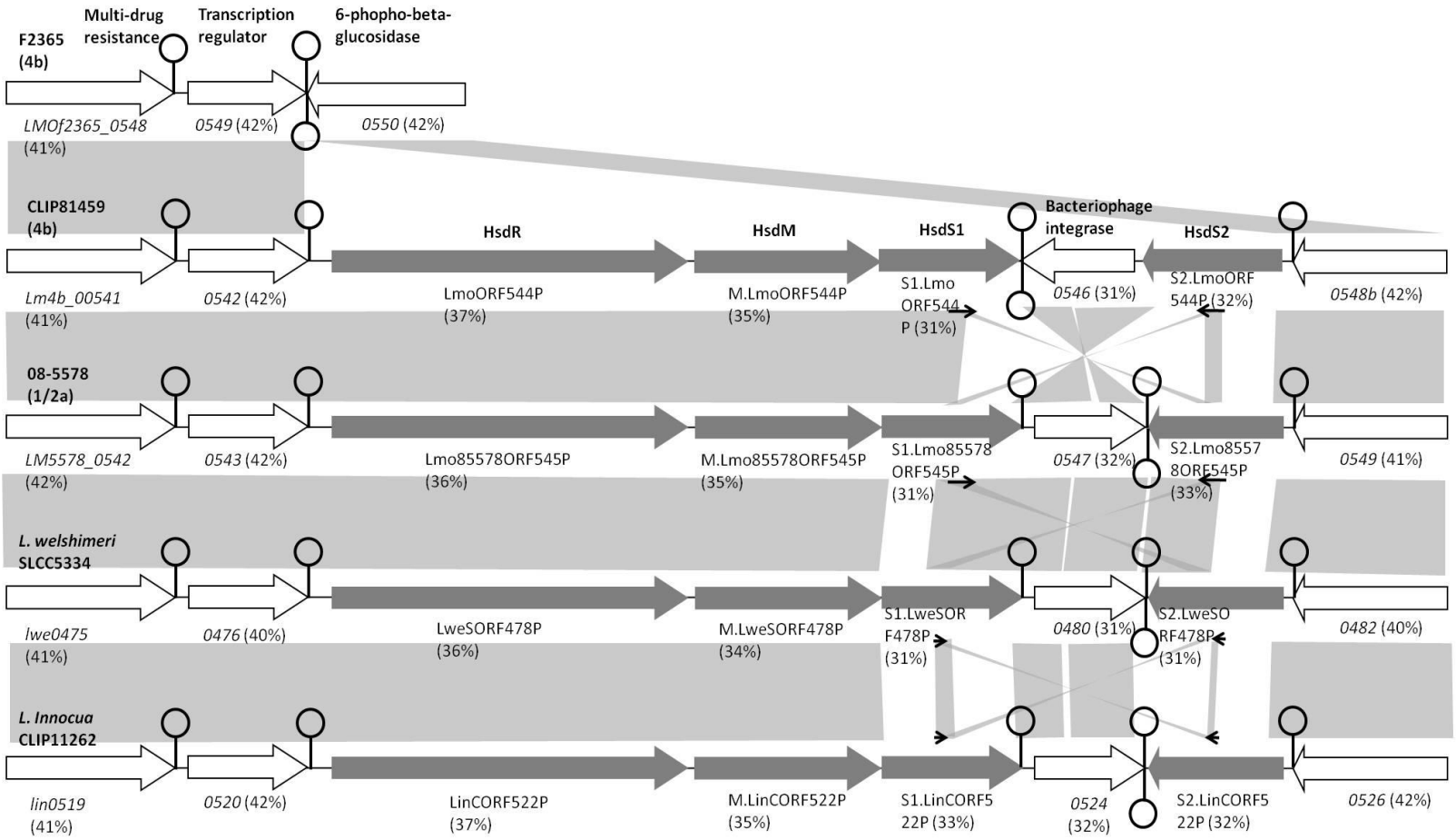

Flanking region
Up_M.Lse3954ORF279P
Up_M.LmoORF679P
Dn_M.LmoHORF40P
Up_LmoORF544P
Up_LinCORF522P
Up_Lmo85923ORF544P
Up_Lmo85578ORF545P
Up_LweSORF478P
Up_M.LweSORF291P
Up_LmoHORF2754P
Up_Lmo85923McrBP
Up_Lmo85578McrBP
Up_LmoEMcrBP
Up_LmoF6McrBP
Up_LmoHMcrBP
Up_LmoFORF327P
Dn_M.LmoFORF327P
Up_Lmo10403McrBP
Up_Lse3954ORF1765P
Up_M.Lmo85923ORF1805P
Up_M.Lmo85578ORF1853P
Up_M.LinCORF1737P
RC_Up_M.Lmo23ORF2958P
Dn_M.Lmo85923ORF1805P
Dn_M.Lmo85578ORF1853P
Up_M.Lmo85923ORF1853P
Up_M.Lmo85578ORF1901P
Dn_LmoN1ORF1065P
Dn_M.LinCORF1737P
Up_M.LmoEORF2316P
Up_LmoF6ORF2379P
RC_Dn_M.Lmo23ORF2958P
Dn_M.LmoEORF2316P
Dn_M.LmoF6ORF2379P
Up_M.LmoN1ORF1065P
Up_LmoN1ORF1083P
Up_M.LmoEORF1119P
Up_LmoN3ORF1850P
Dn_M.Lmo85923ORF1853P
Dn_M.Lmo85578ORF1901P
Dn_LmoN1ORF1083P
Dn_M.Lse3954ORF1765P
Dn_M.LmoHORF2754P
Dn_S.LinCORF21P
Dn_LmoHORF56P
Dn_Lse3954McrBP
Dn_Lmo85923McrBP
Dn_Lmo85578McrBP
Dn_LmoN3McrB2P
Dn_LmoF6McrBP
Dn_LmoEMcrBP
Dn_LmoHMcrBP
Dn_Lmo10403McrBP
Dn_LweSORF291P
Dn_S2.LmoORF544P
Dn_S2.LinCORF522P
Dn_S2.Lmo85923ORF544P
Dn_S2.Lmo85578ORF545P
Dn_S2.LweSORF478P
Up_M.Lmo85923ORF1134P
Up_M.Lmo85578ORF1180P
RC_Up_M.Lmo23ORF2320P
Dn_M.Lmo85578ORF677P
Dn_M.LinCORF88P
RC_Up_M.Lmo23ORF1402P
Dn_LmoORF679P
Up_M.LinCORF1262P
Up_M.LmoB054ORFAP
Dn_M.LinCORF1262P
Dn_M.LmoB054ORFAP
Dn_M.LmoA006ORFAP
Dn_M.LmoA118I
Up_M.Lmo85578ORF677P
Up_M.LinCORF88P
Up_M.LmoA006ORFAP
Up_M.LmoA118I
Dn_LmoPSAORF38P
Up_M.LmoHORF40P
Dn_Lmo85923ORF1134P
Dn_Lmo85578ORF1180P
Up_M.LmoEORF470P
RC_Dn_Lmo23ORF2320P
Up_LinCORF21P
Dn_M.LinCORF29P
Dn_M.LmoEORF1119P
Up_M.LinCORF29P
Up_M.LmoHORF56P
Dn_M.LmoEORF470P
RC_Dn_M.Lmo23ORF1402P
Up_M.LmoPSAORF38P
Second ICR (upstream/ G1)
Region 85 (upstream/ G2 and G11)
Region 85 (downstream/ G2 and G11)
Second ICR (downstream/ G1)

**Figure 6.4 Genetic organizations and diversity in the additional immigration control region (ICR) (A) and inverted repeats in proximity of integrase genes (B, C and D).** In Figure A, RM system genes were colored in grey. Percentage following the gene represents the GC content of the gene of interest and the putative Rho-independent terminators were marked with the stem loop next to the gene which might be under control of the terminator. Shade depicts the homologous region identified through the DNA sequence comparison with WebACT. Arrows signify mirror repeats suggested by the WebACT comparison. The Figures B and C describe the upstream and downstream regions of the bacteriophage integrase gene. In these pictures, start codons were marked in light grey and stop codons were surrounded by boxes. Putative rho-independent terminators were marked in bold and reverse complementary nucleotides within these terminators were underlined. Inverted repeats were colored in dark grey. Figure D evidences the reverse complementary nature of the inverted repeats shown in Figures B and C.

**A.**


F2365 (4b)
Multi-drug resistance
Transcription regulator
6-phopho-beta-glucosidase
LMOf2365_0548 (41%)
0549 (42%)
0550 (42%)
CLIP81459 (4b)
HsdR
HsdM
HsdS1
Bacteriophage integrase
HsdS2
Lm4b_00541 (41%)
0542 (42%)
LmoORF544P (37%)
M.LmoORF544P (35%)
S1.LmoORF544P (31%)
0546 (31%)
S2.LmoORF544P (32%)
0548b (42%)
08-5578 (1/2a)
LM5578_0542 (42%)
0543 (42%)
Lmo85578ORF545P (36%)
M.Lmo85578ORF545P (35%)
S1.Lmo85578ORF545P (31%)
0547 (32%)
S2.Lmo85578ORF545P (33%)
0549 (41%)
L. welshimeri SLCC5334
lwe0475 (41%)
0476 (40%)
LweSORF478P (36%)
M.LweSORF478P (34%)
S1.LweSORF478P (31%)
0480 (31%)
S2.LweSORF478P (31%)
0482 (40%)
L. Innocua CLIP11262
lin0519 (41%)
0520 (42%)
LinCORF522P (37%)
M.LinCORF522P (35%)
S1.LinCORF522P (33%)
0524 (32%)
S2.LinCORF522P (32%)
0526 (42%)

**B.**

```
08-5578         TAC-CAGCAAATGATGAACAAATAAAAATTACAAACTTATTTCAAAAATTAGAATA-TAC 8592
SLCC5334        TAC-CAACGAATGAAGAGCAAATAAAAATTACTAACTTATTACGAAAATTAGAATA-CAC 8582
F2365           ------------------------------------------------------------
CLIP81459       -ATTCAACTTTATATGAACAAGAAAAAATAAGT-TTTTTTTTAGATGAGATAACTATCAT 8618
CLIP11262       TGCCCAATAGTA-ATGAACAACAAAAAATTGGAACTTTTTCCCGACAAATTGACGA-CTT 8568

08-5578         TTCAATACTTCATCAAAACCAAATAGAGATGCTAAAAAAGGTAAAAAAAGCTTATTTGCA 8652
SLCC5334        TTCTACATTTCATCAAAACAAAATAGAACGACTAAAAAAGTTAAAAAAAGCTTATTTGCA 8642
F2365           ------------------------------------------------------------
CLIP81459       TATAAATCTTCATCAAAATAAACTAAAGAAACTCTCTTCTCTAAAAAAAGCCTATTTACA 8678
CLIP11262       AATTGCCCTTCAGCAAAATAAACTTGAGAAACTCTCTTCTCTAAAAAGATTCTATCTAAA 8628

08-5578         AACTATGTTTATCTAATT-AAGAT---ATTTACACTTAA---------ACTAACTCCTAAG 8699
SLCC5334        AACTATGTTCATTTAATTAAAAC---ATTTACACTTAA---------ACTAACTCCTAAG 8690
F2365           ------------------------------------------------------------
CLIP81459       AAATATGTTTATATAATTTAATCCCCATCTATTCTTATGAGATGGGGATTATTTATCACA 8738
CLIP11262       GAACATGTTTATATAATTCAAT----ATTTACACTTAT---------ATCAACTCCTAAG 8675
                          hsdS1

08-5578         ATAATAATAAAAGGTGATTAACTCATCTATTTATTATCTTAGGAGGTTTTTTATGAAAGC 8759
SLCC5334        ATAATAATAAAAGGTGATCAACTCATCTATTTATTATCTTAGGAGGTTTTTTATGAAAGC 8750
F2365           ------------------------------------------------------------
CLIP81459       TTAACATTGATAAATGGCGTATAATTTTATC-ATTATCTT----GATTTTCCA---ATTC 8790
CLIP11262       ATAATAAACAAAGGTGATTAATTCATCTATT-ATTATCTTAGGAGGTTTTTTATGAAAGC 8734
                                                                  Bacteriophage
                                                                  integrase
```

**C.**

```
08-5578         TTATCAATGTTAATGTGATAAGAGAATCCCCCATC-----TATTAAAA--------CTAG 9710
SLCC5334        TTATCAATGTTAATGTGATAAGA-AATCCCCCATC-----TATTAAAA--------CCAG 9700
F2365           ------------------------------------------------------------
CLIP81459       CTAAAATAATAAATAGATGAGTT-AATCACCTTTTATTATTATCTTAGGAGTTAGTTTAA 9732
CLIP11262       TTATCAATGTTAATGTGATAATA-AATCCCCACTTAGCTTTATCTTAG--------CTAG 9689
Bacteriophage integrase

08-5578         ATGGGGG-ATTCTA-CTATATGAACATATTTTGTAAATAGGTCTTTTTTAGAGAAGAGAG 9768
SLCC5334        ATGGGGG-ATTTGAACTATATGAGCATATTTTGTAAATAGGTCTTTTTTAGAGAAGAGAG 9759
F2365           ------------------------------------------------------------
CLIP81459       GTGTAAATGTTTTAATTAAATGAACATAGTTTGCAAATAAGCTTTTTTGAGAATATCAAA 9792
CLIP11262       GTGGGGATTATT---TTATATAAACATATTTTGTAGATAAGCTTTTTTTAATTCTTTAAT 9746
                          hsdS2

08-5578         TTTATTTATTTTGTTTTTATGAAGAGTAATAGTATTGTCCAATCTTTGTAAGAAAGTTCC 9828
SLCC5334        TTTATTTATTTTGTTTTTATGAAGAGTAATAACATTCTCCAATCTTTGTAAGAAAGTTCC 9819
F2365           ------------------------------------------------------------
CLIP81459       TTTATTTAATTTATTTTGATAGAGAATAATGGTATTATCCAATT-GTTTAAAAATTGAAC 9851
CLIP11262       TCTACCTAGTTTATTTTGAAGAAGATTAATAGTAATATCTATTTTATTCAGAAATTCAGC 9806

08-5578         TA-TTTTTTTTTGTCCTTC-ATAACAAGGGAACTTAACCATCAATGTCATCATAACAT-- 9884
SLCC5334        TA-TTTTTTTTTGTTCTTC-ATAACAAGGAAACTTAACTATCAATTTCATCATAACAT-- 9875
```

```
F2365          ------------------------------------------------------------
CLIP81459      CAATTTTTTCTTGTTCTTCAAAGTCAGCAATATAAATTTTTTGAT----TTAAGAAATGC 9907
CLIP11262      TA-TTTTCTTTTGCTCCTCTA--TTGATGGTACTGATATTAAAATATCCTCAAGATACGA 9863
```

**D.**

```
08-5578: 86/127 (67%)
    8544 atgatgaacaaataaaaattacaaacttatttcaaaaattagaatatacttcaatacttcatcaaaaccaaatagagatgctaaaaaaggtaaaaaaagcttatttgcaaactatgtttatctaata 8670
         |||| | ||||| ||||||   ||  || || |||| ||| ||  |||||   |  |||||| ||||| ||||| | |  ||        |||||||  | ||||| |||| |||||| || || ||
    9848 tacttcctgttttttttatccttgaaagaatgtttctaacctgttatgataatgagaagtatttttgttttatttatttgagagaagagatttttttctggataaatgttttatacaagtatatcat 9722



SLCC5334: 91/130 (70%)
    8534 atgaagagcaaataaaaattactaac-ttattacgaaaattagaatacacttctacatttcatcaaaacaaaatagaacgactaaaaaagttaaaaaaagcttatttgcaaactatgttcatttaattaaaa 8664
         ||||||| |||| ||||| ||  ||| || |||| || ||| ||  |   |  |||  ||||| |||||||||| |   |||        |||||||  | ||||| |||| |||| |||| || || |||
    9839 tacttcttgttt-ttttttatccttgaaagaatgtttctaacctcttacaataatgagaagtatttttgttttatttatttgagagaagagatttttttctggataaatgttttatacgagtatatcaagttt 9709



CLIP81459: 88/125 (70%)
    8575 gaacaagaaaaaataagttttttttttagatgagataactatcattataaatcttcatcaaaataaactaaagaaactctcttctctaaaaaaagcctatttacaaaatatgtttatataatttaa 8699
         |||||||||||||  |||   ||||  |  |  |   ||  |  || | |||  |||||||||| |||| ||| |   |  ||| |||||||| ||||| |||| |||||| || ||||| ||
    9867 cttgttctttttttaaccaagttaaaaatttgttaacctattatggtaataagagatagttttatttaatttatttaaactataagagttttttcgaataaacgtttgatacaagtaaattaattt 9743



CLIP11262: 28/35 (80%)
    8610 taaaaagattctatctaaagaacatgtttatataa 8644
         |||||| |   |||||| | || ||||||||||||
    9736 atttttttcgaatagatgttttatacaaatatatt 9702
```

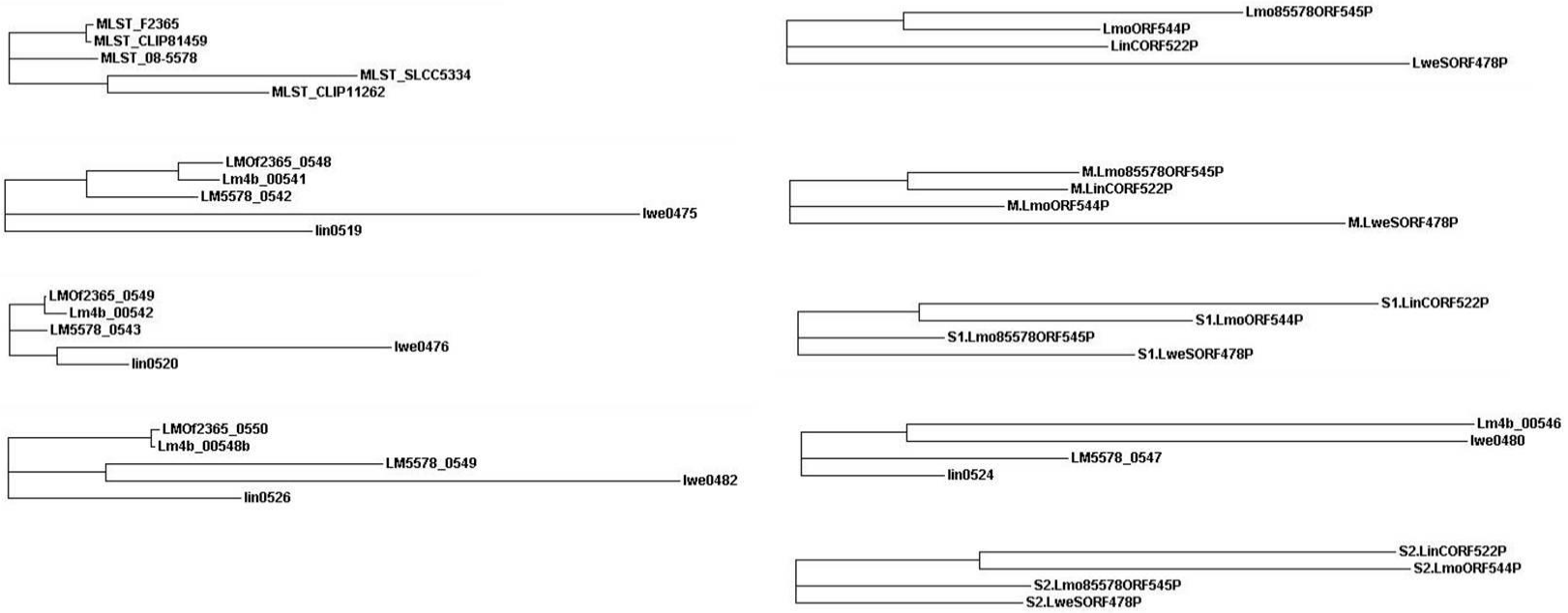



**Figure 6.5 Comparisons of the homologous genes within the second ICR.** MLST represents the concatenated target genes of the MLST implemented by Ragon et al. (38).

## REFERENCES


1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403-410.

2. **Broad Institute.** *Listeria monocytogenes* database <http://www.broadinstitute.org/annotation/genome/listeria_group/MultiHome.html>. Last accessed on Jan 31, 2011.

3. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics. **21:**3422-3423.

4. **Cheng, Y., R. M. Siletzky, and S. Kathariou.** 2008. Genomic Divisions/Lineages, Epidemic Clones, and Population Structure, p. 337-358. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

5. **Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins, and J. D. Thompson.** 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. **31:**3497-3500.

6. **den Bakker, H. C., X. Didelot, E. D. Fortes, K. K. Nightingale, and M. Wiedmann.** 2008. Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. BMC Evol. Biol. **8:**277.

7. **Dorscht, J., J. Klumpp, R. Bielmann, M. Schmelcher, Y. Born, M. Zimmer, R. Calendar, and M. J. Loessner.** 2009. Comparative genome analysis of *Listeria* bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. J. Bacteriol. **191:**7206-7215.

8. **Doumith, M., C. Buchrieser, P. Glaser, C. Jacquet, and P. Martin.** 2004. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. J. Clin. Microbiol. **42:**3819-3822.

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

10. **Furuta, Y., K. Abe, and I. Kobayashi.** 2010. Genome comparison and context analysis reveals putative mobile forms of restriction-modification systems and related rearrangements. Nucleic Acids Res. **38:**2428-2443.

11. **Gilmour, M. W., M. Graham, G. Van Domselaar, S. Tyler, H. Kent, K. M. Trout-Yakel, O. Larios, V. Allen, B. Lee, and C. Nadon.** 2010. High-throughput genome sequencing of two *Listeria monocytogenes* clinical isolates during a large foodborne outbreak. BMC Genomics. **11:**120.

12. **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:**849-852.

13. **Grkovic, S., M. H. Brown, and R. A. Skurray.** 2002. Regulation of bacterial drug export systems. Microbiol. Mol. Biol. Rev. **66:**671-701, table of contents.
14. **Hain, T., C. Steinweg, C. T. Kuenne, A. Billion, R. Ghai, S. S. Chatterjee, E. Domann, U. Karst, A. Goesmann, T. Bekel, D. Bartels, O. Kaiser, F. Meyer, A. Puhler, B. Weisshaar, J. Wehland, C. Liang, T. Dandekar, R. Lampidis, J. Kreft, W. Goebel, and T. Chakraborty.** 2006. Whole-genome sequence of *Listeria welshimeri* reveals common steps in genome reduction with *Listeria innocua* as compared to *Listeria monocytogenes*. J. Bacteriol. **188:**7405-7415.
15. **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:**3972-3979.
16. **Hofreuter, D., S. Odenbreit, and R. Haas.** 2001. Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. Mol. Microbiol. **41:**379-391.
17. **Hoskisson, P. A., and M. C. Smith.** 2007. Hypervariation and phase variation in the bacteriophage 'resistome'. Curr. Opin. Microbiol. **10:**396-400.
18. **Kathariou, S.** 2003. Foodborne outbreaks of listeriosis and epidemic-associated lineages of *Listeria monocytogenes*, p. 243-256. *In* M. E. Torrence and R. E. Isaacson (ed.),

Microbial food safety in animal agriculture-current topics, 1st ed. Iowa State Press, Ames, Iowa.

19. **Kathariou, S.** 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J. Food Prot. **65:**1811-1829.
20. **Kibbe, W. A.** 2007. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res. **35:**W43-6.
21. **Kim, J. W., R. M. Siletzky, and S. Kathariou.** 2008. Host ranges of *Listeria*-specific bacteriophages from the turkey processing plant environment in the United States. Appl. Environ. Microbiol. **74:**6623-6630.
22. **Kobayashi, I.** 2004. Restriction-modification systems as minimal forms of life. Nucleic Acids and Molecular Biology. **14:**19-62.
23. **Kobayashi, I.** 2001. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. Nucleic Acids Res. **29:**3742-3756.
24. **Kong, H., L. F. Lin, N. Porter, S. Stickel, D. Byrd, J. Posfai, and R. J. Roberts.** 2000. Functional analysis of putative restriction-modification system genes in the *Helicobacter pylori* J99 genome. Nucleic Acids Res. **28:**3216-3223.
25. **Kuenne, C., S. Voget, J. Pischimarov, S. Oehm, A. Goesmann, R. Daniel, T. Hain, and T. Chakraborty.** 2010. Comparative analysis of plasmids in the genus *Listeria*. PLoS One. **5:**e12511.
26. **Loessner, M. J., and C. E. D. Rees.** 2005. *Listeria* phages: basics and applications, p. 362-379. *In* K. M. Waldor, D. I. Friedman, and S. L. Adhya (ed.), Phages: their role in bacterial pathogenesis and biotechnology. ASM Press, Washington, DC.

27. **Loessner, M. J., R. B. Inman, P. Lauer, and R. Calendar.** 2000. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. Mol. Microbiol. **35:**324-340.

28. **Lorenz, M. G., and W. Wackernagel.** 1994. Bacterial gene transfer by natural genetic transformation in the environment. Microbiol. Rev. **58:**563-602.

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

30. **Meidanis, J., M. D. Braga, and S. Verjovski-Almeida.** 2002. Whole-genome analysis of transporters in the plant pathogen *Xylella fastidiosa*. Microbiol. Mol. Biol. Rev. **66:**272-299.

31. **Meinersmann, R. J., R. W. Phillips, M. Wiedmann, and M. E. Berrang.** 2004. Multilocus sequence typing of *Listeria* monocytogenes by use of hypervariable genes reveals clonal and recombination histories of three lineages. Appl. Environ. Microbiol. **70:**2193-2203.

32. **Murray, N. E.** 2002. 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. Microbiology. **148:**3-20.

33. **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:**2386-2395.

34. **Nikaido, H.** 2009. Multidrug resistance in bacteria. Annu. Rev. Biochem. **78:**119-146.
35. **Orsi, R. H., M. L. Borowsky, P. Lauer, S. K. Young, C. Nusbaum, J. E. Galagan, B. W. Birren, R. A. Ivy, Q. Sun, L. M. Graves, B. Swaminathan, and M. Wiedmann.** 2008. Short-term genome evolution of *Listeria monocytogenes* in a non-controlled environment. BMC Genomics. **9:**539.
36. **Orsi, R. H., Q. Sun, and M. Wiedmann.** 2008. Genome-wide analyses reveal lineage specific contributions of positive selection and recombination to the evolution of *Listeria monocytogenes*. BMC Evol. Biol. **8:**233.
37. **Piddock, L. J.** 2006. Multidrug-resistance efflux pumps - not just for resistance. Nat. Rev. Microbiol. **4:**629-636.
38. **Ragon, M., T. Wirth, F. Hollandt, R. Lavenir, M. Lecuit, A. Le Monnier, and S. Brisse.** 2008. A new perspective on *Listeria monocytogenes* evolution. PLoS Pathog. **4:**e1000146.
39. **Roberts, A., K. Nightingale, G. Jeffers, E. Fortes, J. M. Kongo, and M. Wiedmann.** 2006. Genetic and phenotypic characterization of *Listeria monocytogenes* lineage III. Microbiology. **152:**685-693.
40. **Roberts, R. J., M. Belfort, T. Bestor, A. S. Bhagwat, T. A. Bickle, J. Bitinaite, R. M. Blumenthal, S. K. Degtyarev, D. T. Dryden, K. Dybvig, K. Firman, E. S. Gromova,**

**R. I. Gumport, S. E. Halford, S. Hattman, J. Heitman, D. P. Hornby, A. Janulaitis, A. Jeltsch, J. Josephsen, A. Kiss, T. R. Klaenhammer, I. Kobayashi, H. Kong, D. H. Kruger, S. Lacks, M. G. Marinus, M. Miyahara, R. D. Morgan, N. E. Murray, V. Nagaraja, A. Piekarowicz, A. Pingoud, E. Raleigh, D. N. Rao, N. Reich, V. E. Repin, E. U. Selker, P. C. Shaw, D. C. Stein, B. L. Stoddard, W. Szybalski, T. A. Trautner, J. L. Van Etten, J. M. Vitor, G. G. Wilson, and S. Y. Xu.** 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. Nucleic Acids Res. **31:**1805-1812.

41. **Roberts, R. J., T. Vincze, J. Posfai, and D. Macelis.** 2010. REBASE--a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res. **38:**D234-6.
42. **Schmid, M. W., E. Y. Ng, R. Lampidis, M. Emmerth, M. Walcher, J. Kreft, W. Goebel, M. Wagner, and K. H. Schleifer.** 2005. Evolutionary history of the genus *Listeria* and its virulence genes. Syst. Appl. Microbiol. **28:**1-18.
43. **Senger, M., P. Rice, A. Bleasby, and M. Uludag.** 2009. Soaplab: Open Source Web Services Framework for Bioinformatics Programs. *In* The 10th Annual Bioinformatics Open Source Conference.
44. **Srikhanta, Y. N., K. L. Fox, and M. P. Jennings.** 2010. The phasevarion: phase variation of type III DNA methyltransferases controls coordinated switching in multiple genes. Nat. Rev. Microbiol. **8:**196-206.

45. **Steele, C. L., J. R. Donaldson, D. Paul, M. M. Banes, T. Arick, S. M. Bridges, and M. L. Lawrence.** 2011. Genome Sequence of Lineage III *Listeria monocytogenes* Strain HCC23. J. Bacteriol. **193:**3679-3680.

46. **Steinweg, C., C. T. Kuenne, A. Billion, M. A. Mraheil, E. Domann, R. Ghai, S. B. Barbuddhe, U. Karst, A. Goesmann, A. Puhler, B. Weisshaar, J. Wehland, R. Lampidis, J. Kreft, W. Goebel, T. Chakraborty, and T. Hain.** 2010. Complete genome sequence of *Listeria seeligeri*, a nonpathogenic member of the genus *Listeria*. J. Bacteriol. **192:**1473-1474.

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

48. **Thomas, C. M., and K. M. Nielsen.** 2005. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. Nat. Rev. Microbiol. **3:**711-721.

49. **Tock, M. R., and D. T. Dryden.** 2005. The biology of restriction and anti-restriction. Curr. Opin. Microbiol. **8:**466-472.

50. **U.S. Food and Drug Administration.** 2006. Food additives permitted for direct addition to food for human consumption; bacteriophage preparation. Fed. Regist. **71:**47729-47732.

51. **Verghese, B., M. Lok, J. Wen, V. Alessandria, Y. Chen, S. Kathariou, and S. Knabel.** 2011. *comK* prophage junction fragments as markers for *Listeria monocytogenes* genotypes unique to individual meat and poultry processing plants and a model for rapid niche-specific adaptation, biofilm formation, and persistence. Appl. Environ. Microbiol. **77:**3279-3292.

52. **Wagner, M., and J. MacLauchlin.** 2008. Biology, p. 3-25. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

53. **Ward, T. J., T. F. Ducey, T. Usgaard, K. A. Dunn, and J. P. Bielawski.** 2008. Multilocus genotyping assays for single nucleotide polymorphism-based subtyping of *Listeria monocytogenes* isolates. Appl. Environ. Microbiol. **74:**7629-7642.

54. **Ward, T. J., L. Gorski, M. K. Borucki, R. E. Mandrell, J. Hutchins, and K. Pupedis.** 2004. Intraspecific phylogeny and lineage group identification based on the *prfA* virulence gene cluster of *Listeria monocytogenes*. J. Bacteriol. **186:**4994-5002.

55. **Warriner, K., and A. Namvar.** 2009. What is the hysteria with *Listeria*? Trends in Food Science & Technology. **20:**245-254.

56. **Yildirim, S., D. Elhanafi, W. Lin, A. D. Hitchins, R. M. Siletzky, and S. Kathariou.** 2010. Conservation of genomic localization and sequence content of Sau3AI-like restriction-modification gene cassettes among *Listeria monocytogenes* epidemic clone I and selected strains of serotype 1/2a. Appl. Environ. Microbiol. **76:**5577-5584.

57. **Zimmer, M., E. Sattelberger, R. B. Inman, R. Calendar, and M. J. Loessner.** 2003. Genome and proteome of *Listeria monocytogenes* phage PSA: an unusual case for programmed + 1 translational frameshifting in structural protein synthesis. Mol. Microbiol. **50:**303-317.

# Chapter 7 Prevalence and Chronological Trends of *Listeria monocytogenes* Serotype 4b Clonal Groups


## ABSTRACT

*Listeria monocytogenes* is an etiological agent of a potentially severe foodborne illness (listeriosis) for humans and animals. Of 13 serotypes, serotype 4b has been involved with numerous listeriosis outbreaks and sporadic cases, and four epidemic-associated clonal groups (Epidemic Clones (ECs) I, II, Ia and V) of serotype 4b have been identified. Recently, three additional clonal groups were reported in serotype 4b. We employed a panel of DNA probes to analyze the prevalence of clonal groups among 278 serotype 4b *L. monocytogenes* isolates from clinical, environment/food, and animal sources. ECI, II and Ia were identified among these source categories at high frequencies, suggesting that these clonal groups might be environmental generalists that inhabit a wide range of environmental niches. Neither of these clonal groups was over-represented among clinical isolates as compared with non-clinical isolates, suggesting that human exposure rather than enhanced pathogenicity might be essential in causing listeriosis. Temporal trend analysis was conducted with serotype 4b human clinical isolates collected in the United States between 2003 to 2008. ECI and ECII were encountered each year of this period. Pronounced fluctuations were noted in prevalence of ECI and isolates belonging to neither of the aforementioned clonal groups during this period. In addition, we noted a trend towards increasing prevalence of ECIa between 2006-2008. Our findings could further our understanding of the complexities of the ecology and temporal dynamics of serotype 4b *L. monocytogenes*.

## INTRODUCTION

*Listeria monocytogenes* is a Gram-positive bacterium that can infect humans and animals via foods (15, 22). The disease caused by this bacterium (listeriosis) is manifested by severe symptoms including bacteremia, meningitis, and maternal-fetal infections resulting in stillbirths and abortions (8, 15, 22). In the United States, listeriosis is relatively infrequent (a mean of 1662 cases/year); however, its hospitalization and fatality rates (mean of 94% and 16%, respectively) were ranked first and third among foodborne illnesses (19). *L. monocytogenes* is ubiquitously present in the environment including food processing facilities (23) and is capable of tolerating a range of environmental stresses, including cold, acid, and osmotic stress (21). Hence, control of this bacterium has been challenging both for regulatory agencies and the food industry (23).

The population structure of *L. monocytogenes* consists of three lineages each of which is comprised of unique sets of serotypes and differs in epidemiological characteristics (2). Within Lineage I (serotypes 4b, 4d, 4e, 1/2b and 3b), serotype 4b is over-represented among clinical isolates (2). Lineage II (serotypes 1/2a, 3a, 1/2c, and 3c) is over-represented among food or environmental isolates; however, serotype 1/2a and 1/2c are often associated with clinical cases (2). Lineage III includes serotypes 4a, 4c, and certain strains of serotype 4b; and is frequently implicated with animal listeriosis (2, 26). Recently, a subset of lineage III was classified into a separate population lineage (linage IV) (24).

Over 95% of the clinical cases of human listeriosis are caused by serotypes 1/2a, 1/2b and 4b (22). In particular, serotype 4b has attracted a considerable attention since strains of this serotype are accountable for a majority of listeriosis outbreaks as well as many sporadic

cases (2). Comparison of outbreak strains revealed four major clonal groups of serotype 4b that were associated with multiple outbreaks that occurred in geographically distant regions and at different times. These clonal groups were designated Epidemic Clone (EC) I, II, Ia and V (2). Genome sequence data are currently available for strains F2365, H7858, and HPB2262 representing ECI, ECII and ECIa, respectively (5, 17). Recently, our laboratory discovered three novel clonal groups (Groups 1-3) that harbor a gene cassette (*lmo0734-lmo0739*) that had been regarded to be specific to lineage II (Chapter 2) (7).

To date, several DNA probes have been described for ECI and ECII (3, 11, 27), whereas highly specific DNA probes for ECIa, ECV and novel clonal groups have not been described yet. However, a combination of previously characterized ECI and ECII probes was suggested to identify ECIa and ECV (3). Furthermore, DNA probes that can differentiate ECIa from other serotype 4b strains have been identified through characterization of genetic diversity in Region 85, a genomic region of *L. monocytogenes* that frequently harbors restriction-modification (RM) systems (described in detail in Chapter 3). Also, Groups 1-3 revealed unique hybridization patterns for certain ECI- and ECII-specific probes and Region 85-based probes as well as a DNA probe originating from the *lmo0734-lmo0739* cassette. Information on ecology and temporal dynamics of serotype 4b will provide an insight into characteristics of these groups and facilitate establishing effective control measures against *L. monocytogenes*. Several studies relied on various subtyping tools such as multilocus sequence typing (MLST), genetic markers for ECI and ECII, and multilocus genotyping (MLGT) to address these aspects (5, 12, 18, 25, 27); nonetheless, only limited knowledge on these topics is currently available. In this study, we attempted to elucidate distribution of

serotype 4b clonal groups from different sources with a panel of DNA probes that reflect a recent advance in the genetic markers to discriminate the epidemic clonal groups as well as the novel clonal groups. We also analyzed temporal trends in clonal group prevalence among clinical serotype 4b isolates collected in the United States during six consecutive years (2003-2008).

## MATERIALS AND METHODS

**Bacterial isolates.** A total of 278 serotype 4b *L. monocytogenes* isolates were selected from the strain collection of our laboratory (Table 1). The isolates included in this study were composed of 178 clinical isolates; 77 isolates from foods and food processing facilities; 16 animal isolates; and 7 isolates of unknown origin. In particular, 142 clinical isolates from the Centers for Disease Control and Prevention (CDC) represent clinical isolates from epidemiologically unrelated cases that occurred in the US between 2003 to 2008, and included 19 to 27 isolates from each year, with each isolate typically originating from a different state. Bacteria were routinely cultured overnight in brain heart infusion (BHI; Becton, Dickinson and Company, Sparks, MD) media or on BHI solidified with 1.2% agar (Becton, Dickinson and Company).

**Genomic DNA extraction and Polymerase Chain Reaction (PCR).** Genomic DNA was extracted from overnight bacterial cultures using the DNeasy Blood and Tissue kit (QIAGEN, Valencia, CA). PCR amplification included an initial denaturation at 95ºC for 5 min, 31 cycles (each 95ºC for 1 min, 52ºC for 1 min, and 72ºC for 1 min) and a final extension of 72ºC for 10 min. PCR used exTaq (Takara, Otsu, Shiga, Japan) and a

thermocycler (Biometra, Goettingen, Germany). Annealing temperatures and extension times for repetitive cycles were determined by melting temperatures of the primer pair and anticipated PCR amplicon size, respectively. After PCR, the products were loaded for visualization onto the 1XTris/Borate/EDTA gel containing 0.8% agarose (Apex BioResearch Products, Research Triangle Park, NC). PCR primers (Eurofins MWG Operon, Huntsville, AL) are listed in Table 2.

**Enzymatic digestion of the genomic DNA.** Genomic DNA was restricted with MboI and Sau3AI (New England Biolabs, Ipswich, MA) following the protocol supplied by the manufacturer. The digested genomic DNA was run on the 0.8% agarose gel in parallel with intact genomic DNA.

**DNA-DNA dot blots.** DNA probes used in this study are listed in Table 2 and were prepared by gel-purifying PCR amplicons with the QIAquick gel extraction kit (QIAGEN), denaturing in boiling water for 10 min, and labeling with a digioxigenin (DIG) labeling kit (Roche, Florence, SC). The probes were hybridized with genomic DNA immobilized on nylon membranes (GE Water &Process Technologies; Trevose, PA) as described in Chapter 4.

**Pulse-Field Gel Electrophoresis (PFGE) and multilocus genotyping (MLGT).** PFGE was performed with AscI (New England Biolabs) and ApaI (Roche) as described by Graves et al. (13) and the results were analyzed with BioNumerics (Applied Maths, Austin, TX). MLGT was conducted as described (9, 24).

**Statistical analysis.** Fisher's exact test was performed with SAS (SAS Institute Inc., Cary, NC) to compare the distribution of each clonal group among different sources (animal, clinical and environment/food) and among different isolation years.

## RESULTS

**Clonal group determination.** Clonal groups were designated mainly based on hybridization with nine DNA probes that could discriminate among known serotype 4b clonal groups although the clonal groups of some environment/food isolates were determined by a few highly selective DNA probes (85M, Non-ECIIC-WAP, Hsp001219, LMSG_01573, McrB and lmo0737) (Table 1). For the longitudinal study, additional DNA probes based on the Immigration Control Region (Region 85) were employed to diversity of the isolates (Table 5). Each clonal group displayed a unique hybridization profile consistent to previously identified representative strains; however, one exception was observed in ECI strain OLM 65, which was positive for probes LMSG_01573 and McrB but displayed otherwise the same profile as other ECI strains (Table 1). Also, the PFGE and MLGT assay supported that OLM 65 belonged to ECI (data not shown). MboI and Sau3AI digestion results were included in the analysis as auxiliary means since certain clonal groups were characteristically resistant to either of these enzymes (Table 1). For instance, ECI was resistant to Sau3AI and susceptible to MboI whereas ECV and Group 3displayed the opposite (Chapter 3, (10)).

**Distribution of serotype 4b clonal groups among different sources.** To study the ecology of serotype 4b clonal groups, we compared the frequency of each group among a total of 278 isolates from various sources. Among these isolates, ECI was the most common

clonal group and was represented by 77 isolates (77/278, 28%) (Table 3) and the ECI prevalence was significantly higher than ECII or ECIa ($p=0.0353$ and 0.0052 in comparison with ECII and ECIa, respectively). Within each source category, ECI accounted for 19% of animal isolates (3/16), 29% of human clinical isolates (51/178) and 25% of environment/food isolates (19/77) (Table 3). Statistical analysis found no significant difference in frequency of ECI depending on sources ($p=0.6568$). Also, the prevalence of ECI did not differ significantly in the comparison between clinical vs. non-clinical isolates (isolates of animals and environment/food) ($p=0.4710$).

ECII and Ia were represented at the second and third highest frequency among the isolates analyzed in this study (Table 3). Statistical analysis revealed that their overall percentages are similar to each other ($p=0.4817$) and no animal isolates belonged to ECII or ECIa (Table 3). Specifically, a total of 53 ECII isolates (53/278, 19%) were identified, of which 34 were of human clinical origin (19%, 34/178) and 19 were environment/food isolates (25%, 19/77) (Table 3). Statistical analysis suggested that the prevalence of ECII strains did not differ significantly between clinical vs. non-clinical isolates ($p=0.8720$) and between clinical vs. environment/food isolates ($p=0.3179$). Forty six ECIa isolates (46/278, 17%) included 16 human clinical isolates (16/178, 9% of clinical isolates) and 27 isolates of environment/food origin (27/77, 35% of environment/food isolates) (Table 3). In the statistical analysis, the prevalence of ECIa was over-represented in non-clinical isolates as compared with clinical isolates ($p=3.80*10^{-5}$) and in environment/food isolates as compared with clinical isolates ($p=1.92*10^{-6}$).

The other clonal groups were identified in small numbers (4-9 isolates, 1-3% of the total isolates) with the hybridization data (Table 3). Clonal groups ECV, Group 1, and Group 2 were only identified within clinical isolates whereas Group 3 included both clinical (5/178, 3% of clinical isolates) and environment/food isolates (4/77, 5% of environment/food isolates) (Table 3). Neither of these minor clonal groups was significantly over-represented among clinical isolates compared with non-clinical isolates (p=0.3023, 0.0995, 0.0995, and 0.4983 for ECV and Groups 1-3, respectively). The prevalence of Group 3 isolates was similar when clinical isolates were compared with environment/food isolates (p=0.4597).

Seventy five isolates (75/278, 27%) belonged to no known serotype 4b clonal group (Table 3). These non-clonal group isolates accounted for 81% of animal isolates (13/16), 30% of clinical isolates (54/178) and 10% of environment/food isolates (8/77) (Table 3). Statistical analysis indicated that non-clonal group isolates were over-represented among animal isolates in comparison with the other two source categories ($p=9.11*10^{-5}$ for clinical isolates and $p=3.27*10^{-8}$ for environment/food isolates). The non-clonal group isolates were significantly more often represented among clinical isolates in comparison with the environment/food isolates ($p=4.50*10^{-4}$). Thirty six of the non-clonal group isolates were subjected to MLGT analysis, which designated 17 and 1 isolates into lineages III and IV, respectively (Table 1). These confirmed lineage III and IV isolates consisted of 13 animal isolates, 4 clinical isolates and one environment/food isolate (Table 1). It is noteworthy that all the non-clonal group isolates of animal origin belonged to lineages III or IV (Table 1). Also, several isolates exhibited the same MLGT types (1.5_4b [n=4], 1.62_4b [n=2], 1.7_4b [n=2], 1.70_4b [n=3], Lm3.1 [n=2] and Lm3.42 [n=4]) (Table 1).

**Temporal changes in the prevalence.** Temporal dynamics of the clonal groups were investigated with a randomly chosen panel of 142 serotype 4b isolates from human listeriosis cases in different states of the United States between 2003 and 2008 (Table 4). All the aforementioned clonal groups were identified among these isolates, with the notable exception of Group 3. ECI, II and IV were the three most predominant clonal groups (Tables 5), similarly to the results from the analysis of the entire panel of serotype 4b isolates as described above (Table 3).

During the period examined, only ECs I and II were consistently identified each year although their annual prevalence differed from year to year (Table 5). ECI represented 58% of isolates in 2003 but its prevalence decreased to 26% in 2004 (Table 5). This clonal group was identified at the prevalence of 14-35% between 2005 to 2007 but increased again to 48% in 2008 (Table 5). On the other hand, ECII accounted for only 16% of the 2003 isolates but its prevalence increased to 26% in 2004 and remained at lower levels (4 -20%) from 2004 to 2008 (Table 5). The observed prevalence shifts were significant for ECI ($p=0.0262$) but not for ECII ($p=0.3428$). ECIa isolates were encountered each year except for 2003, and seemed to increase in prevalence between 2003 to 2008, especially for the last three years (Table 5), but the yearly changes were not statistically significant ($p=0.3760$). With the exception of the year of 2005, isolates belonging to either the other clonal groups (ECV, Group 1and Group 2) were identified during the period analyzed at the annual prevalence of 4-11%.

Among the isolates subjected to temporal change analysis, 32% of the isolates (46/142) did not belong to any of the known serotype 4b clonal groups (Table 5). These isolates include the three lineage III isolates, which belonged to the same MLGT type

(Lm3.42) (Table 4). In the hybridizations with all the DNA probes used in this study, the non-clonal isolates exhibited 4-6 profiles and the same hybridization pattern did not always correspond to the same MLGT and vice versa. These non-clonal group isolates usually represented 20-30% of the isolates each year; however, in 2005 and 2007, their prevalence was 48 and 50%, respectively (Table 5), even though the temporal changes were not statistically significant (p=0.1037).

## DISCUSSION

In this study we examined the distribution of known clonal groups among serotype 4b isolates of different sources. Among all the clonal groups analyzed, ECI was ranked the first in terms of its frequency and was represented among all the source categories at comparable percentages. This agrees with previous studies reporting that a high percentage of serotype 4b of various origins belong to this clonal group (5, 12, 27). Using ECI and ECII-specific markers, Franciosa et al. found that ECI accounted for 16 (18%) of 89 serotype 4b *L. monocytogenes* isolates from clinical, food, and environmental sources in Europe (12). Prevalence of this clonal group was similar among clinical, food or environmental isolates (18, 18 and 20%, respectively) (12). In a collection of 195 *L. monocytogenes* (40 serotype 4b isolates and 21 isolates of unknown serotype belonging to lineages I and III) which were characterized with MLST by den Bakker et al. (4), 23 ECI isolates were identified representing a variety of sources (5). The prevalence of ECI amounted to 38% within serotype 4b isolates (15/40), the highest among the serotype 4b clonal groups examined by den Bakker et al. (4, 5). Surveys of the food/environment isolates also suggested that ECI is a

major clonal group in the L. monocytogenes isolates of food/environment origins (25, 27). Yildirim et al. employed ECI-specific markers for 34 serotype 4b isolates from food and food-processing facilities and classified 20 isolates (20/34, 59%) into ECI (27). Also, Ward et al. conducted MLGT-based analysis of 501 *L. monocytogenes* isolates from ready-to-eat meat (RTE) products and food-processing facilities and found 9 ECI isolates (26.5%, 9/34) (25).

ECII and ECIa constituted other dominant clonal groups after ECI among the serotype 4b isolates that we analyzed, and included isolates of clinical, food or environmental origin. In the survey with serotype 4b isolates from Europe, Franciosa et al. identified 15 ECII isolates (17%) consisting of seven clinical isolates, seven food isolates, and one environmental isolate (21, 14 and 20% within each source category) (12). In the MLST study, den Bakker et al. identified eight ECII isolates and five ECIa isolates among all the isolates examined. Seven ECII isolates and four ECIa isolates were confirmed to be serotype 4b (18% and 10% of the 40 serotype 4b isolates) (4, 5). These clonal groups were the second and third common among the serotype 4b isolates (4, 5). Origin of these isolates included clinical samples, food, environment and animals (5). Meanwhile, in the survey using MLGT of RTE products and food –processing facilities, Ward et al. identified ECIa most frequently among serotype 4b isolates (41.2%, 14/34). In his study, the prevalence of ECII isolates (26.5%, 9/34) was the same to that of ECI isolates (25).

In our study, ECI was the only known serotype 4b clonal group that included animal isolates and the other animal isolates belonged to lineages III or IV. This finding contrasts with the results obtained by den Bakker et al. that revealed two ECII strains, one from an

animal clinical case and the other from farm environments, in addition to four ECI animal isolates (5). Under-representation often results from an insufficient number of isolates; however, this does not seem to be likely in this case. In fact, the maximal number of serotype 4b animal isolates amounts to 11 in the study by den Bakker et al., which is fewer than the animal isolates examined by us (n=16), given that they examined seven serotype 4b isolates of animal-related origins and four animal isolates of unknown serotype belonging to lineages I and III (4). Rather, this discrepancy might potentially originate from the strain collections that differ in geographical and chronological characteristics. A majority of the *L. monocytogenes* isolates characterized by den Bakker et al. were collected from a confined region (New York state) between 1999 and 2003 (4).On the other hand, the isolates examined by us consist of isolates collected from a broader region encompassing multiple states of the US and other countries for an extended period of time. However, it deserves a particular attention that ECI was identified among animal isolates in both studies, strengthening the association between ECI and animal isolates. Indeed, the multilocus enzyme electrophoresis study by Boerlin et al. revealed that ECI was frequently presented among animal isolates (1).

Taken together, three most frequently identified clonal groups (ECs I, II and Ia) seem to belong to environmental generalists that can be encountered in a variety of environmental niches including foods and food processing facilities. This may be due to their abilities to endure stresses expected for such environments, for instance, temperature fluctuations, pH and salt. On the other hand, minor clonal groups except for Group 3 were composed only of clinical isolates. This phenomenon might be due to the rarity of these clonal groups and the dominance of clinical isolates in our samples; however, we still cannot exclude the

possibility that this finding might reflect a particularly high association between these clonal groups and clinical cases.

Two factors have been hypothesized to contribute to the association of *L. monocytogenes* epidemic clonal group with listeriosis: frequent human exposure and enhanced pathogenicity. It is expected that an epidemic clonal group of greater pathogenicity would be identified in disproportionately higher prevalence among clinical isolates than among non-clinical isolates. Our data revealed that none of the serotype 4b epidemic clonal groups was over-represented among clinical isolates, raising a possibility that none possess unique virulence factors to result in over-representation among clinical isolates. Instead, for major epidemic clonal groups, frequent human exposure seems to play a particularly important role in causing human listeriosis. Indeed, ECs II and IV were significantly over-represented among environment/food isolates compared with among clinical isolates.

Following a similar reasoning den Bakker et al. reached different conclusions, suggesting that ECIa was over-represented among clinical isolates whereas others (ECs I, II and V) were not (5). However, these findings could not be compared with ours since each analysis is based on different criteria. Den Bakker et al. included in their analysis isolates belonging to lineages I and II, i.e., isolates of serotype 4b and other serotypes (5); in contrast, all our isolates belonged to serotype 4b. Therefore, in the analysis by den Bakker et al., the prevalence based on source criteria was also affected by the source bias for *L. monocytogenes* serotypes. For instance, it is well known that serotype 4b is over-represented among clinical isolates but under-represented among environment/food isolates (2, 9). In fact, within the serotype 4b isolates that den Bakker et al. examined, the prevalence of ECIa

was 13% (4/30) among clinical isolates and 0% (0/10) among non-clinical isolates (4, 5). When these frequencies were analyzed with Fisher's exact test, we found no statistical evidence that ECIa was over-represented among serotype 4b clinical isolates in comparison with among serotype 4b non-clinical isolates ($p=0.5558$).

Our analysis also provided an insight into non-clonal group isolates of serotype 4b which accounted for 27% of the isolates that we examined. These non-clonal group isolates were most frequently represented among animal isolates and all the non-clonal group animal isolates belonged to lineages III and IV. Another intriguing finding is that the prevalence of non-clonal isolates was merely 10% among environment/food isolates, which was significantly lower than that among clinical isolates. This phenomenon indicates that a majority of serotype 4b environment/food isolates belong to known clonal groups whereas additional serotype 4b clonal groups might be present among clinical isolates as exemplified by ECV, Group 1, and Group 2. Some of such clonal groups might possess unique virulence factors that enhance pathogenicity or other attributes that diminish their fitness in foods or the environment. However, it should be emphasized that our serotype 4b strain panel included four times as many clinical isolates as environment/food isolates, suggesting that non-clonal group isolates are more likely to be discovered among clinical isolates than among environment/food isolates.

In this study we addressed the temporal dynamics of serotype 4b clonal groups among clinical serotype 4b isolates collected in the US between 2003-2008. ECI and II were consistently identified among the subset of isolates from each year, and ECIa was also identified each year except for 2003, with a trend towards increasing prevalence between

2006-2008. An earlier longitudinal study on *L. monocytogenes* was conducted by Ragon et al. by analyzing with MLST 126 maternal-fetal case isolates (of which 62 isolates belonged to serotype 4b) that were isolates mostly from France between 1987 and 2005 (18). Consistent to our findings, ECI isolates (22/62, 35% of the serotype 4b isolates) were continuously identified within this 19-year period except for 1993, 2001 and 2002 whereas ECIa (8/62, 13% of the serotype 4b isolates) increased in prevalence (18). However, unlike our results, ECII (4/62, 6% of the serotype 4b isolates) was identified only in three years (1998, 2003 and 2004) during the entire period examined by Ragon et al. (18).

Annual prevalence of a clonal group can be influenced by two main factors: (1) specific events (e.g. outbreaks) that occur in a particular year which may promote or suppress the proliferation of the clonal group, and contrariwise events (currently poorly understood) that may diminish prevalence; and (2) competition with other isolates. Our data on the ecology of serotype 4b clonal groups indicated that ECI, II and Ia generally predominate and are capable of colonizing diverse environmental niches. Thus, temporal dynamics of these clonal groups may be mainly attributable to environmental changes whose effects might be favorable to one clonal group or unfavorable to another. For instance, certain environmental changes in 2004 might have repressed the growth of ECI in a specific reservoir but promoted the multiplication of ECs II and Ia in their respective (and currently unidentified) reservoirs. Similarly, environmental changes in 2008 may have increased the prevalence of ECI and ECIa but reduced that of ECII. However, environmental changes that may affect prevalence of a certain clonal group remain to be identified, and, as mentioned above, we currently lack information on reservoirs that may be clone-specific.

On the other hand, competition with the aforementioned major clonal groups might be particularly instrumental in explaining the temporal changes of other strains. To illustrate, decreases in the major clonal groups might provide advantages for minor clonal groups, such as more nutrients and availability of environmental niches, whereas increases would do the opposite. The observed increase in prevalence of non-clonal group isolates in 2005 and 2007 may have been due to favorable environmental factors and/or reduced competition with the major epidemic clonal groups of serotype 4b (ECI, II and Ia).

In summary, our study provided evidence that three major clonal groups (ECI, II and Ia) of serotype 4b are environmental generalists. Also, our temporal trend analysis of the clinical isolates revealed complex dynamics of the serotype 4b clonal groups in annual prevalence for the six year period examined. Further analysis of more ecologically representative serotype 4b *L. monocytogenes* will be needed in order to accurately characterize the ecology of clonal groups of serotype 4b. In accompany with such analyses, larger-scale longitudinal studies including isolates from clinical cases, foods, the environment and animals will provide data needed for implementing effective control measures and assessing the public health risk associated with *L. monocytogenes*.

**Table7.1 List of the isolates analyzed in this study and clonal group classification**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| OLM 126* | 1959 | Canada | A | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| OLM 20 | 1937 | USA | A | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| OLM 65 | 1952 | Japan | A | ECI | 1.13_4b_Sw87_EC1 | - | + | ■ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | |
| 2008-114 | 2008 | NC | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| 2008-451 | 2008 | NC | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| 2010-0055 | 2010 | NC | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| 2010-0072A | 2010 | NC | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| J2213 | 2003 | AZ | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2269 | 2003 | GA | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2275 | 2003 | PA | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2282 | 2003 | MD | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2288 | 2003 | TX | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2302 | 2003 | CA | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2313 | 2003 | TX | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2327 | 2003 | MI | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2353 | 2003 | IL | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2391 | 2003 | TX | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2584 | 2003 | VT | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (20) |
| J2854 | 2004 | AZ | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| J2985 | 2004 | IL | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3082 | 2004 | GA | C | ECI | | - | + | P | | P | N[e] | N | N | N | | P | |
| J3106 | 2004 | NY | C | ECI | | - | + | | | | | | | | | | |
| J3133 | 2004 | TX | C | ECI | | - | + | | | | | | | | | | |
| J3180 | 2004 | CO | C | ECI | | - | + | P | | P | N | N | N | N | | P | |
| J3232 | 2004 | OK | C | ECI | | - | + | | | | | | | | | | |
| J3410 | 2005 | SC | C | ECI | | - | + | | | | | | | | | | |
| J3535 | 2005 | RI | C | ECI | | - | + | | | | | | | | | | |
| J3559 | 2005 | GA | C | ECI | | - | + | | | | | | | | | | |
| J3592 | 2005 | ME | C | ECI | | - | + | | | | | | | | | | |
| J3709 | 2005 | NH | C | ECI | | - | + | | | | | | | | | | |
| J3799 | 2005 | CT | C | ECI | | - | + | | | | | | | | | | |
| J3916 | 2006 | NM | C | ECI | | - | + | | | | | | | | | | |
| J4001 | 2006 | TX | C | ECI | | - | + | | | | | | | | | | |
| J4099 | 2006 | VA | C | ECI | | - | + | | | | | | | | | | |
| J4116 | 2006 | ME | C | ECI | | - | + | | | | | | | | | | |
| J4187 | 2006 | WI | C | ECI | | - | + | | | | | | | | | | |
| J4253 | 2006 | TN | C | ECI | | - | + | | | | | | | | | | |
| J4274 | 2006 | NH | C | ECI | | - | + | | | | | | | | | | |
| J4297 | 2006 | PA | C | ECI | | - | + | | | | | | | | | | |
| J4316 | 2006 | IA | C | ECI | | - | + | | | | | | | | | | |
| J4429 | 2007 | OR | C | ECI | | - | + | | | | | | | | | P | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J4600 | 2007 | OK | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J4685 | 2007 | MO | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J4950 | 2008 | WI | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J4977 | 2008 | NC | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J4979 | 2008 | TX | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5043 | 2008 | MA | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5080 | 2008 | NM | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5095 | 2008 | MD | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5136 | 2008 | SC | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5202 | 2008 | MS | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5354 | 2008 | UT | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5392 | 2008 | KY | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| J5478 | 2008 | IL | C | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | P | |
| AT-16 | 1988 | N/A[f] | E/F | ECI | 1.13_4b_Sw87_EC1 | - | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | | |
| F6812 | 1988 | N/A | E/F | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | | |
| FDA 34 | 1988 | N/A | E/F | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| FDA 5 | 1986 | N/A | E/F | ECI | | - | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | | |
| FDA 93 | 1986 | N/A | E/F | ECI | | - | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | | |
| FDA 96 | 1988 | N/A | E/F | ECI | | - | | ■ | ■ | N | □ | N | N | □ | □ | P | |
| FDA 97 | 1986 | N/A | E/F | ECI | | - | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | | |
| G3990 | 1985 | CA | E/F | ECI | Lm1.11 | - | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | | (9, 17) |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L0228 | 2003 | VA | E/F | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (10) |
| L0328 | 2003 | VA | E/F | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (10) |
| L0616 | 2004 | VA | E/F | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (10) |
| L0617 | 2004 | VA | E/F | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | (10) |
| LW-A1 | 2000 | N/A | E/F | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| LW-A32 | 2003 | N/A | E/F | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| LW-A4 | 2001 | N/A | E/F | ECI | | - | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| LW-A45 | 2001 | N/A | E/F | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| LW-A46 | 2001 | N/A | E/F | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| LW-A6 | 2001 | N/A | E/F | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| LW-A69 | 2001 | N/A | E/F | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| LW-A77 | N/A | N/A | N/A | ECI | | - | + | ■ | ■ | N | □ | N | N | □ | □ | P | |
| NC0064408a-3 | N/A | NC | N/A | ECI | | N | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| NC0064408a-5 | N/A | NC | N/A | ECI | | N | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| NC0065706 48h H+ | N/A | NC | N/A | ECI | | N | N | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | |
| 2009-0079 | 2009 | NC | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| 2009-0117 | 2009 | NC | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| H3396 | 1997 | CA | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| H6383 | 1996 | OH | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| H7550 | 1998 | USA | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20, 26) |
| H7738 | 1998 | USA | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H7762 | 1998 | USA | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J1815 | 2002 | N/A | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J1925 | 2002 | N/A | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J2206 | 2003 | NJ | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J2230 | 2003 | MA | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J2446 | 2003 | OH | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J2685 | 2004 | NY | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J3006 | 2004 | TX | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J3033 | 2004 | IL | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J3075 | 2004 | CT | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3170 | 2004 | MI | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3200 | 2004 | CT | C | ECII | | P | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | (20) |
| J3333 | 2004 | MN | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3527 | 2005 | WI | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3565 | 2005 | IL | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3606 | 2005 | MI | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3785 | 2005 | MA | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3840 | 2005 | MD | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J3881 | 2006 | CO | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J4101 | 2006 | WV | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |
| J4105 | 2006 | AL | C | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | ■ | □ | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J4120 | 2006 | OH | C | ECII | | + | + | | | | | | | | | | |
| J4210 | 2006 | GA | C | ECII | | + | + | | | | | | | | | | |
| J4485 | 2007 | MA | C | ECII | | + | + | | | | | | | | | P | |
| J4610 | 2007 | ME | C | ECII | | + | + | | | | | | | | | P | |
| J5172 | 2008 | MN | C | ECII | | + | + | | | | | | | | | P | |
| NRRL B-33589 | 2003 | MD | C | ECII | | + | + | | | | | | | | | | |
| NRRL B-33591 | 2003 | GA | C | ECII | | + | + | | | | | | | | | | |
| #8 MI 15 | 2005 | MI | E/F | ECII | | + | + | | | | | | | | | | |
| 34-2b | 2003 | NC | E/F | ECII | | + | + | | | | | | | | | | |
| 34-6a | 2003 | NC | E/F | ECII | | + | + | | | | | | | | | | (10) |
| Beef (franks) | 1992 | N/A | E/F | ECII | | + | N | | | | | | | | | | |
| L0226 | 2003 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0315 | 2003 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0327 | 2003 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0603 | 2004 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0623 | 2004 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0704 | 2004 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0719 | 2004 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0720 | 2004 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0724 | 2004 | VA | E/F | ECII | | + | + | | | | | | | | | | (10) |
| L0928 | 2004 | VA | E/F | ECII | | + | + | | | | | | | | | | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| L0929 | 2004 | VA | E/F | ECII | | + | + | | | | ■ | ■ | ■ | | ■ | | |
| L0930 | 2004 | VA | E/F | ECII | | + | + | | | | ■ | ■ | ■ | | ■ | | |
| LW-A105 | 2004 | N/A | E/F | ECII | | + | + | | | N | ■ | N | N | | ■ | P | |
| LW-A61 | 2003 | N/A | E/F | ECII | | + | + | | | N | ■ | N | N | | ■ | P | |
| LW-A86 | 2004 | N/A | E/F | ECII | | + | + | | | N | ■ | N | N | | ■ | P | |
| 2007-583 | 2007 | NC | C | ECV | | + | - | | | | | ■ | ■ | | ■ | | |
| J3215 | 2004 | NJ | C | ECV | | + | - | | | | | ■ | ■ | | ■ | | |
| J4045 | 2006 | MO | C | ECV | | + | - | | | | | ■ | ■ | | ■ | | |
| WS1 | 2000 | NC | C | ECV | | + | - | | | | | ■ | ■ | | ■ | | (16) |
| 2007-582 | 2007 | NC | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | | |
| 4b1 | 1962 | N/A | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | | |
| B068 | 1983 | MA | C | ECIa | Lm1.1 | N | N | | ■ | ■ | | | | ■ | | | (9) |
| J2967 | 2004 | CA | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | | |
| J3290 | 2004 | ME | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | | |
| J3419 | 2005 | CA | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | | |
| J3768 | 2005 | CO | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | | |
| J3921 | 2006 | CT | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | | |
| J4503 | 2007 | NYC, NY | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | P | |
| J4559 | 2007 | MA | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | P | |
| J4811 | 2007 | MA | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | P | |
| J4948 | 2008 | GA | C | ECIa | | + | + | | ■ | ■ | | | | ■ | | P | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J4954 | 2008 | CT | C | ECIa | | + | + | | | | | | | | | P | |
| J5074 | 2008 | WI | C | ECIa | | + | + | | | | | | | | | P | |
| J5375 | 2008 | WA | C | ECIa | | + | + | | | | | | | | | P | |
| OLM 120 | 1957 | Canada | C | ECIa | | + | + | | | | | | | | | | |
| 82a-2 | 2003 | NC | E/F | ECIa | | + | + | | | | | | | | | | |
| F7194 | 1989 | N/A | E/F | ECIa | | + | + | | | | | | | | | | |
| F7635 | 1989 | N/A | E/F | ECIa | | + | + | | | | | | | | | | |
| F7666 | 1989 | N/A | E/F | ECIa | | + | + | | | | | | | | | | |
| F7870 | 1989 | N/A | E/F | ECIa | | + | + | | | | | | | | | | |
| FDA 10 | 1998 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| FDA 100 | 1986 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| FDA 11 | 1987 | N/A | E/F | ECIa | | + | N | | | | | | | | | | |
| FDA 35 | 1988 | N/A | E/F | ECIa | | + | N | | | | | | | | | | |
| GA 311 | N/A | N/A | E/F | ECIa | | + | + | | | | | | | | | | |
| LW-A100 | 2005 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A101 | 2005 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A102 | 2005 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A103 | 2005 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A104 | 2005 | N/A | E/F | ECIa | | + | N | | | | | | | | | | |
| LW-A112 | 2006 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A130 | 2006 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LW-A84 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A85 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A87 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A88 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A89 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A90 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A91 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A92 | 2004 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A98 | 2005 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| LW-A99 | 2005 | N/A | E/F | ECIa | | + | + | | | N | | N | N | | | P | |
| BUG912 | N/A | N/A | N/A | ECIa | | + | + | | | | | | | | | | |
| BUG914 | N/A | N/A | N/A | ECIa | | + | + | | | | | | | | | | |
| BUG922 | N/A | N/A | N/A | ECIa | | + | + | | | | | | | | | | |
| 2007-0904 | 2007 | NC | C | Group 1 | 1.17_4b | + | + | | | | | | | | | | |
| J2255 | 2003 | GA | C | Group 1 | 1.17_4b | + | + | | | | | | | | | | (20) |
| J3139 | 2004 | WI | C | Group 1 | 1.17_4b | + | + | | | | | | | | | | |
| J3913 | 2006 | IN | C | Group 1 | 1.17_4b | + | + | | | | | | | | | | |
| J4016 | 2006 | SC | C | Group 1 | 1.17_4b | + | + | | | | | | | | | | |
| J4500 | 2007 | NE | C | Group 1 | 1.17_4b | + | + | | | | | | | | | | |
| J5000 | 2008 | VA | C | Group 1 | 1.17_4b | + | + | | | | | | | | | | |
| 2007-618 | 2007 | NC | C | Group 2 | 1.46_4b | + | + | | | | | | | | | | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3026 | 2004 | CT | C | Group 2 | 1.46_4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | ■ | ■ | |
| J3053 | 2004 | MI | C | Group 2 | 1.46_4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | ■ | ■ | |
| J3195 | 2004 | OH | C | Group 2 | 1.46_4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | ■ | ■ | |
| J4458 | 2007 | MN | C | Group 2 | 1.46_4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | ■ | ■ | |
| J4490 | 2007 | SC | C | Group 2 | 1.46_4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | ■ | ■ | |
| J4953 | 2008 | OH | C | Group 2 | 1.46_4b | + | + | □ | □ | ■ | □ | □ | ■ | ■ | ■ | ■ | |
| 2001-7R | 2001 | NC | C | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| 2001-8R | 2001 | NC | C | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| 2006-296 | 2006 | NC | C | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| 2007-454 | 2007 | NC | C | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| 2008-894 | 2008 | NC | C | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| 128b-1 | 2004 | NC | E/F | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | (10) |
| 18-2a | 2003 | NC | E/F | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| 363b-1 | 2005 | NC | E/F | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| 491a-5 | 2006 | NC | E/F | Group 3 | 1.60_4b | + | - | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | |
| J3085 | 2004 | MD | C | | | + | + | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | |
| J3155 | 2004 | IL | C | | 1.59_4b | + | + | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | |
| J3182 | 2004 | OR | C | | | + | + | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | |
| J3422 | 2005 | LA | C | | | + | + | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | |
| J3728 | 2005 | NY | C | | 1.45_4b | + | + | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | |
| J3751 | 2005 | OH | C | | | + | + | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3793 | 2005 | WV | C | | | + | + | | | | | | | | | | |
| J3976 | 2006 | MA | C | | 1.6_4b | + | + | | | | | | | | | | |
| J4179 | 2006 | RI | C | | | + | + | | | | | | | | | | |
| J4291 | 2006 | NY | C | | | + | + | | | | | | | | | | |
| J4432 | 2007 | PA | C | | | + | + | | | | | | | | | | |
| J4548 | 2007 | WI | C | | | + | + | | | | | | | | | P | |
| J4696 | 2007 | SD | C | | | + | + | | | | | | | | | | |
| J4713 | 2007 | RI | C | | | + | + | | | | | | | | | P | |
| J4726 | 2007 | FL | C | | | + | + | | | | | | | | | | |
| J5032 | 2008 | PA | C | | | + | + | | | | | | | | | | |
| J5421 | 2008 | FL | C | | | + | + | | | | | | | | | | |
| J5502 | 2008 | NM | C | | | + | + | | | | | | | | | P | |
| OLM 15 | 1934 | USA | C | | | + | + | | | | | | | | | | |
| OLM 18 | 1934 | USA | C | | | + | + | | | | | | | | | | |
| OLM 93 | 1954 | Canada | C | | | + | + | | | | | | | | | | |
| #5 MI 6-1 | 2005 | MI | E/F | | | + | + | | | | | | | | | | |
| BS-26 | N/A | N/A | E/F | | | + | + | | | N | | N | N | | | P | |
| LW-A125 | 2006 | N/A | E/F | | | + | + | | | N | | N | N | | | P | |
| NRRL B-33408 | 1997 | N/A | A | | Lm4.3 | + | + | | | | | | | | | | (24) |
| J3720 | 2005 | PA | C | | Lm3.42 | + | + | | | | | | | | | | |
| J2621 | 2003 | OR | C | | 1.5_4b | + | + | | | | | | | | | | (20) |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J3917 | 2006 | CA | C | | 1.5_4b | + | + | | | | | | | | | | |
| J4010 | 2006 | NE | C | | 1.5_4b | + | + | | | | | | | | | | |
| J4109 | 2006 | NJ | C | | 1.5_4b | + | + | | | | | | | | | | |
| J4460 | 2007 | GA | C | | | + | + | | | | | | | | | P | |
| J4875 | 2007 | NY | C | | | + | + | | | | | | | | | P | |
| LW-A13 | 2001 | N/A | E/F | | | + | + | | | N | | N | N | | | P | |
| FDA 107 | 1988 | N/A | E/F | | 1.63_4b | + | N | | | | | | | | | | |
| J2422 | 2003 | RI | C | | 1.7_4b | + | + | | | | | | | | | | (20) |
| J3070 | 2004 | TN | C | | 1.62_4b | + | + | | | | | | | | | | |
| J3115 | 2004 | VA | C | | 1.73_4b | + | + | | | | | | | | | | |
| J3120 | 2004 | WA | C | | | + | + | | | | | | | | | | |
| J3309 | 2005 | IA | C | | | + | + | | | | | | | | | | |
| J3415 | 2005 | TX | C | | 1.7_4b | + | + | | | | | | | | | | |
| J3554 | 2005 | MS | C | | | + | + | | | | | | | | | | |
| J3618 | 2005 | NJ | C | | | + | + | | | | | | | | | | |
| J3684 | 2005 | OR | C | | | + | + | | | | | | | | | | |
| J3762 | 2005 | IN | C | | | + | + | | | | | | | | | | |
| J3795 | 2005 | AZ | C | | | + | + | | | | | | | | | | |
| J3909 | 2006 | OR | C | | | + | + | | | | | | | | | | |
| J4428 | 2007 | OH | C | | 1.62_4b | + | + | | | | | | | | | | |
| J4434 | 2007 | TN | C | | | + | + | | | | | | | | | P | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J4435 | 2007 | WI | C | | | + | + | | | | | | | | | P | |
| J4465 | 2007 | CO | C | | | + | + | | | | | | | | | P | |
| J5038 | 2008 | MI | C | | | + | + | | | | | | | | | P | |
| NRRL B-33077 | N/A | N/A | A | | Lm3.1 | + | - | | | | | | | | | | (24) |
| OLM 8 | 1933 | USA | A | | Lm3.42 | + | + | | | | | | | | | | |
| NRRL B-33185 | N/A | N/A | A | | Lm3.9 | + | + | | | | | | | | | | (24) |
| NRRL B-33190 | N/A | N/A | A | | Lm3.12 | - | + | | | | | | | | | | (24) |
| OLM 39 | 1939 | IA | A | | Lm3.1 | - | + | | | | | | | | | | |
| NRRL B-33181 | N/A | N/A | A | | Lm3.5 | - | + | | | | | | | | | | (24) |
| J2571 | 2003 | KY | C | | Lm3.42 | - | + | | | | | | | | | | (20) |
| (HI) 267 | 1999 | N/A | C | | Lm3.49 | - | N | | | | | | | | | | |
| NRRL B-33105 | N/A | N/A | A | | Lm3.3 | + | - | | | | | | | | | | (24) |
| J2479 | 2003 | MI | C | | Lm3.42 | + | - | | | | | | | | | | (20) |
| NRRL B-33426 | N/A | N/A | E/F | | Lm3.41 | + | - | | | | | | | | | | (24) |
| NRRL B-33092 | N/A | N/A | A | | Lm3.2 | + | + | | | | | | | | | | (24) |
| NRRL B-33187 | N/A | N/A | A | | Lm3.10 | + | + | | | | | | | | | | (24) |
| NRRL B-33191 | N/A | N/A | A | | Lm3.13 | + | + | | | | | | | | | | (24) |
| OLM 37 | 1937 | USA | A | | Lm3.52 | + | + | | | | | | | | | | |
| OLM 115 | 1955 | Canada | C | | 1.70_4b | + | + | | | | | | | | | | |
| OLM 116 | 1956 | Canada | C | | 1.70_4b | + | + | | | | | | | | | | |
| OLM 80 | 1954 | Canada | C | | 1.70_4b | + | + | | | | | | | | | | |

**Table 7.1 Continued**

| Name | Isolation Year | Country/State | Source[a] | Clonal Group | MLGT | Sau3AI[b] | MboI[b] | 85M[c] | Non-ECIIC-WAP[c] | 4bSF18[c] | Hsp001219[c] | ECIIC-WAP[c] | H18RP9/10[c] | LMSG_01573[c] | McrB[c] | lmo0737[d] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| J4221 | 2006 | MD | C | | 1.75_4b | + | + | □ | □ | □ | □ | ■ | ■ | □ | □ | □ | |
| J5166 | 2008 | KS | C | | | + | + | □ | □ | □ | □ | ■ | ■ | □ | □ | P | |
| 2007-584 | 2007 | NC | C | | 1.74_4b | + | + | □ | □ | □ | □ | □ | ■ | □ | □ | □ | |
| LW-A109 | 2006 | N/A | E/F | | | + | + | □ | □ | N | □ | N | N | □ | □ | ■ | |
| NRRL B-33183 | N/A | N/A | A | | Lm3.7 | + | + | □ | □ | □ | □ | □ | □ | □ | □ | □ | (24) |
| LW-A93 | 2004 | N/A | E/F | | | + | + | □ | □ | N | □ | N | N | □ | □ | P | |

[a] A, animal isolates; C, clinical isolates; and E/F, environment/food isolates.

[b] Negative and positive signs represent resistance and susceptibility to the restriction endonuclease digestion, respectively. The letter P indicates partial digestion.

[c] White and black squares signify the absence and presence of the signal in the DNA-DNA hybridization. When the letter P is inside a square, the white and black squares represent the absence and presence of the PCR amplicon with the primers used to amplify a DNA probe.

[d] White and black squares with the letter P inside indicate whether a strain reveals the typical serotype 4b profile or the atypical serotype 4b profile with the addition *lmo0737* band in the multiplex PCR serotyping devised by Doumith et al. (6).

[e] N, Not tested.

**Table 7.1 Continued**

[f] N/A, Information not available.

**Table 7.2 Oligonucleotides and DNA probes used in this study**

| Probe | Primer | Sequence (5' to 3') | Target | Reference |
|---|---|---|---|---|
| **<Clonal group designation probes>** | | | | |
| 85M | F2365_85MF | AATATATTTTCAATGTTTGATGGT | *LMOf2365_0327* | (27) |
| | F2365_85MR | GCTAATTCAATCCCTATTCT | *LMOf2365_0327* | (27) |
| Non-ECIIC-WAP | ECIC-WAPF | ATGGAAATTGGGCATGGC | *LMOf2365_0450* | (3) |
| | ECIC-WAPR | GTAGTTCCAGTGGACATG | *LMOf2365_0450* | (3) |
| 4bSF18 | 4bSF18_F | ACGGGCGTTTTATATTAAATGGG | *LMOf2365_0466* | (11) |
| | 4bSF18_R | AATATCTCGAAAACTCCGAGT | *LMOf2365_0466* | (11) |
| Hsp001219 | Hspec01219F | GAGGCTATCGAAATTGCTCG | *LMOh7858_1168* | (3) |
| | Hspec01219R | AGGATTCGGAATTTCATCCA | *LMOh7858_1168* | (3) |
| ECIIC-WAP | ECIIC-WAPF | GGGAACTTTCCATTAGCC | *LMOh7858_0479* | (3) |
| | ECIIC-WAPR | TTAAATGGGATATGATGT | *LMOh7858_0479* | (3) |
| H18RP9/10 | H18RP9/10F | CTGGATTTGCAGCTTATGAT | *LMOh7858_0487* | (14) |
| | H18RP9/10R | CCTATTCTTTCCATAAGTAAT | *LMOh7858_0487* | (14) |
| LMSG_01573 | LMSG_01573F | TACAATTGGTCGGACACGTG | *LMSG_01573* | This study |
| | LMSG_01573R | AGAATCCGCTCATAAACAGC | *LMSG_01573* | This Study |
| McrB | H7858_0337F | ATATCCATGCCCATCACCAC | *LMOh7858_0337* | This study |
| | H7858_0337R | CGGGAGGAATCTCGTTATAC | *LMOh7858_0337* | This study |
| lmo0737 | lmo0737F | AGGGCTTCAAGGACTTACCC | *lmo0737* | (6) |
| | lmo0737R | ACGATTTCTGCTTGCCATTC | *lmo0737* | (6) |
| **<Additional Region 85 probes>** | | | | |
| J2479Res | J2479Res_F | GACCTAGTAAAGCAGGTGCT | *lmoJ2R* | This study |
| | J2479_3R | AGGTACCCATTCGATAGTCG | *lmoJ2R* | This study |
| J3115Met | J3115_2F | TATCAGGCTTTCCCTGTCAA | *lmoJ3M* | This study |
| | J3115Met_R | GGTACTACAACCGAATTCCC | *lmoJ3M* | This study |
| LMSG_01574C | LMSG_01574_CF | CTAGACAACACTATCGCTC | *LMSG_01574* | This Study |

**Table 7.2 Continued**

| Probe | Primer | Sequence (5' to 3') | Target | Reference |
|---|---|---|---|---|
| | LMSG_01574_CR | ACGCTGTTCCCAATCATCAG | *LMSG_01574* | This Study |
| LMSG_01574 | LMSG_01574F | CCTCTAATTCGAGTATCC | *LMSG_01574* | This Study |
| | LMSG_01574R | GTCCCCTACAGCAACACCTTT | *LMSG_01574* | This Study |
| LMSG_01578 | LMSG_01578F | GCTTGCAGGAAAGAAGGAAC | *LMSG_03073* | This Study |
| | LMSG_01578R | GGTTTTCTTTCGCAAGTGC | *LMSG_03073* | This Study |
| LP | LPF | GTTGTTGATAGTAAGTATA | *LMOf2365_0329* | This Study |
| | LPR | TCTGCTTTGACATCTATGTG | *LMOf2365_0329* | This Study |
| J3976TypeIS_V | J3976_TypeIS_VF | GCAAGTCTAAATCTACCAG | *LmoJ3976S* | This Study |
| | J3976_TypeIS_VR | ATGACTCTGCGGCAAGTAC | *LmoJ3976S* | This Study |
| J3976NG | J3976NG_F | GCAGCATTAGTAGTTTGTAGC | *J3976NG* | This Study |
| | J3976_01575LP_3R | TGAAGTCTATAGTCTCCGGC | *J3976NG* | This Study |
| 2007-640LP | 2007-640_LPF | GTTAGTAGGATGTGGTAAAG | *lin0331* | This Study |
| | 2007-640_LPR | CTGTTCTCTTTGTCATCCAC | *lin0331* | This Study |
| Mrr | LMSG_01576F | CAGAATGGGATATAGCGGAG | *LMSG_01576* | This Study |
| | LMSG_01576R | CAGCTTCCACTCCAACTTTG | *LMSG_01576* | This Study |
| J1-194_1 | J1-194_1 F | CGAGGTCTAGGACATTGTTC | Intergenic region between *LMBG_02168* and *LMBG_02169* | This Study |
| | J1-194_1 R | CATATGGTTACCGTCACTAG | *LMBG_02168* | This Study |

**Table 7.3 Clonal group prevalence among serotype 4b isolates from different sources**

| Clonal group | Animal (%) | Clinical (%) | Environment/foods (%) | Unknown (%) | Total (%) |
|---|---|---|---|---|---|
| ECI | 3 (19) | 51 (29) | 19 (25) | 4 (57) | 77 (28) |
| ECII | 0 | 34 (19) | 19 (25) | 0 (0) | 53 (19) |
| ECIa | 0 | 16 (9) | 27 (35) | 3 (43) | 46 (17) |
| Others | | | | | |
| ECV | 0 | 4 (2) | 0 (0) | 0 (0) | 4 (1) |
| Group 1 | 0 | 7 (4) | 0 (0) | 0 (0) | 7 (3) |
| Group 2 | 0 ( | 7 (4) | 0 (0) | 0 (0) | 7 (3) |
| Group 3 | 0 ( | 5 (3) | 4 (5) | 0 (0) | 9 (3) |
| Non-clonal | 13 (81) | 54 (30) | 8 (10) | 0 (0) | 75 (27) |
| Total | 16 | 178 | 77 | 7 | 278 |

## Table 7.4 List of the serotype 4b clinical isolates collected in the United States, 2003-2008

| Hybridization Pattern | Name | Isolation Year | Country/State | Clonal Group | MLGT | Sau3AI[a] | MboI[a] | 85M[b] | Non-ECIIC-WAP[b] | 4bSF18[b] | Hsp001219[b] | ECIIC-WAP[b] | H18RP9/10[b] | J2479Res[b] | J3115Met[b] | LMSG_01573[b] | LMSG_01574C[b] | LMSG_01574[b] | LMSG_01578[b] | LP[b] | J3976TypeIS_V[b] | J3976NG[b] | 2007-6401LP[b] | Mrr[b] | McrB[b] | J1-194_1[b] | lmo0737[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | J2213 | 2003 | AZ | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2269 | 2003 | GA | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2275 | 2003 | PA | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2282 | 2003 | MD | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2288 | 2003 | TX | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2302 | 2003 | CA | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2313 | 2003 | TX | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2327 | 2003 | MI | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2353 | 2003 | IL | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2391 | 2003 | TX | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2584 | 2003 | VT | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 1 | J2854 | 2004 | AZ | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | |
| 1 | J2985 | 2004 | IL | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | |
| 1 | J3082 | 2004 | GA | ECI | | - | + | P | | P | N[d] | N | N | N | N | N | N | N | N | P | N | N | N | N | | N | P | |
| 1 | J3106 | 2004 | NY | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | |
| 1 | J3133 | 2004 | TX | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | |
| 1 | J3180 | 2004 | CO | ECI | | - | + | P | | P | N | N | N | N | N | N | N | N | N | P | N | N | N | N | | N | P | |
| 1 | J3232 | 2004 | OK | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | |
| 1 | J3410 | 2005 | SC | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | |
| 1 | J3535 | 2005 | RI | ECI | | - | + | | | | | | | | | | | | | | | | | | | | | |

## Table 7.4 Continued

| Hybridization Pattern | Name | Isolation Year | Country/State | Clonal Group | MLGT | Sau3AI[a] | MboI[a] | 85M[b] | Non-ECIIC-WAP[b] | 4bSF18[b] | Hsp001219[b] | ECIIC-WAP[b] | H18RP9/10[b] | J2479Res[b] | J3115Met[b] | LMSG_01573[b] | LMSG_01574C[b] | LMSG_01574[b] | LMSG_01578[b] | LP[b] | J3976TypeIS_V[b] | J3976NG[b] | 2007-640LP[b] | Mn[b] | McrB[b] | J1-194_1[b] | lmo0737[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | J3559 | 2005 | GA | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J3592 | 2005 | ME | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J3709 | 2005 | NH | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J3799 | 2005 | CT | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J3916 | 2006 | NM | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4001 | 2006 | TX | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4099 | 2006 | VA | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4116 | 2006 | ME | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4187 | 2006 | WI | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4253 | 2006 | TN | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4274 | 2006 | NH | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4297 | 2006 | PA | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4316 | 2006 | IA | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4429 | 2007 | OR | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J4600 | 2007 | OK | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J4685 | 2007 | MO | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J4950 | 2008 | WI | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J4977 | 2008 | NC | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J4979 | 2008 | TX | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J5043 | 2008 | MA | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | |
| 1 | J5080 | 2008 | NM | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |

## Table 7.4 Continued

| Hybridization Pattern | Name | Isolation Year | Country/State | Clonal Group | MLGT | Sau3AI[a] | MboI[a] | 85M[b] | Non-ECIIC-WAP[b] | 4bSF18[b] | Hsp001219[b] | ECIIC-WAP[b] | H18RP9/10[b] | J2479Res[b] | J3115Met[b] | LMSG_01573[b] | LMSG_01574C[b] | LMSG_01574[b] | LMSG_01578[b] | LP[b] | J3976TypeIS_V[b] | J3976NG[b] | 2007-640LP[b] | Mn[b] | McrB[b] | J1-194_1[b] | lmo0737[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | J5095 | 2008 | MD | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J5136 | 2008 | SC | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J5202 | 2008 | MS | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J5354 | 2008 | UT | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J5392 | 2008 | KY | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 1 | J5478 | 2008 | IL | ECI | | - | + | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | P | |
| 2 | J2206 | 2003 | NJ | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | (20) |
| 2 | J2230 | 2003 | MA | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | (20) |
| 2 | J2446 | 2003 | OH | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | (20) |
| 2 | J2685 | 2004 | NY | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | (20) |
| 2 | J3006 | 2004 | TX | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | (20) |
| 2 | J3033 | 2004 | IL | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | (20) |
| 2 | J3075 | 2004 | CT | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |
| 2 | J3170 | 2004 | MI | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |
| 2 | J3200 | 2004 | CT | ECII | | P | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | (20) |
| 2 | J3333 | 2004 | MN | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |
| 2 | J3527 | 2005 | WI | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |
| 2 | J3565 | 2005 | IL | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |
| 2 | J3606 | 2005 | MI | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |
| 2 | J3785 | 2005 | MA | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |
| 2 | J3840 | 2005 | MD | ECII | | + | + | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | |

## Table 7.4 Continued

| Hybridization Pattern | Name | Isolation Year | Country/State | Clonal Group | MLGT | Sau3AI[a] | MboI[a] | 85M[b] | Non-ECIIC-WAP[b] | 4bSF18[b] | Hsp001219[b] | ECIIC-WAP[b] | H18RP9/10[b] | J2479Res[b] | J3115Met[b] | LMSG_01573[b] | LMSG_01574C[b] | LMSG_01574[b] | LMSG_01578[b] | LP[b] | J3976TypeIS_V[b] | J3976NG[b] | 2007-640LP[b] | Mn[b] | McrB[b] | J1-194_1[b] | lmo0737[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | J3881 | 2006 | CO | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 2 | J4101 | 2006 | WV | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 2 | J4105 | 2006 | AL | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 2 | J4120 | 2006 | OH | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 2 | J4210 | 2006 | GA | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 2 | J4485 | 2007 | MA | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 2 | J4610 | 2007 | ME | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | P |
| 2 | J5172 | 2008 | MN | ECII | | + | + | | | | | | | | | | | | | | | | | | | | | P |
| 3 | J3215 | 2004 | NJ | ECV | | + | - | | | | | | | | | | | | | | | | | | | | | |
| 3 | J4045 | 2006 | MO | ECV | | + | - | | | | | | | | | | | | | | | | | | | | | |
| 4 | J2967 | 2004 | CA | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 4 | J3290 | 2004 | ME | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 4 | J3419 | 2005 | CA | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 4 | J3768 | 2005 | CO | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 4 | J3921 | 2006 | CT | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 4 | J4503 | 2007 | NYC, NY | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | P |
| 4 | J4559 | 2007 | MA | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | P |
| 4 | J4811 | 2007 | MA | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | P |
| 4 | J4948 | 2008 | GA | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | P |
| 4 | J4954 | 2008 | CT | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | P |
| 4 | J5074 | 2008 | WI | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | | P |

## Table 7.4 Continued

| Hybridization Pattern | Name | Isolation Year | Country/State | Clonal Group | MLGT | Sau3AI[a] | MboI[a] | 85M[b] | Non-ECIIC-WAP[b] | 4bSF18[b] | Hsp001219[b] | ECIIC-WAP[b] | H18RP9/10[b] | J2479Res[b] | J3115Met[b] | LMSG_01573[b] | LMSG_01574C[b] | LMSG_01574[b] | LMSG_01578[b] | LP[b] | J3976TypeIS_V[b] | J3976NG[b] | 2007-640LP[b] | Mn[b] | McrB[b] | J1-194_1[b] | lmo0737[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 | J5375 | 2008 | WA | ECIa | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 5 | J3913 | 2006 | IN | Group 1 | 1.17_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 6 | J2255 | 2003 | GA | Group 1 | 1.17_4b | + | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 6 | J3139 | 2004 | WI | Group 1 | 1.17_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 6 | J4016 | 2006 | SC | Group 1 | 1.17_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 6 | J4500 | 2007 | NE | Group 1 | 1.17_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 6 | J5000 | 2008 | VA | Group 1 | 1.17_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 7 | J3026 | 2004 | CT | Group 2 | 1.46_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 7 | J3053 | 2004 | MI | Group 2 | 1.46_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 7 | J3195 | 2004 | OH | Group 2 | 1.46_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 7 | J4458 | 2007 | MN | Group 2 | 1.46_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 7 | J4490 | 2007 | SC | Group 2 | 1.46_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 7 | J4953 | 2008 | OH | Group 2 | 1.46_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 8 | J3085 | 2004 | MD | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 8 | J3728 | 2005 | NY | | 1.45_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 8 | J3793 | 2005 | WV | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 8 | J4291 | 2006 | NY | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 8 | J4548 | 2007 | WI | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 8 | J5502 | 2008 | NM | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 9 | J4713 | 2007 | RI | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 10 | J2621 | 2003 | OR | | 1.5_4b | + | + | | | | | | | | | | | | | | | | | | | | | (20) |

## Table 7.4 Continued

| Hybridization Pattern | Name | Isolation Year | Country/State | Clonal Group | MLGT | Sau3AI[a] | MboI[a] | 85M[b] | Non-ECIIC-WAP[b] | 4bSF18[b] | Hsp001219[b] | ECIIC-WAP[b] | H18RP9/10[b] | J2479Res[b] | J3115Met[b] | LMSG_01573[b] | LMSG_01574C[b] | LMSG_01574[b] | LMSG_01578[b] | LP[b] | J3976TypeIS_V[b] | J3976NG[b] | 2007-640LP[b] | Mn[b] | McrB[b] | J1-194_1[b] | lmo0737[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 10 | J3917 | 2006 | CA | | 1.5_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 10 | J4010 | 2006 | NE | | 1.5_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 10 | J4109 | 2006 | NJ | | 1.5_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 11 | J4460 | 2007 | GA | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 11 | J4875 | 2007 | NY | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 12 | J3155 | 2004 | IL | | 1.59_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 12 | J3422 | 2005 | LA | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 12 | J4432 | 2007 | PA | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 12 | J4696 | 2007 | SD | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 12 | J5421 | 2008 | FL | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 13 | J3182 | 2004 | OR | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 13 | J3751 | 2005 | OH | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 13 | J3976 | 2006 | MA | | 1.6_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 13 | J4179 | 2006 | RI | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 13 | J4726 | 2007 | FL | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 13 | J5032 | 2008 | PA | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 14 | J3720 | 2005 | PA | | Lm3.42 | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J2422 | 2003 | RI | | 1.7_4b | + | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 15 | J3070 | 2004 | TN | | 1.62_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J3120 | 2004 | WA | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J3309 | 2005 | IA | | | + | + | | | | | | | | | | | | | | | | | | | | | |

**Table 7.4 Continued**

| Hybridization Pattern | Name | Isolation Year | Country/State | Clonal Group | MLGT | Sau3AI[a] | MboI[a] | 85M[b] | Non-ECIIC-WAP[b] | 4bSF18[b] | Hsp001219[b] | ECIIC-WAP[b] | H18RP9/10[b] | J2479Res[b] | J3115Met[b] | LMSG_01573[b] | LMSG_01574C[b] | LMSG_01574[b] | LMSG_01578[b] | LP[b] | J3976TypeIS_V[b] | J3976NG[b] | 2007-640LP[b] | Mn[b] | McrB[b] | J1-194_1[b] | lmo0737[c] | Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | J3554 | 2005 | MS | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J3618 | 2005 | NJ | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J3684 | 2005 | OR | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J3762 | 2005 | IN | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J3795 | 2005 | AZ | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J3909 | 2006 | OR | | | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J4428 | 2007 | OH | | 1.62_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 15 | J4434 | 2007 | TN | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 15 | J4435 | 2007 | WI | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 15 | J4465 | 2007 | CO | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 15 | J5038 | 2008 | MI | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 16 | J3415 | 2005 | TX | | 1.7_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 17 | J4221 | 2006 | MD | | 1.75_4b | + | + | | | | | | | | | | | | | | | | | | | | | |
| 17 | J5166 | 2008 | KS | | | + | + | | | | | | | | | | | | | | | | | | | | P | |
| 18 | J2571 | 2003 | KY | | Lm3.42 | - | + | | | | | | | | | | | | | | | | | | | | | (20) |
| 19 | J2479 | 2003 | MI | | Lm3.42 | + | - | | | | | | | | | | | | | | | | | | | | | (20) |
| 20 | J3115 | 2004 | VA | | 1.73_4b | + | + | | | | | | | | | | | | | | | | | | | | | |

[a] Negative and positive signs represent resistance and susceptibility to the restriction endonuclease digestion, respectively. The letter P indicates partial digestion.

**Table 7.4 Continued**

[b] White and black squares signify the absence and presence of the signal in the DNA-DNA hybridization, respectively. When the letter P is inside a square, the white and black squares represent the absence and presence of the PCR amplicon, respectively, with the primers used to amplify a DNA probe.

[c] White and black squares with the letter P inside indicate whether a strain reveals the typical serotype 4b profile or the atypical serotype 4b profile with the addition *lmo0737* band in the multiplex PCR serotyping devised by Doumith et al. (6).

[d] N, Not tested.

**Table 7.5 Prevalence of serotype 4b clonal groups among 142 serotype 4b isolates from human listeriosis cases in the United States, 2003-2008**

| Hybridization Pattern | Clonal Group | B5M | Non-ECIIC-WAP | 4bSF18 | Hsp001219 | ECIIC-WAP | H18RP9/10 | J2479Res | J3115Met | LMSG_01573 | LMSG_01574C | LMSG_01574 | LMSG_01578 | LP | J3976TypeIS_V | J3976NG | 2007-640LP | Mrr | McrB | J1-194_1 | lmo0737 | 2003 (%) | 2004 (%) | 2005 (%) | 2006 (%) | 2007 (%) | 2008 (%) | Total (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | ECI | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | 11 (58) | 7 (26) | 6 (24) | 9 (35) | 3 (14) | 11 (48) | 47 (33) |
| 2 | ECII | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | 3 (16) | 7 (26) | 5 (20) | 5 (19) | 2 (9) | 1 (4) | 23 (16) |
| 3 | ECV | □ | □ | □ | □ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | ■ | □ | □ | 0 | 1 (4) | 0 | 1 (4) | 0 | 0 | 2 (1) |
| 4 | ECIa | □ | ■ | ■ | □ | □ | □ | □ | □ | ■ | ■ | ■ | ■ | ■ | □ | □ | □ | ■ | □ | □ | □ | 0 | 2 (7) | 2 (8) | 1 (4) | 3 (14) | 4 (17) | 12 (8) |
| 5 | Group 1 | □ | □ | □ | □ | □ | ■ | □ | □ | ■ | □ | □ | □ | □ | ■ | ■ | ■ | ■ | ■ | □ | ■ | 0 | 0 | 0 | 1 (4) | 0 | 0 | 1 (1) |
| 6 | Group 1 | □ | □ | □ | □ | □ | ■ | □ | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | ■ | ■ | □ | ■ | 1 (5) | 1 (4) | 0 | 1 (4) | 1 (5) | 1 (4) | 5 (4) |
| 7 | Group 2 | □ | □ | ■ | □ | □ | ■ | □ | □ | ■ | ■ | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | ■ | 0 | 3 (11) | 0 | 0 | 2 (9) | 1 (4) | 6 (4) |
| <Non-clonal isolates> | | | | | | | | | | | | | | | | | | | | | | 4 (21) | 6 (22) | 12 (48) | 8 (31) | 11 (50) | 5 (22) | 46 (32) |
| 8 | | □ | ■ | ■ | □ | □ | □ | □ | □ | ■ | ■ | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | □ | 0 | 1 (4) | 2 (8) | 1 (4) | 1 (5) | 1 (4) | 6 (4) |
| 9 | | □ | ■ | ■ | □ | □ | □ | □ | □ | ■ | ■ | ■ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | □ | 0 | 0 | 0 | 0 | 1 (5) | 0 | 1 (1) |
| 10 | | □ | □ | ■ | □ | □ | ■ | □ | □ | ■ | ■ | ■ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | □ | 1 (5) | 0 | 0 | 3 (12) | 0 | 0 | 4 (3) |
| 11 | | □ | □ | □ | □ | ■ | ■ | □ | □ | ■ | ■ | □ | ■ | ■ | □ | □ | □ | ■ | ■ | □ | □ | 0 | 0 | 0 | 0 | 2 (9) | 0 | 2 (1) |
| 12 | | □ | ■ | ■ | □ | □ | □ | □ | □ | ■ | ■ | □ | □ | □ | □ | ■ | ■ | ■ | ■ | □ | □ | 0 | 1 (4) | 1 (4) | 0 | 2 (9) | 1 (4) | 5 (4) |
| 13 | | □ | ■ | ■ | □ | □ | □ | □ | □ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | ■ | ■ | □ | □ | 0 | 1 (4) | 1 (4) | 2 (8) | 1 (5) | 1 (4) | 6 (4) |
| 14 | | □ | □ | ■ | □ | □ | □ | □ | □ | ■ | ■ | ■ | □ | □ | □ | ■ | ■ | ■ | ■ | □ | □ | 0 | 0 | 1 (4) | 0 | 0 | 0 | 1 (1) |
| 15 | | □ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | 1 (5) | 2 (7) | 6 (24) | 1 (4) | 4 (18) | 1 (4) | 15 (11) |
| 16 | | □ | ■ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | ■ | □ | □ | □ | □ | □ | □ | 0 | 0 | 1 (4) | 0 | 0 | 0 | 1 (1) |
| 17 | | □ | □ | □ | □ | ■ | ■ | □ | □ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | 0 | 0 | 0 | 1 (4) | 0 | 1 (4) | 2 (1) |
| 18 | | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | ■ | □ | 1 (5) | 0 | 0 | 0 | 0 | 0 | 1 (1) |
| 19 | | □ | □ | ■ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | 1 (5) | 0 | 0 | 0 | 0 | 0 | 1 (1) |
| 20 | | □ | ■ | ■ | □ | □ | □ | □ | ■ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | □ | 0 | 1 (4) | 0 | 0 | 0 | 0 | 1 (1) |
| Total | | | | | | | | | | | | | | | | | | | | | | 19 | 27 | 25 | 26 | 22 | 23 | 142 |

## REFERENCES


1. **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:**1624-1629.

2. **Cheng, Y., R. M. Siletzky, and S. Kathariou.** 2008. Genomic Divisions/Lineages, Epidemic Clones, and Population Structure, p. 337-358. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

3. **Cheng, Y., J. W. Kim, S. Lee, R. M. Siletzky, and S. Kathariou.** 2010. DNA probes for unambiguous identification of *Listeria monocytogenes* epidemic clone II strains. Appl. Environ. Microbiol. **76:**3061-3068.

4. **den Bakker, H. C., X. Didelot, E. D. Fortes, K. K. Nightingale, and M. Wiedmann.** 2008. Lineage specific recombination rates and microevolution in *Listeria monocytogenes*. BMC Evol. Biol. **8:**277.

5. **den Bakker, H. C., E. D. Fortes, and M. Wiedmann.** 2010. Multilocus sequence typing of outbreak-associated *Listeria monocytogenes* isolates to identify epidemic clones. Foodborne Pathog. Dis. **7:**257-265.

6. **Doumith, M., C. Buchrieser, P. Glaser, C. Jacquet, and P. Martin.** 2004. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. J. Clin. Microbiol. **42:**3819-3822.

7. **Doumith, M., C. Cazalet, N. Simoes, L. Frangeul, C. Jacquet, F. Kunst, P. Martin, P. Cossart, P. Glaser, and C. Buchrieser.** 2004. New aspects regarding evolution and

virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. Infect. Immun. **72:**1072-1083.

8. **Drevets, D. A., and M. S. Bronze.** 2008. *Listeria monocytogenes*: epidemiology, human disease, and mechanisms of brain invasion. FEMS Immunol. Med. Microbiol. **53:**151-165.

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

10. **Eifert, J. D., P. A. Curtis, M. C. Bazaco, R. J. Meinersmann, M. E. Berrang, S. Kernodle, C. Stam, L. A. Jaykus, and S. Kathariou.** 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog. Dis. **2:**192-200.

11. **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:**2383-2390.

12. **Franciosa, G., C. Scalfaro, A. Maugliani, F. Floridi, A. Gattuso, S. Hodzic, and P. Aureli.** 2007. Distribution of epidemic clonal genetic markers among *Listeria monocytogenes* 4b isolates. J. Food Prot. **70:**574-581.

13. **Graves, L. M., and B. Swaminathan.** 2001. PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. Int. J. Food Microbiol. **65:**55-62.

14. **Kathariou, S., L. Graves, C. Buchrieser, P. Glaser, R. M. Siletzky, and B. Swaminathan.** 2006. Involvement of closely related strains of a new clonal group of *Listeria monocytogenes* in the 1998-99 and 2002 multistate outbreaks of foodborne listeriosis in the United States. Foodborne Pathog. Dis. **3:**292-302.

15. **Liu, D.** 2008. Epidemiology, p. 27-59. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

16. **MacDonald, P. D., R. E. Whitwam, J. D. Boggs, J. N. MacCormack, K. L. Anderson, J. W. Reardon, J. R. Saah, L. M. Graves, S. B. Hunter, and J. Sobel.** 2005. Outbreak of listeriosis among Mexican immigrants as a result of consumption of illicitly produced Mexican-style cheese. Clin. Infect. Dis. **40:**677-682.

17. **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:**2386-2395.

18. **Ragon, M., T. Wirth, F. Hollandt, R. Lavenir, M. Lecuit, A. Le Monnier, and S. Brisse.** 2008. A new perspective on *Listeria monocytogenes* evolution. PLoS Pathog. **4:**e1000146.

19. **Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones, and P. M. Griffin.** 2011. Foodborne illness acquired in the United States-major pathogens. Emerg. Infect. Dis. **17:**7-15.

20. **Sperry, K. E., S. Kathariou, J. S. Edwards, and L. A. Wolf.** 2008. Multiple-locus variable-number tandem-repeat analysis as a tool for subtyping *Listeria monocytogenes* strains. J. Clin. Microbiol. **46:**1435-1450.

21. **Stack, H. M., C. Hill, and C. G. M. Gahan.** 2008. Stress Responses, p. 61-96. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

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

23. **Wagner, M., and J. MacLauchlin.** 2008. Biology, p. 3-26. *In* D. Liu (ed.), Handbook of *Listeria monocytogenes*, 1st ed. CRC Press, Boca Raton, FL.

24. **Ward, T. J., T. F. Ducey, T. Usgaard, K. A. Dunn, and J. P. Bielawski.** 2008. Multilocus genotyping assays for single nucleotide polymorphism-based subtyping of *Listeria monocytogenes* isolates. Appl. Environ. Microbiol. **74:**7629-7642.

25. **Ward, T. J., P. Evans, M. Wiedmann, T. Usgaard, S. E. Roof, S. G. Stroika, and K. Hise.** 2010. Molecular and phenotypic characterization of *Listeria monocytogenes* from U.S. Department of Agriculture Food Safety and Inspection Service surveillance of ready-to-eat foods and processing facilities. J. Food Prot. **73:**861-869.

26. **Ward, T. J., L. Gorski, M. K. Borucki, R. E. Mandrell, J. Hutchins, and K. Pupedis.** 2004. Intraspecific phylogeny and lineage group identification based on the *prfA* virulence gene cluster of *Listeria monocytogenes*. J. Bacteriol. **186:**4994-5002.

27. **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:**4158-4164.