uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Sarah McIlwham
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Microbiology and Immunology)
GRADE / DEGREE

Department of Biochemistry, Microbiology and Immunology
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Defining the Tropism of *Listeria monocytogenes* Using a Comparative Genomics Approach

TITRE DE LA THÈSE / TITLE OF THESIS

Dr. Jeffrey Farber
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Dr. Susan Logan

Dr. Thien-Fah Mah

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Defining the Tropism of *Listeria monocytogenes* Using a Comparative Genomics Approach

Sarah C. McIlwham

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the M.Sc. Degree in Microbiology and Immunology

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Your file Votre référence
ISBN: 978-0-494-48481-4
Our file Notre référence
ISBN: 978-0-494-48481-4

ABSTRACT

Listeria monocytogenes is a human foodborne pathogen capable of causing invasive and non-invasive infections. Ubiquitous in the environment, strains are often classified into lineages I (serovars 1/2a, 1/2c, 3a; food); II (containing serovars 1/2b, 4b, 3b; human illness); and III (4a, 4c; food-production). It is not clearly understood how serovars become tropic to a particular niche. The objective of this project was to investigate strain differences in clinical, environmental and food isolates, with the aim of identifying genomic features responsible for the tropism of the microorganism for clinical, food and environmental niches.

Comparative genomic techniques, namely mixed-genome microarrays, dot-blot hybridization and suppressive subtractive hybridization were used to compare the genomes of *L. monocytogenes* isolates of different serotypes from a variety of different sources. Over 30 genome sequences, potentially involved in the tropism of the microorganism, were identified using Versions 2 and 3 of a *Listeria* mixed-genome array. These included genes with similarities to cystathionine β-lyase and homoserine O-acetyltransferase, a putative outer-membrane lipoprotein, conserved hypothetical proteins, a peptidoglycan lytic protein and a glycosyltransferase family 65 enzyme.

The glycosyltransferase family 65 gene was found to be present in isolates belonging to lineage II but was absent in lineage I isolates. By creating an isogenic deletion mutant of this gene, it was found that this enzyme is involved in cell wall biosynthesis and could be responsible for increased survival of lineage II strains in the host environment, thus rendering these isolates more prevalent in cases of human listeriosis.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisors Dr. Franco Pagotto and Dr. Jeff Farber for providing me with the opportunity to do this research. It would not have been possible without their support.

I would like to thank all the members (past and present) of the *Listeria* and *Enterobacter* laboratory at the Bureau of Microbial Hazards, namely, Nathalie Corneau, Karine Hébert, Elaine Daley, Kevin Tyler, Raquel Lenati, Paul Manninger and Amy Commodore for their technical assistance, advice and support. I would also like to thank all my other colleagues at the Bureau of Microbial Hazards, especially Dr. Catherine Carrillo, for their thoughts and guidance during this project.

Thank-you to my thesis advisory committee Drs. Craig Lee, Syed Sattar and Min Lin for their time and advice.

Many thanks to Turki Abujamel and Dr. Min Lin (Canadian Food Inspection Agency, Ottawa, Canada) for providing me with a protocol for the deletion mutant and to Dr. Chakraborty (Institute for Medical Microbiology, Giessen, Germany) for providing us with the pAUL-A plasmid.

Finally, thank-you to my family for their lasting support and understanding during the time I spent working on this project.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AFLP: amplified fragment length polymorphism

BHI: brain heart infusion

CFU: colony forming unit

CNS: central nervous system

dNTP: deoxynucleotide triphosphate

GT65: glycosyltransferase family 65

HPB#: Health Products Branch Number

IFN: interferon

IL: interleukin

LB: Luria-Bertrani

LPS: lipopolysaccharide

MLEE: multi-locus enzyme electrophoresis

MLST: multi-locus sequence typing

MRSA: methicillin-resistant *Staphylococcus aureus*

msDNA: multi-copy single-stranded DNA

NK: natural killer

OD: optical density

PBP: penicillin-binding protein

PCR: polymerase chain reaction

PFGE: pulsed-field gel electrophoresis

PGRP: peptidoglycan recognition protein

PTS: phosphotransferase system

RFLP: restriction fragment length polymorphism

rRNA: ribosomal RNA

RTE: ready-to-eat

RT-PCR: reverse transcriptase PCR

SSH: suppressive subtractive hybridization

SDS: sodium dodecyl sulphate

TA: tryptose agar

TBE: Tris-Borate-EDTA

TE: Tris-EDTA

TSB-YE: Trypticase soy broth with yeast extract

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction to *Listeria monocytogenes*

Background, Taxonomy and Phenotypic Characteristics of *Listeria monocytogenes*

Listeria monocytogenes, a foodborne pathogen, is responsible for approximately 2,500 cases of listeriosis per year in the United States, with approximately 20% of these cases resulting in death (14). *L. monocytogenes* belongs to the genus *Listeria*, which currently includes a total of six species: *L. monocytogenes*, *L. ivanovii*, *L. seeligeri*, *L. innocua*, *L. welshimeri* and *L. grayi*. Of these six species, *L. monocytogenes* and *L. ivanovii* are the only species known to be pathogenic to humans and other animals (77). Table 1 shows the major characteristics that can be used to differentiate *L. monocytogenes* from other *Listeria* species.

Bergey's Manual of Determinative Bacteriology (36) places the genus *Listeria* into Group 19, "Regular, non-spore-forming, facultatively anaerobic rods". The organism varies in size (0.4 × 0.5-2 μm), is motile between 20 and 25°C and is able to grow at temperatures ranging from < 0 to 45°C (57). The ability of the organism to grow and survive over a wide range of temperatures and in acidic environments provides a many challenges when attempting to control this organism in the food-supply (63). Many countries have adopted a "zero-tolerance" policy for the organism in food, while others have set tolerable limits dependant on the food item (57).

Murray *et al.* (53) were the first to describe *L. monocytogenes*, naming it *Bacterium monocytogenes*. Over the years, its name evolved from *Listerella hominis*, *Corynebacterium parvulum*, *Erysipelothrix monocytogenes* and *Corynebacterium infantisepticum* (64). From this time until the early 1980's, cases were primarily reported and studied involving workers (i.e., farmers and veterinarians) in contact with infected

animals. The first epidemiologically validated outbreak of listeriosis from a food source occurred in 1981, in the Canadian Maritime Provinces.

Table 1. Characteristics Used to Differentiate Different Species of *Listeria*

Characteristic	*L. monocytogenes*	*L. innocua*	*L. ivanovii*	*L. seeligeri*	*L. welshimeri*	*L. grayi*	*L. murrayi*
Gram stain	+	+	+	+	+	+	+
β-hemolysis	+	-	+	v	-	-	-
Acid production from sugar:							
Mannitol	-	-	-	-	-	+	+
L-Rhamnose	+	v	-	-	v	-	v
D-Xylose	-	-	+	+	+	-	-
CAMP-test (*S. aureus*)	+	-	-	+	-	-	-

v- results variable: 11-89% of strains are positive.

Adapted from Pagotto *et al.*, (57).

In this outbreak, 34 of the cases were perinatal and 7 were adult. The case fatality rates for the perinatal and adult cases were 27% and 28.6%, respectively. The outbreak was linked to coleslaw containing cabbage that had been contaminated with *L. monocytogenes* at a farm with a flock of sheep, two of which had died from listeriosis (28). However, the bacterium still remained largely unknown until 1985, when it was identified as the causative agent of a large foodborne outbreak in the United States involving Mexican-style soft cheese. This outbreak resulted in 142 cases of listeriosis with 48 deaths (28).

Listeriosis

The condition listeriosis refers to a variety of clinical manifestations caused by *L. monocytogenes* ranging from gastroenteritis to meningitis and sepsis, possibly resulting in death. Listeriosis can thus manifest itself in both an invasive and a non-invasive form. The non-invasive form of listeriosis usually presents itself as a febrile gastroenteritis (57). It is hypothesized that this form of listeriosis results from a limited invasion of the gut mucosa (77). The infectious dose of *L. monocytogenes* required to cause a case of gastroenteritis remains unknown. Foods recovered from outbreaks have shown a range in levels of the microorganism, from 3×10^1 CFU/mL to 1.6×10^9 CFU/mL, with an average of 10^5 CFU/mL (60). Invasive listeriosis is a more serious infection and is primarily seen in the immune-compromised. It occurs when the bacterium spreads from the gut to the liver and infects the bloodstream and central nervous system (CNS). Table 2 summarizes how the disease can present itself.

Table 2. Clinical Symptoms of Listeriosis

Host	Clinical Symptoms
Pregnant women	Fever, chills, flu-like symptoms, myalgia, preterm delivery, abortions, stillbirth, amnionitis, intrauterine/cervical infections
Immune-compromised	Meningitis, encephalitis, sepsis
Newborns	Sepsis, meningitis, encephalitis, granulomatosis, pneumonia
Healthy adults	Diarrhea, fever, gastroenteritis, fatigue, headache, general malaise, nausea, vomiting, meningitis, septicaemia
Other clinical symptoms	Cutaneous lesions, conjunctivitis, convulsions, hepatitis, osteomyelitis, upper respiratory tract symptoms

Adapted from Pagotto *et al.*, (57).

The dose response of *L. monocytogenes* required to cause an invasive infection in an immune-compromised host is unknown however, it is presumed to be lower than the average dose required for a non-invasive infection (60). The factors responsible for a case of gastroenteritis versus an invasive infection remain unclear. At present, risk factors for non-invasive listeriosis are unknown and strain differences in isolates causing invasive versus non-invasive listeriosis have not been found. In fact, strains causing non-invasive gastroenteritis have also caused invasive cases of listeriosis (60).

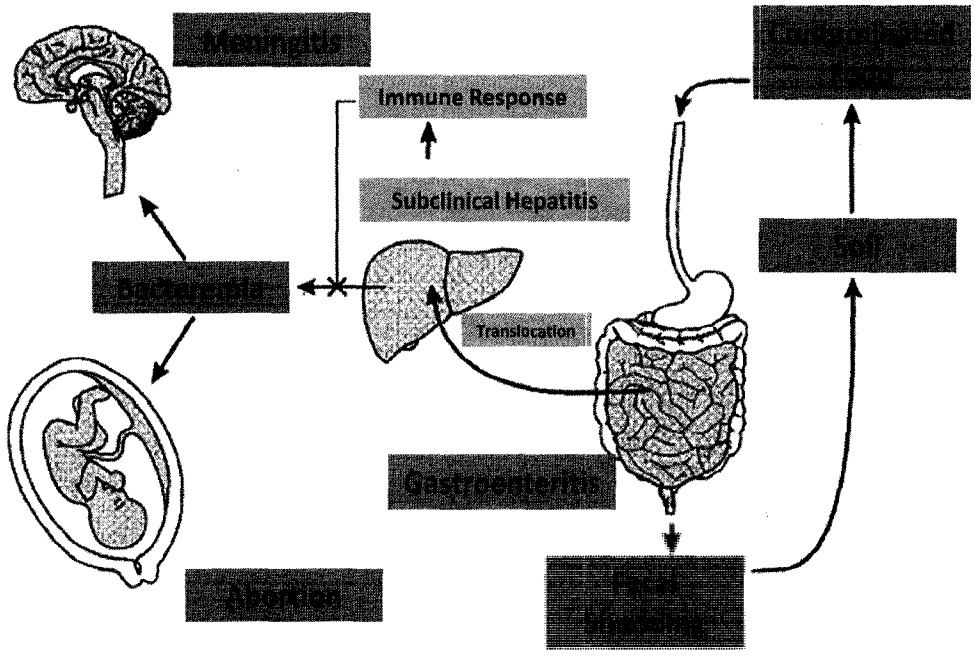
L. monocytogenes is exceptional in its ability to infect the host by being able to cross the three major protective barriers of the human body: gastrointestinal, materno-fetal and blood-brain (57). The majority of cases of listeriosis occur after the consumption of contaminated food (51). The organism is able to survive the non-specific immune system cells in the gut and multiply within the intestine, as well as be shed in the feces of the host. During the infection process, the organism is believed to invade the Peyer's patches in the intestinal epithelium, enter the lymphatic system and disseminate to the liver and spleen via the bloodstream. In most cases, the infection is limited at this point and eventually cleared by the host, but in the immunocompromised, elderly and pregnant women, clinical symptoms of listeriosis can appear (57). Figure 1 depicts the pathophysiology of the organism after ingestion by the host.

Listeriosis in non-pregnant adults is most frequently manifested as an infection of the CNS, with a mortality rate of 20%, but rates can reach as high as 60% if there is underlying disease (77). In 15-50% of cases, listeriosis manifests itself as an infection in the bloodstream, causing bacteremia. This can result in death in up to 70% of cases (77). In fetomaternal and neonatal listeriosis, infection of the mother is usually asymptomatic, but

may sometimes present as mild flu-like symptoms with miscarriage occurring between 2 and 14 days later (77). Occasionally, late onset neonatal listeriosis will occur 1-8 weeks postpartum. The symptoms include fever, meningitis and in some cases gastroenteritis and pneumonia. Mortality rates for late onset neonatal listeriosis are lower than other types of listeriosis, ranging from 10-20% (77).

Treatment for listeriosis involves a cocktail of ampicillin and gentamicin, however, the organism is susceptible to the majority of β-lactam antibiotics. Therapy in successful cases is usually provided for 2-3 weeks in order to prevent a relapse of infection. Since the infection with this microorganism has such a high case-fatality rate, it is important that treatment be administered promptly. Fortunately, antibiotic resistance of this microorganism is not currently a problem, unlike for many other microorganisms (48).

Figure 1. Pathophysiology of *L. monocytogenes* throughout the host after ingestion



Adapted from Vazquez-Boland *et al.*, (77).

Prevalence of *Listeria* in the Environment and Food Products

L. monocytogenes is able to survive and multiply over a wide range of temperatures, particularly low temperatures, a wide range of pH and low water activity, which enables it to inhabit and thrive in wide a variety of niches (57). Its ability to grow and survive in the environment enables the microorganism to contaminate food at many stages of the food production process such as agricultural ecosystems on farms (57, 71), abbatoirs, processing plants and at the retail level (44, 57). At the farm level, *L. monocytogenes* is predominately associated with pig and cattle feces. In one study, 52% of cattle feces contained *L. monocytogenes* (71) while in another, 20% of fresh pig and cattle feces were positive for the presence of the microorganism (75). While fresh manure appears to be a natural niche for the microorganism, samples obtained from storage tanks containing manure and soil fertilized with manure were shown to be negative for the presence of *L. monocytogenes*, indicating that the microorganism was not able to survive in manure during long-term storage (75).

Listeria spp. can also be found in water environments such as river water, sewage sludge, fresh water and sediments (4, 5, 17, 57, 78). Colburn *et al.* (17), in a study on sampled fresh water from the Humboldt-Arcata Bay in California, found that 81% of samples were positive for *Listeria* spp., with *L. monocytogenes* being the dominant species, isolated from 62% of the fresh water samples. The sediment samples taken from the same area as the fresh water samples had a lower percentage of *Listeria* spp. contamination, with 8/46 (17.4%) of the samples testing positive for the presence of *Listeria* spp. *Listeria* spp. was recovered more frequently from the water and sediments when domestic animals were located nearby. This emphasizes the nature of farm-to-fork contamination with *Listeria*

spp. Animals shed the microorganism into the environment, contaminating the water and soil, which in turn can contaminate fresh produce at the farm level.

Raw milk and related raw milk products, such as soft cheeses, can also serve as a niche for *Listeria*. This pathogen is able to contaminate the milk through unhygienic practices on the farm, from a cow with mastitis, encephalitis or an aborted fetus caused by a listeriosis infection (33, 57). The level of *L. monocytogenes* in raw milk has been shown to range from 1 to 62 CFU/mL (29, 56). The microorganism is able to survive, and possibly proliferate, during the ripening of certain soft cheeses because of the suitable level of water activity in the cheese (57). While it is more difficult for the microorganism to contaminate and proliferate in harder cheeses, such as cheddar, due to the pasteurisation process and the low water activity levels, it is still possible for the cheese to become contaminated after it has been pasteurised. *Listeria* spp. have been shown to colonize different niches in the manufacturing plant and possibly contaminate hard cheeses during the manufacturing process (57).

Control of *Listeria* spp. contamination in foods is difficult due to the ubiquitous nature of the microorganism in the environment. While raw milk and raw milk cheeses are a primary concern for the presence of *L. monocytogenes*, there have been many other cases of listeriosis from consumption of other ready-to-eat foods (Table 3). Due to the serious nature of the illness and the nature of the pathogen, it is important for the food industry to take measures to control and eliminate contamination of produce and other ready-to-eat foods with *L. monocytogenes.*

Table 3. Incidence of *L. monocytogenes* in various RTE food products

Type of Food	Incidence of *L. monocytogenes* (%)	Country
Retail vacuum-packaged processed meats	78/175 (44.6)	Australia
Small cooked sausages	243/6820 (2.1)	United States
Raw milk	21/220 (9.5)	Brazil
Raw milk	6/455 (1.3)	Canada
Soft Cheese	2/45 (4.0)	United Kingdom
Fresh shrimp	8/74 (10.8)	United States
Smoked salmon	3/33 (9.1)	Norway
Vegetable salads	1/63 (1.6)	Canada
Prepared salads	15/146 (10.2)	Spain

Adapted from Pagotto *et al.* (57).

Host Immune Response, Pathogenesis and Virulence of *L. monocytogenes*

The primary route of infection with *L. monocytogenes* in humans is via ingestion of contaminated food. During colonization of the gut, it is believed that *L. monocytogenes* invades the intestinal epithelium and/or the Peyer's Patches of the intestine. It then replicates rapidly and establishes an infection in the lymphoid tissues of the gut (77). It is believed at this stage that invasion of the epithelial cells induces the activation of NF-κB, resulting in an up-regulation of IL-15. At the same time, the activated lymphocytes in the intestinal epithelium produce Th1 cytokines, important in host defense against further infection (77). The microorganisms that are able to cross the intestinal barrier are carried by the lymph or blood to the mesenteric lymph nodes, spleen and liver (76). In healthy individuals, with an uncompromised immune system, the bacterium is rapidly cleared from the blood by macrophages. However, 90% of the accumulation of the microorganism occurs in the liver, where it is able to spread intracellularly and multiply. In healthy individuals, the clearance of the infection from the hepatocytes is rapid and the infection is resolved quickly (76). Neutrophils are an important factor in the initial innate immune response to *L. monocytogenes* infection in the hepatocytes. They are able to engulf the bacterium and destroy it using reactive nitrogen and oxygen intermediates, and the release of chemokines by the neutrophils brings macrophages to the site of infection. *In vivo* studies in mice have shown that resident macrophages, particularly Kupffer cells in the infected hepatocytes, release TNFα and IL-12. These cytokines induce natural killer (NK) cells to release IFNγ, which further activates the macrophages and increases their bactericidal activity (82).

In terms of adaptive immunity, the development of a *Listeria*-specific T-cell response is essential for final elimination of the bacterium and the generation of memory T-cells to protect the host in case of reinfection (82). When *L. monocytogenes* is inside the cytosol of the host cell, the proteins secreted by the microorganism are degraded by proteosomes and presented on the cell surface by the MHC class I molecules to CD8 T-cells. It is believed that the CD8 T-cells aid in clearing a *L. monocytogenes* infection by lysing infected host cells with perforin and granzymes to expose the bacteria to activated macrophages, more CD8 T-cells and that the CD8 T-cells also secrete IFNγ to activate and recruit more macrophages (82). CD4 T-cells are also believed to be involved in adaptive immunity to *L. monocytogenes*, however, their exact role remains unclear. It is known that CD4 cells also secrete IFNγ for macrophage activation, but other potential functions of the CD4 cells remain unknown (82).

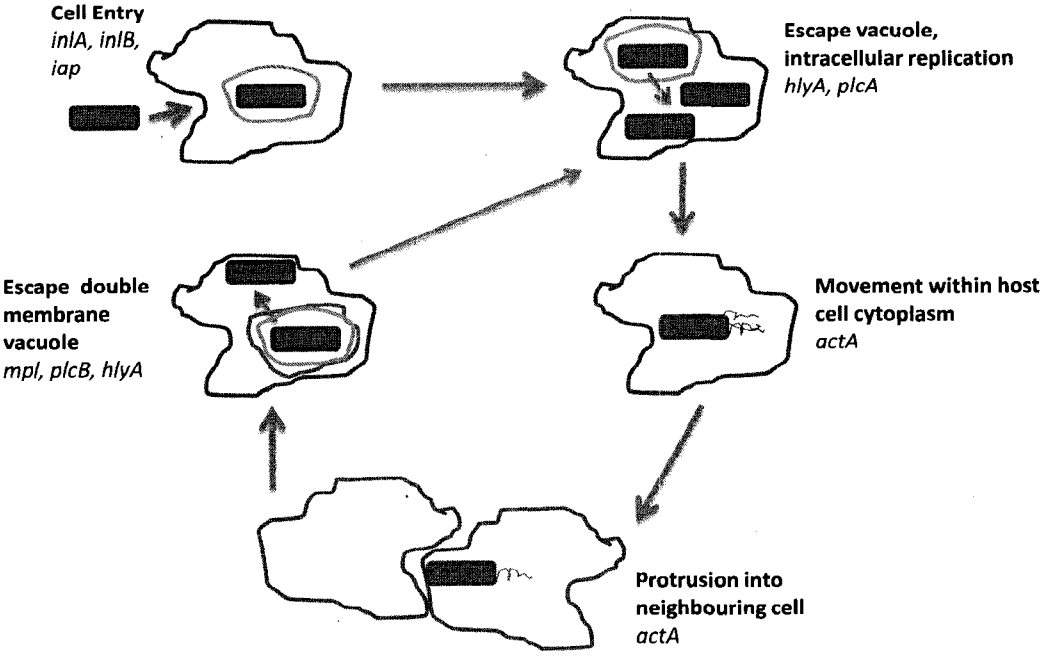
Due to the intracellular nature of the pathogen, antibody production is not believed to play a role in control of a *L. monocytogenes* infection. In hepatocytes, the microorganism goes through a complete intracellular lifecycle and spreads intercellularly from one hepatocyte to the next. In this manner, the pathogen is able to avoid humoral aspects of the immune system as it spreads and forms infectious pockets in the host hepatocytes, rendering the use of *L. monocytogenes*-specific antibodies futile in controlling the spread of the bacterium (77, 82).

Unfortunately for immune-compromised individuals, the clearance of *L. monocytogenes* from the liver does not occur with the same efficiency as a healthy individual. The proliferation of the microorganism is unhindered in the liver, and *L.*

monocytogenes is released back into the bloodstream where it can invade other organs, possibly causing meningitis, septicaemia or death (77).

The first step in the intracellular lifecycle of *L. monocytogenes* is entry into the cell. In phagocytic cells, such as macrophages, the bacterium is internalized by membrane-bound vacuoles. In non-phagocytic cells, *L. monocytogenes* gains entry through the action of two cell-surface proteins, InlA and InlB (Figure 2), encoded by the *inlA* and *inlB* genes (57). InlA is believed to mediate entry into epithelial cells, such as those found in the gut, whereas InlB is believed to be involved with entry into hepatocytes and fibroblasts (57). InlA contains an LPXTG motif near the C-terminal, seen in many Gram-positive bacteria, responsible for attachment of the protein to the cell envelope of the bacterium. On the other hand, the surface protein InlB does not possess the LPXTG motif, but appears to attach itself to the bacterial cell surface by association with the lipoteichoic acid in the cell wall (42). The mureine hydrolase protein p60, encoded by *iap*, has also been shown to be essential to the invasion process, but its exact role remains unclear (80). Once inside the host cell, *L. monocytogenes* is able to escape the membrane-bound vacuoles by releasing listeriolysin O (LLO), a hemolysin, encoded by the gene *hlyA*. Phospholipase C, encoded by the gene *plcA*, also plays a role in the initial escape from the vacuole. While the bacterium is in the host cell cytoplasm, it grows, multiplies and moves intracellularly through the activity of actin polymerization and the *actA* gene product, pushing the bacterium into adjacent host cells. Upon entry into a new host cell, the bacterium is double-encapsulated and the membranes are dissolved by a lecithinase encoded by the *plcB* gene. The *plcB* gene is activated by a metalloprotease, encoded by the *mpl* gene. Listerolysin O (*hlyA*) may also play a role in the breakdown of the double-membrane at this stage (57).

Figure 2. The Intracellular Lifestyle and Cell-to-Cell Spread of *L. monocytogenes*
Virulence genes associated with each step are indicated. The vacuole surrounding the bacterium upon entry into the cell is depicted in green.



Adapted from Pagotto *et al* (57).

All of the major virulence factors in *L. monocytogenes* involved in pathogenicity are under the control of positive regulatory factor A (PrfA), encoded by the *prfA* gene. Currently, it is the only confirmed transcription factor that is directly involved in virulence gene expression of *L. monocytogenes* (77). The transcription factor binds to a PrfA box, a 14-bp palindrome centred at the -14 position relative to the transcription start site of the promoter region and of the genes *actA*, *inlA*, *hlyA*, *plcA*, *mpl*, including the *prfA* promoter region (57). There are also many external factors that influence transcription of virulence genes. For example, the presence of fermentable sugars in media and a low pH will both repress the transcription of *prfA*, and therefore the expression of the other virulence genes, mentioned above, will not occur. On the other hand, *prfA* transcription is activated in minimal media, the presence of activated charcoal and at temperatures above 30°C (57).

The pathogenesis of *L. monocytogenes* is very well understood as compared to other foodborne pathogens. However, despite the known pathogenesis mechanisms, the virulence of the microorganism remains poorly understood. The mechanism of the main virulence gene regulator, *prfA*, has been well studied, but the current data fails to provide an accurate explanation for the variations in virulence seen in different food, clinical and environmental isolates, including the manifestation of invasive versus non-invasive forms of the disease (41).

Serotype Prevalence of *L. monocytogenes*

Studies into different aspects of virulence and pathogenicity of *L. monocytogenes* have greatly improved the understanding of how the microorganism causes human disease. However, the data that is currently available still fails to provide an accurate picture of why

(i) certain strains are more likely to cause invasive listeriosis, (ii) certain serovars are more pathogenic than others and (iii) some strains survive more readily than others in the environment (41).

The genetic structure of *L. monocytogenes* has been classified into three clonal lineages. Interestingly, these genetic divisions correspond with serotype, rarely seen in pathogens possessing multiple serovars (41). Analysis of the gene content patterns in *L. monocytogenes* by Hain *et al.* (35), allowed for the greater subdivision of the three known lineages. There are 14 serovars reported for *L. monocytogenes*: 1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4b, 4ab, 4c, 4b(x), 4d, 4e and 7 (58). It has been proposed that serotypes 1/2a and 3a belong to lineage I.1, serotypes 1/2c and 3c to lineage I.2, serotypes 4b, 4d and 4e to lineage II.1, serotypes 1/2b, 3b and 7 to lineage II.2, and serotypes 4a and 4c to lineage III (35). Lineage II the contains serotypes most commonly seen in human illness (1/2b and 4b), while lineage I contains serotypes (1/2a, 1/2c and 3a) more commonly isolated from foods (59, 61). The third lineage represents serotypes commonly seen in cases of animal listeriosis. Several studies have confirmed the presence of these three distinct lineages (57), using techniques such as multilocus enzyme electrophoresis (MLEE) (6), ribotyping (32), pulsed-field gel electrophoresis (PFGE) (10) and amplified fragment length polymorphism (AFLP) (1).

To date, the majority of studies with *L. monocytogenes* have used strains EGD (serotype 1/2a) or LO28 (serotype 1/2c). While these studies have provided a wealth of information about the microorganism, they have not included other serotypes, such as 1/2b and 4b, most commonly seen in human cases of listeriosis (41). Significant genomic and virulence trait differences may exist between serotypes. Thus, it is essential that these

differences be investigated to gain a more complete understanding of the differences in virulence between strains of *L. monocytogenes* and to discover potential factors leading to the behaviour of certain strains for outbreak situations in survival in the food manufacturing environment.

Objectives and Hypotheses

The objectives of this study are:

1. To identify differences, at the genomic level, in *L. monocytogenes*, that account for serovar prevalence in human illness; and
2. To elucidate a possible role for these genomic differences in the tropism of *L. monocytogenes* for a food, environmental or host niche.

The hypotheses of this study are:

1. Differences at the genomic level in *L. monocytogenes* isolates are responsible for the variations in the organism's tropism for a clinical, food or environmental niche.
2. Comparative genomic techniques can be used to discover these genomic differences.
3. The deletion of an identified genome marker found only in some isolates can alter the ability of the organism to survive in a particular niche.

Chapter 2: Comparative Genomics of *L. monocytogenes*

The information in this chapter was presented in poster form at the 2006 Federal Food Safety and Nutrition Research Meeting (Ottawa, ON), the 2007 International Symposium on the Problems of Listeriosis XVI (Savannah, GA), the 2007 Federal Food Safety and Nutrition Research Meeting (Winnipeg, MB) and the 2007 Health Canada Science Forum (Ottawa, ON). The information was presented in oral form at the 2007 International Association of Food Protection 94th Annual Meeting (Lake Buena Vista, FL).

INTRODUCTION

DNA Microarrays and Full Genome Sequence Bioinformatics

Differences in environmental fitness and virulence amongst strains of *L. monocytogenes* may be attributed to variations in gene expression. However, there is also the possibility that the variations are a direct result of differences between strains at the genomic DNA level (12), i.e., due to the presence or absence of specific genes. While the option of sequencing entire bacterial genomes is available, due to the cost, it is still not feasible for most laboratories. Different DNA sub-typing methods have been used on a number of different bacteria. Techniques such as multi-locus sequence typing (MLST), AFLP analysis, restriction fragment length polymorphism analysis (RFLP), comparisons of 16S rRNA genes and PFGE are presently used to investigate genomic differences of many bacteria (73). However, these techniques do not allow for an in-depth look at minute variations that may occur between isolates of the same bacterial species, which may be responsible for the different characteristics seen amongst isolates.

The use of DNA microarrays to study bacteria is becoming increasingly popular. Whole genome comparisons are often able to identify sets of "core genes" for a particular bacterial species and other genes that may only be present in a few strains of the same species (22). These types of studies are important as they identify potential factors that may explain why some strains are more pathogenic or have greater environmental fitness, than others of the same species. Furthermore, the evolution of particular bacterial genomes can be studied, providing clues about the emergence of bacterial pathogens (22).

Comparative genomics studies using genome sequence data and DNA microarray have been used successfully in the study of many microorganisms such as *Campylobacter jejuni*, *Escherichia coli, Vibrio cholera, Helicobacter pylori, Staphylococcus aureus* and various *Listeria* spp. (11, 12, 22, 73). In the study of the *C. jejuni* genome, a full genome array was created from a single strain and used for hybridization of many different isolates. It was shown that large regions of the *C. jejuni* genome were very stable. However, several genes that were divergent and highly variable were discovered, and these may be of use for sub-typing different strains of the microorganism (73). *E. coli* was the first microorganism to have experiments performed on its genome using the Affymetrix™ oligonucleotide-based chip. Many studies of the *E. coli* genome have used The GeneChip® *E. coli* Genome Array to study the intricacies of the genome of this microorganism (22). In studies of *V. cholera*, genomic arrays using PCR products of the open reading frames (ORFs) were designed based on the available genome sequence of a pandemic strain; the genomic DNA of other strains was hybridized to the array. The results showed a high degree of conservation of genes between all isolates, with divergence only occurring in <1% of the genes, the majority of which were associated with virulence factors of pandemic strains (22, 25). The study of the *S. aureus* genome used more of a mixed-genome approach, encompassing virulence genes from four important strains of the microorganism (67). They found that approximately 22% of the genome was composed of accessory genes such as phage genes, integrases and transposases. Their data also showed that the problematic methicillin-resistant *S. aureus* (MRSA) did not evolve from a single ancestor, but rather that these strains have evolved independently on several different occasions. Also, it was suggested that EMRSA-15, one of the most problematic strains in the UK, emerged from a

large recombination event within the genome (22, 67). The study of the *H. pylori* genome using a whole-genome microarray of a pathogenic isolate revealed that 22% of the genome was composed of accessory genes. When the genomes of two isolates that caused different levels of inflammation in gerbils were compared, they found a large deletion in the *cag* pathogenicity island in the isolate causing less inflammation (22, 38).

One of the major advances in the study of *Listeria* was the publication of the *L. monocytogenes* genome from strain EGD and the closely related, but non-pathogenic *L. innocua*. Sequence analysis and other comparative genomics tools can be performed using the available sequence data to establish a link between pathogenicity, biochemistry, physiology and the genomic DNA sequences (11). In 2003, Buchrieser *et al.*, compared the genomic DNA sequences of *L. monocytogenes* strain EGD and *L. innocua* strain CLIP 11262 to improve the understanding of the genetic basis of the virulence of *L. monocytogenes* (11). It was found that a large number of surface proteins were encoded by the *Listeria* genome. In total, there were 133 surface proteins encoded in the *L. monocytogenes* EGD genome, of which 30 were absent from the *L. innocua* CLIP11262 genome. Of these 30 proteins absent from *L. innocua* CLIP11262, 20 belonged to the LPXTG surface protein family motif, common cell surface proteins in Gram-positive bacteria. Both genomes also encoded a large number of transport proteins, predominately the ABC transport system and the phosphoenolpyruvate-dependant phosphotransferase system (PTS). The ABC transport system was conserved in the *L. innocua* and *L. monocytogenes* genomes, however not all of the PTS genes were conserved. It is believed that the presence of such a large number of transport system genes aids in the survival of the organism in diverse environmental conditions, such as in a wide range of temperatures,

acidities and low levels of free water (11). It should also be noted that the pathogenicity loci found in *L. monocytogenes* are usually absent in *L. innocua*, particularly the *L. monocytogenes*-specific allele of *iap*, as well as *inlA*, *inlB*, *inlC*, *hlyA* and *plcA*. However, recently an *L. innocua* strain was found to possess and express two of these virulence genes, *hlyA* and *plcA* (39). The strain was found to be avirulent, but the finding suggests that caution must be taken when detecting haemolytic *Listeria* in foods to ensure that it is not a non-pathogenic *L. innocua* being identified as *L. monocytogenes* (39).

While these studies were useful in identifying some genes that were species-specific to *Listeria* and other genes that were highly conserved across two species, it did not compare genomes of different *L. monocytogenes* strains to move towards an understanding of strain specific or serotype specific virulence or environmental adaptation genome features.

Doumith *et al.* (24) performed a comparative genomics experiment with *L. monocytogenes* to gain further insight into the evolution and virulence of the microorganism. Initially, they compared the genome sequences of two different serotypes of *L. monocytogenes*, the serotype 4b epidemic partially sequenced isolate and the serotype 1/2a fully sequenced EGD isolate. Approximately 8% of the sequences were serotype 4b specific, indicating a divergence between the serotypes (24). A DNA array was then constructed based on these differing parts of the genomes and multiple *Listeria* isolates were screened for the presence of these genes. Many *L. monocytogenes* specific genes were found and three surface proteins specific to 4b were identified (24).

The scope of the study by Doumith *et al.* (24), was narrowed by the comparison of only two sequences and the use of genes specific to only *L. monocytogenes* serotype 4b in

the DNA array experiments. To help overcome some of these factors, in the present study, a mixed-genome array was used to investigate the differences between *L. monocytogenes* isolates. This approach, which included pooled genomic DNA from 15 different isolates of *L. monocytogenes*, provides a more diverse comparison of sequences potentially responsible for the tropism of the microorganism for food, environmental and host environments.

Suppressive Subtractive Hybridization

Suppressive subtractive hybridization is a suppression PCR-based technique used to amplify pieces of DNA that do not hybridize to each other (i.e., are different from one another). It has been used in creating subtractive cDNA libraries of differentially regulated and tissue-specific genes in eukaryotes (21). In bacterial systems, it has recently been successfully used to study microbial population diversity in rumen animals (30), genomic differences in *H. pylori* (2), *Pseudomonas flourescens* (50), genomic differences between *Salmonella enterica* serovar Typhi and *S. enterica* serovar Typhimurium (27) and genome differences between *E. coli* and *S. enterica* serovar Typhimurium (7). In this study, the technique was used to find genomic sequence differences between *L. innocua* and *L. monocytogenes*, as well as in intraspecies comparisons of *L. monocytogenes.*

Confirmation of Array Data Using Dot-Blot Hybridization and PCR Screening of *L. monocytogenes* Isolates

In order to confirm the hybridization tendencies seen in the mixed-genome array experiments, dot-blot hybridization was used. Although this technique is the reverse of the

array-based experiments, with the fragmented genomic DNA being blotted onto a membrane and the target PCR product being used as the probe, it is a useful tool in confirming the hybridization results. Dot-blot assays can also address the sensitivity of the microarray experiments.

PCR was also used as a confirmatory tool for the microarray data for targets that were identified as being present in some isolates and absent in others. After sequencing the targets of interest from the microarray, primers were designed to use for the PCR amplification and the *L. monocytogenes* isolates used in the array hybridization were screened for the presence or absence of a particular sequence. The PCR screening was expanded to include 45 diverse strains of *L. monocytogenes*. This was a quicker and more feasible method to screen isolates for a specific sequence rather than using the mixed-genome array to screen a large number of isolates for a handful of selected target sequences.

MATERIALS AND METHODS

Creation of a Genomic DNA Library

A custom random shotgun genomic library, composed of a mixture of 15 different *L. monocytogenes* isolates, was constructed by Amplicon Express (Pullman, WA, USA). Fragments ranging in size from 200-1300 bp were inserted into the Stratagene (Cedar Creek, TX) vector pPCR-Script (Figure 3) and cloned into *E. coli* to generate a genomic library containing approximately 5800 clones. Pertinent information about the 15 strains used to create the library can be seen in Table 4.

Generation of Features for the Mixed Genome Array

Clones were picked from a frozen stock and grown overnight at 37°C in a 96-well plates containing 100 µL of BHI with ampicillin (100 µg/mL). After incubation, 5 µL of the overnight growth was added to a new 96-well plate containing PCR Master Mix. Each 100 µL PCR reaction contained five µL of overnight culture, ddH_20 up to 100 µL, 1X PCR buffer for Platinum® Taq DNA Polymerase (Invitrogen Canada Inc., Burlington, ON), 1.5 mM $MgCl_2$ (Invitrogen Canada Inc.), 0.2 mM dNTPS (Roche Applied Science, Laval, QB), 0.55 µM of M13 forward and reverse primers (Sigma-Genosys, Sigma-Genosys, Sigma-Aldrich Inc., Oakville, ON) and 5U of Platinum® Taq DNA Polymerase (Invitrogen Canada Inc.). Primer sequences are as follows: M13F 5'-TTG TAA AAC GAC GGC CAG T-3' and M13R 5'-GGA AAC AGC TAT GAC CAT G-3'. Primers were designed using the Primer3 software (http://frodo.wi.mit.edu/). The PCR conditions were as follows: 5 min at 95°C followed by 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 min. After the 35 cycles, an extra extension of 72°C for 10 min was added.

Table 4. Description of the 15 *L. monocytogenes* Isolates Used to Create the Mixed Genome Library

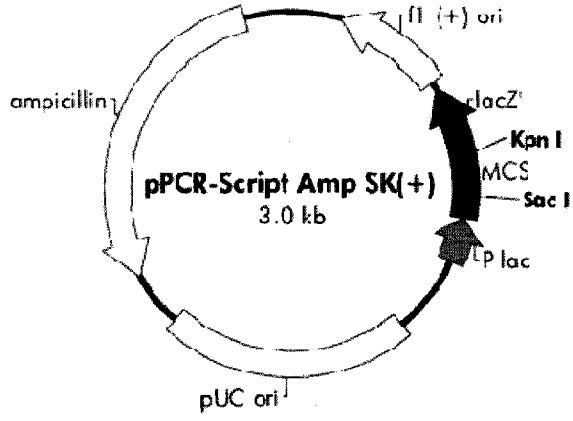
HPB #	Characteristics	C/F/A[1]	Serotype	Lineage
12	Human spinal fluid (Scotland)	C	1/2c	I
308	ATCC 35152 (guinea pig, England)	A	1/2a	I
2411	Case No. 2 Caesarian	C	1/2a	I
2863	Finland (butter-related)	C	3a	I
3	Human outbreak (Boston, 1983)	C	4b	II
22	Jalisco Cheese outbreak (1985)	F	4b	II
850	Swiss epidemic strain	C	4b	II
774	Pâté, U.K. outbreak	F	4b	II
1174	Glazed pork tongue	F	4b	II
1380	Crab meat packages	F	1/2b	II
1945	Rainbow trout outbreak	F	4b	II
2142	Hotdogs (Ball Park)	F	4b	II
2129	Italy outbreak, rice salad	F	1/2b	II
2	Coleslaw outbreak (Halifax)	F	4b	II
2905	Tiny Tomme Cheese outbreak	F	4b	II

[1]C= Clinical isolate; F= Food isolate; A= animal isolate

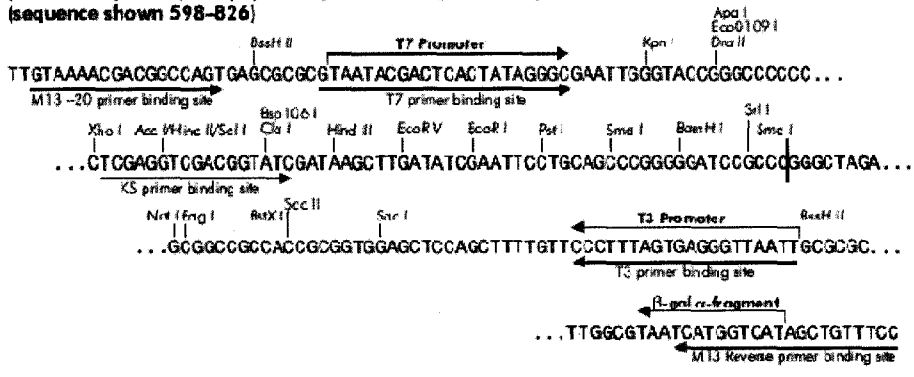
[2]Lineage designation based on Hain *et al.* (35).

Figure 3. Map of Stratagene pPCR-Script Vector Used for Creating the *L. monocytogenes* Mixed Genome Library

f1 (+) origin 135–441
β-galactosidase α-fragment 460–816
multiple cloning site 653–760
lac promoter 817–938
pUC origin 1158–1825
ampicillin resistance (*bla*) ORF 1978–2633



pPCR-Script Amp SK(+) Multiple Cloning Site Region
(sequence shown 598–826)



Samples were held at 4°C until analysis by agarose gel electrophoresis. PCR products were analyzed on a 1% agarose gel in 1X TBE buffer, run at 100 V for 30 minutes. PCR products between 200 and 1200 bp were selected to use for printing as features in the mixed genome array.

Prior to printing, all PCR products were cleaned using phenol-chloroform-isoamyl alcohol (25:24:1) (Fisher Scientific, Ottawa, ON) in MaXtract High Density tubes (Qiagen, Mississauga, ON). The MaXtract tube was spun at 18,300 × g for 1 min, then 100 µL of PCR product was added to the tube along with one volume of phenol-chloroform-isoamyl alcohol (25:24:1) and mixed well. The tube was then spun at 18,300 × g for 5 min. The top phase above the gel in the MaXtract tube was collected and placed in a new 1.5 mL Eppendorf tube. The PCR product was precipitated by adding 1/10 volume of 3 M sodium acetate (pH 5.2) and 2 volumes of ice-cold 95% EtOH, then spun at 21,000 × g for 20 min at 4°C. The supernatant was discarded and the DNA pellet was washed twice with cold 70% EtOH. The pellet was dried and resuspended in 40 µL ddH_2O and stored at -20°C. DNA concentration after cleaning and precipitation was measured using the NanoDrop® Spectrophotometer ND-1000 (Thermo Fisher Scientific, Waltham, MD). PCR products with a minimal concentration of 40 ηg/µL were selected for use as targets in the mixed genome array.

Printing the Mixed Genome Arrays

PCR products to be used as targets in the mixed genome array were selected as described above. A volume of PCR product containing 1.2 µg of DNA was aliquoted into a 1.5 mL Eppendorf tube and the sample was completely dried using the Savant SpeedVac

Concentrator (Fisher Scientific). The pellets were all resuspended in 12 μL of Corning Pronto! Universal Spotting Solution (Fisher Scientific) and transferred to a 384-well plate. Buffer spots and empty wells were incorporated into the array design as negative and background controls, respectively. The arrays were printed using the VersArray ChipWriter™ Compact System (BioRad Laboratories Canada Ltd., Mississauga, ON) and the printing program was designed using the VersArray ChipWriter™ Compact software. Arrays were printed in triplicate on Corning UltraGAPS™ (Fisher Scientific) amino-silane coated glass slides.

Preparation of Slides for Hybridization

After printing, the slides were dried for 48 h in a dessicator and the DNA was subsequently fixed to the slides by UV crosslinking at 225 mJ using the UV Stratalinker® 1800 from Stratagene. Slides were stored in light-protective casing at room temperature until hybridization.

Immediately prior to hybridization, the PCR products printed on the array were denatured by plunging into water at 95°C for 2 min and then quick chilled by rinsing for 2 min in 95% EtOH. The slides were dried by centrifugation inside 50 mL conical tubes for 5 min at 50 × g. To reduce background hybridization, the slides were incubated in a blocking solution containing 5X SSC, 0.1% SDS, 0.1 mg/mL BSA and 40% formamide for 45 minutes at 45°C. The slides were rinsed twice with ddH_20 and once with isopropanol and dried by centrifugation in a 50 mL conical tube for 5 min at 50 × g. The slides were placed into hybridization cassettes and loaded into the GeneTAC™ Hybridzation Station (BioRad Laboratories Canada Ltd).

Preparation of Genomic DNA for Use as a Probe in Array Hybridization

A single colony of *L. monocytogenes* was picked from a TA plate and inoculated into 50 mL of TSB-YE broth and incubated overnight at 37°C. The cells from overnight culture were spun down at 3,900 × g for 20 min. The supernatant was decanted and the pellet was resuspended in three mL of lysis buffer (6.7% sucrose, 50 mM Tris pH 8, 1 mM EDTA and 40 mg/mL lysozyme) and transferred to a 1.5 mL Eppendorf tube. The cells were incubated at 37°C for 15 min. A total of 100 µL of 10% SDS was added to the solution and incubated at 60°C until the solution turned clear (approximately 20 min). After incubation with SDS, 20 µL of 10 mg/mL Proteinase K was added and incubated at 60°C for 15 min. To clean the genomic DNA, MaXtract tubes were used in combination with phenol:chloroform:isoamyl alcohol (25:24:1). The MaXtract tubes were prepared as previously described and the genomic DNA was added to the tube along with one volume of phenol:chloroform:isoamyl alcohol (25:24:1) and mixed well. The tubes were spun at 18,300 × g for 5 min in a table top centrifuge and the top layer was transferred to a new 1.5 mL Eppendorf tube. To remove RNA from the genomic DNA, 50 µL of 500 ηg/µL RNase, DNase free (Roche) was added and incubated at 37°C for 30 min. The DNA was cleaned again using phenol:chloroform:isoamyl alcohol (25:24:1) and MaXtract tubes as described above. The cleaned DNA was transferred to a new 1.5 mL Ependorf tube. The genomic DNA was precipitated by the addition of two volumes of ice-cold 95% EtOH and 1/10 volume of sodium acetate (3 mM, pH 5.2), then spun at 21,000 × g for 20 min at 4°C. The supernatant was discarded and the DNA pellet was washed twice with cold 70% EtOH. The pellet was dried and resuspended in a minimum 100 µL of ddH_2O and stored at -20°C. If the DNA did not go into solution, additional ddH_2O was added or the pellet was heated

at 60°C for 5 min until the pellet went into solution. DNA concentration after cleaning and precipitation was measured using the NanoDrop® Spectrophotometer ND-1000 (Thermo Fisher Scientific).

Before the genomic DNA was labelled with Cy3 for use as a probe, it was broken down into smaller fragments by sonication using the Sonifier 250 (Fisher Scientific). A total of 20 μg of genomic DNA was added to the sonication buffer (50mM Tris, 10mM EDTA) up to a final volume of 100 μL. The sample was placed on ice and sonicated at a % duty cycle of 20 and an output control of 2 for 10 seconds. After the sonication was complete, 5 μL of the sonicated DNA was run on a 1% agarose gel to ensure that adequate fragmentation had been achieved. The sonicated sample was precipitated by the addition of two volumes of ice-cold 95% EtOH and 1/10 volume of sodium acetate (3 mM, pH 5.2), then spun at 21,000 × g for 20 min at 4°C. The supernatant was discarded and the DNA pellet was washed twice with cold 70% EtOH. The pellet was dried and resuspended in 60 μL of ddH_2O and stored at -20°C until use. DNA concentration after precipitation was measured using the NanoDrop® Spectrophotometer ND-1000 (Thermo Fisher Scientific).

Preparation of Cy3 labelled Probe

Four-hundred ηg of sonicated genomic DNA was labelled with the Cy3 dye using the Mirus Label IT® Cy3 Nucleic Acid Labelling Kit (Fisher Scientific). In total, 2 μL of Label Buffer A, 2 μL of Cy3 dye, 400 ηg of sonicated genomic DNA and ddH_2O up to 25 μL, were mixed together and incubated for 3 h at 37°C in the dark. After incubation, 25 μL of Label buffer A was added to bring the total volume to 50 μL and the probe was purified using AutoSeq™ Microspin G-50 columns (Amersham Pharmacia Biotech Inc.,

Piscataway, NJ) according to the manufacturer's guidelines. The purified probe was dried in the Savant SpeedVac Concentrator in the dark (Fisher Scientific). The probe was then resuspended in 100 µL of 1X hybridization buffer (GeneTAC™ Hybridization Solutions, NextGen Sciences Inc., Ann Arbor, MI) containing 40% formamide. The probe was chemically denatured using the reagents provided in the Mirus Label IT® Cy3 Nucleic Acid Labelling Kit. Prior to hybridization, 0.1 volume of Denaturation Buffer D1 was added to the probe, mixed well and incubated for 5 min at room temperature. The sample was then placed on ice for 2 min and following this, 0.1 volume of Neutralization Buffer N1 was added and incubated at room temperature for 5 min. Following chemical denaturation, the probes were heat-denatured for 5 min at 100°C, then immediately chilled on ice for 2 min. A total of 120 µL of probe was loaded onto the GeneTAC™ Hybridization Station (BioRad Laboratories Canada Ltd) and hybridized to the array for 18 h with a ramp-down hybridization temperature from 65°C to the holding temperature of 45°C. Automated washes with Medium and High Stringency Wash Buffers and a Post Wash Buffer (GeneTAC™ Hybridization Solutions, NextGen Sciences Inc.) were performed after hybridization.

When the hybridization and washes were complete, the slides were removed from the hybridization station and dried by centrifugation at 50 × g for 5 min. The slides were scanned using the VersArray ChipReader™ Unit and ChipReader™ Version 3.1 Software (BioRad Laboratories Canada Ltd) on the Cy3 channel. Images of the scanned results were saved for analysis.

Analysis of Array Results

Raw images generated from the scan of the array were loaded into ArrayPro Analyzer Software (Media Cybernetics Inc., Silver Spring, MD). The sum of the net intensity from all spots on the 3 sub-grids of the array was calculated and the net intensity of each target was converted into a % net intensity of the total array intensity. The background signal from each sub-grid for each target was subtracted from the intensity for each spot. In order to be considered significant, the cut-off for hybridization intensity of a target was set to three times higher than the background signal. The average net intensity for each target was calculated from each of the three sub-grids. Hybridization experiments for each strain of *L. monocytogenes* were performed in triplicate, so the final net intensity for each target on the array was compiled from an average of 9 hybridizations. Array targets with different net intensities between isolates, or showing hybridization in some strains and no hybridization in others, were considered for further analysis by dot-blot hybridization and sequencing.

Version 3 of the *Listeria* mixed-genome array was analyzed in a slightly different manner. Instead of converting each net intensity into a percentage value of the total net intensity, the net intensities for each isolate were compared to one another. Array probes which had a negative net intensity for some isolates, but a positive value for other isolates, were sent for sequencing. A negative net intensity indicates that hybridization did not occur between the isolate and the probe, so those probes which did not hybridize to all the *L. monocytogenes* strains tested were deemed to be of interest.

Dot Blot Hybridization

A total of 2 μg of sonicated *L. monocytogenes* genomic DNA, prepared as described above, was denatured by treating with 0.4 M NaOH and 10 mM EDTA and added onto a vacuum slot blot apparatus with a positively-charged nylon membrane (Roche Applied Science, Laval, QC). The DNA was bound to the membrane by vacuum suction of the 0.4 M NaOH and 10 mM EDTA solution. The membrane was removed when all the liquid had passed through and washed with 2X SSC. Targets of interest identified in the mixed-genome array were DIG-labelled (Roche Applied Science) and hybridized according to the manufacturer's protocols (DIG High Prime DNA Labelling and Detection Starter Kit I for Color Detection with NBT/BCIP, Roche Applied Science, December 2005). The hybridizations were performed overnight at 45°C in a hybridization oven. Post-hybridization washes, blocking of membranes, and colorimetric detection were also performed according to the manufacturer's protocols.

Suppressive Subtractive Hybridization

Suppressive subtractive hybridization was performed on selected serovars of *L. monocytogenes* isolates, including one *L. innocua* isolate. In each experiment, one strain was designated as the "Driver" and the isolate of interest was designated as the "Tester". The PCR-Select™ Bacterial Genome Subtraction Kit (Clontech Laboratories Inc., Mountain View, CA) was used for the experiments and the manufacturer's protocols were followed (Clontech Laboratories Inc, Protocol number PT-3170-1, Version PR621453).

The *E. coli* control genomic DNA provided in the kit was used as a positive control for each experiment.

Sequencing of Selected Features Showing Different Hybridization Tendencies in the Microarray

PCR products that showed different hybridization tendencies to different *L. monocytogenes* isolates were selected for further investigation. The PCR product generated using M13 primers was sent for sequencing to DNA Landmarks Inc (St-Jean-sur-Richelieu, QC). The ABI trace files were analysed using Chromas software from Technelysium Pty Ltd (Australia) and the consensus sequences were generated using Tera Term Software. Identities to the consensus sequences were determined by using the NCBI Nucleotide MegaBLAST algorithm, searching the Nucleotide Collection and Reference Genomic Sequences databases.

Screening *L. monocytogenes* Isolates for Selected Target Sequences

In total, 45 unrelated serovars of *L. monocytogenes* (Table 5) were selected for PCR screening of target sequences, selected based on differing array hybridization tendencies. PCR primers were designed using the Primer3 software for the following array target sequences: 001-G23, 002-E8, 002-C16 and 003-F3. The primer sequences and properties are described in Table 6.

Isolates selected for PCR screening were grown from frozen stock on TA plates for 48 h at 30°C. A single pure colony was picked from the plate and suspended in 50 µL of Tris Buffer (10 mM, pH 8). The cells were lysed by boiling for 10 min at 100°C and stored

at -20°C until use in the PCR reaction. The PCR conditions were as follows and were kept consistent for all primer sets: 2 min at 95°C followed by 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 min. After the 35 cycles, an extra extension of 72°C for 10 min was added. Samples were held at 4°C until analysis by agarose gel electrophoresis.

Table 5. *L. monocytogenes* Isolates Selected for PCR Screening for the Presence or Absence of Selected Target Sequences Obtained from Microarray Data

These isolates were also used in the antibiotic susceptibility testing described in Chapter 3.

HPB#	Serotype	Source[1]	Lineage[2]
2	4b	F	II
3	4b	C	II
22	4b	F	II
850	4b	C	II
774	4b	F	II
1174	4b	F	II
1945	4b	F	II
2142	4b	F	II
2905	4b	F	II
426	4b	E	II
842	4b	C	II
1015	4b	F	II
1026	4b	C	II
3724	4b	E	II
3763	4b	E	II
1380	1/2b	F	II
2129	1/2b	F	II
425	1/2b	E	II
394	1/2b	F	II
3830	1/2b	E	II
1031	3b	F	II
19	4d	A	II
1088	4d	C	II
2182	4e	C	II
11	1/2c	A	I
12	1/2c	C	I
3768	1/2c, 3c[3]	E	I
3812	1/2c	E	I
3826	1/2c	E	I
3883	1/2c	E	I
2411	1/2a	C	I
427	1/2a	E	I
337	1/2a	C	I
977	1/2a	F	I
1012	1/2a	A	I
1045	1/2a	F	I
2443	1/2a	A	I
3726	1/2a	E	I
3731	1/2a, 3a[3]	E	I
3744	1/2a	E	I
3838	1/2a	E	I
3893	1/2a	E	I
3943	1/2a, 3a[3]	E	I
3949	1/2a, 3a[3]	E	I
2863	3a	C	I

[1]C=Clinical isolate; F=Food isolate; E=Environmental isolate; A=Animal isolate

[2]Lineage designation based on Hain *et al.* (35).

[3]Serotype based on multiplex PCR as described by Doumith *et al.* (23).

Table 6. Primers Used in Screening Selected *L. monocytogenes* for the Presence of Target Sequences Identified in the Mixed Genome Array

Primer pair name	Target ID	Product size (bp)	Tm (°C)	Sequence (5'-3')
001G23	Similar to cystathionine β-lyase	218	60	**F**[1]: CTTGCGCGTATTCTTTAGCC **R**[2]:GAAAGTTTAGGCGGTGTGGA
002C16	Putative cell surface protein	388	60	**F:** TCAGGAAGCTCTCCGTTTTG **R:** TCACTGGAATGCCTGCTTTA
002E8	Reverse transcriptase	373	60	**F:** CTCGTTCGAAATTCCCAAAA **R:** AGGTGCTCCCTGTGGTAATG
003F3	Glycosyltransferase family 65	360	60	**F:** TCGAAAAAGCTTCATCAGCA **R:** TCTCGACATCCGCTTTATCC

[1]F: Forward primer.
[2]R: Reverse primer.

Each 100 μL PCR reaction contained 1 μL of boil prep DNA, ddH_2O up to 100 μL, 1X PCR buffer for GoTaq® Flexi DNA Polymerase (Promega), 2 mM $MgCl_2$ (Promega), 0.2 mM dNTPS (Roche Applied Science), 1 μM of forward and reverse primers (Sigma-Genosys Inc.) and 5 Units of Promega GoTaq® Flexi DNA Polymerase (Fisher Scientific). PCR products were analyzed on a 1% agarose gel in 1X TBE buffer, run at 100V for 30 min. All PCR products were sent to DNA Landmarks Inc for sequencing. The ABI trace files were analysed using Chromas software and the consensus sequences were generated using Tera Term Software. Homologies to the consensus sequences were determined by using the NCBI Nucleotide MegaBLAST algorithm, searching the Nucleotide Collection and Reference Genomic Sequences databases to ensure that the correct sequences had been amplified.

RESULTS

Listeria monocytogenes Mixed-Genome Array

Array Version 1

Version 1 of the *L. monocytogenes* mixed-genome microarray was intentionally designed as a very small array in order to optimize the correct hybridization conditions and parameters for the purpose of our study. A total of 56 features were printed for this array. A schematic of the array design is shown in Figure 4. The PCR products were spotted on to the array in random order, with negative control spots from the printing buffer and printing pin alone placed in throughout the array. The optimal conditions were found to be overnight ramp-down hybridization from 65°C to 45°C, with the probe reconstituted in 1X hybridization buffer containing 40% formamide. Cross-hybridization was noted with the Spot Report Alien 7 DNA PCR product with several isolates of *L. monocytogenes*. Thus, these features were removed from the second version of the array.

Array Version 2

The second version of the mixed-genome microarray was larger, consisting of 154 features printed on to the array (Figure 5). Multiple target sequences of interest were discovered from the differential hybridization tendencies to the features on this array. A complete list of target sequences with different hybridization tendencies to various *L. monocytogenes* isolates and serotypes is shown in Table 7. Typical images of arrays used for analysis are shown in Figure 6.

Figure 4. *Listeria* Mixed-Genome Array Version 1 Test Chip

Spot Report Alien DNA in various concentrations are shown in red. Negative control spots are indicated in yellow and purple. PCR products from the mixed-genome library are light blue.

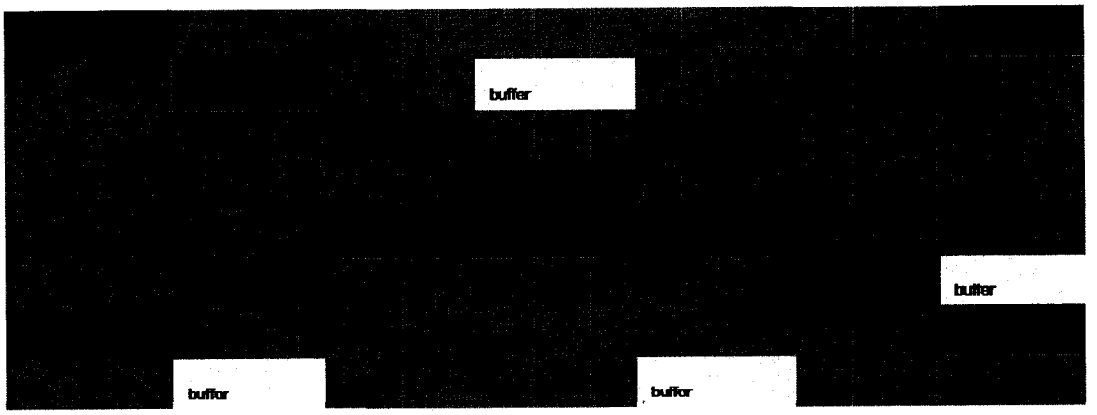

buffer
buffer
buffer
buffer

Figure 5. *Listeria* Mixed-Genome Array Version 2

Spot Report Alien DNA in various concentrations are shown in red. The number refers to the concentration of Alien DNA that was spiked in ηg/μL. Negative control spots are indicated in yellow and dark blue. PCR products from the mixed-genome library are light blue.

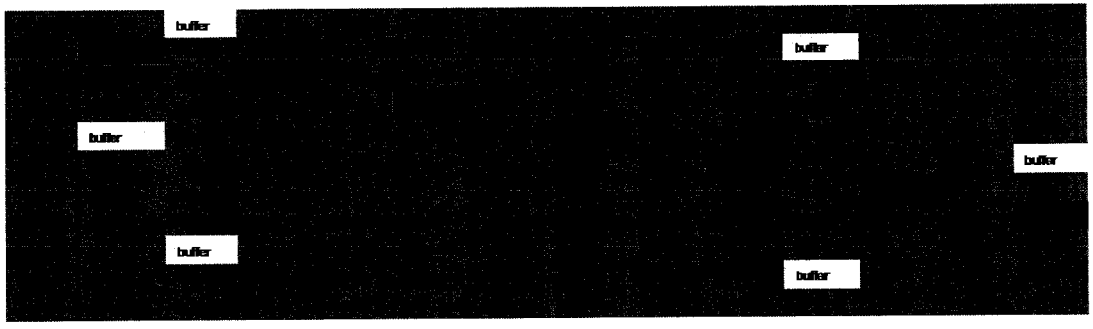

buffer
buffer
buffer
buffer
buffer
buffer

Table 7. Summary of Net Intensities and Sequence Identities of Selected Features from *Listeria* Mixed Genome Array Version 2

Feature ID	BLAST Search	% Normalized Signal Intensity					
HPB# (serotype, source)		HPB#2863 (3a, C[3])	HPB#3 (4b, C)	HPB#850 (4b, C)	HPB#1380 (1/2b, F)	HPB#17 (4c,ATCC)	HPB#1945 (4b, F)
001G23[1]	Similar to cystathionine β-lyase	0.70	0.05	0[4]	0	0.56	0.03
001H8[1]	Similar to molybdenum cofactor biosynthesis protein A	0.23	0.05	0.01	0.00	0.31	0.03
001N20[1]	Probable glucarate dehydratase	0.00	0.00	0.05	0.00	0.25	0.04
002B24[2]	Putative outer-membrane lipoprotein	0.07	0.12	0.14	0.00	0.81	0.01
002C8[2]	Transcriptional regulator, LysR family	0.11	0.22	0.47	0.00	0.89	0.05
002C16[1]	Hypothetical protein/putative cell surface protein	0.52	0.31	0.32	0.12	1.06	0.04
002E8[2]	RNA-directed DNA polymerase from retron ec67	0.05	0	0	0	0	0.33
003C6[2]	Anti-ECFsigma factor, ChrR	0.02	0.02	0.01	0.00	0.27	0.00
003D5[1]	23S rRNA	1.08	0.45	0.51	0.33	0.00	1.19
003D12[1]	Similar to conserved hypothetical protein	0.03	0.01	0.00	0.00	0.46	0.00
003F3[1]	Glycosyl transferase, family 65	0.01	0.03	0.00	0.00	0.08	0.21
003L20[1]	Subunit B of DNA gyrase	0.01	0.03	0.00	0.00	0.18	0.04
003M9[1]	Conserved hypothetical protein; β-glucosidase	0.14	0.46	0.36	0.00	0.00	0.70
003P10[1]	Conserved hypothetical protein	0.38	0.12	0.48	0.00	0.00	0.71

[1] Sequence identities were obtained using nucleotide BLAST (NCBI) on the target sequences.

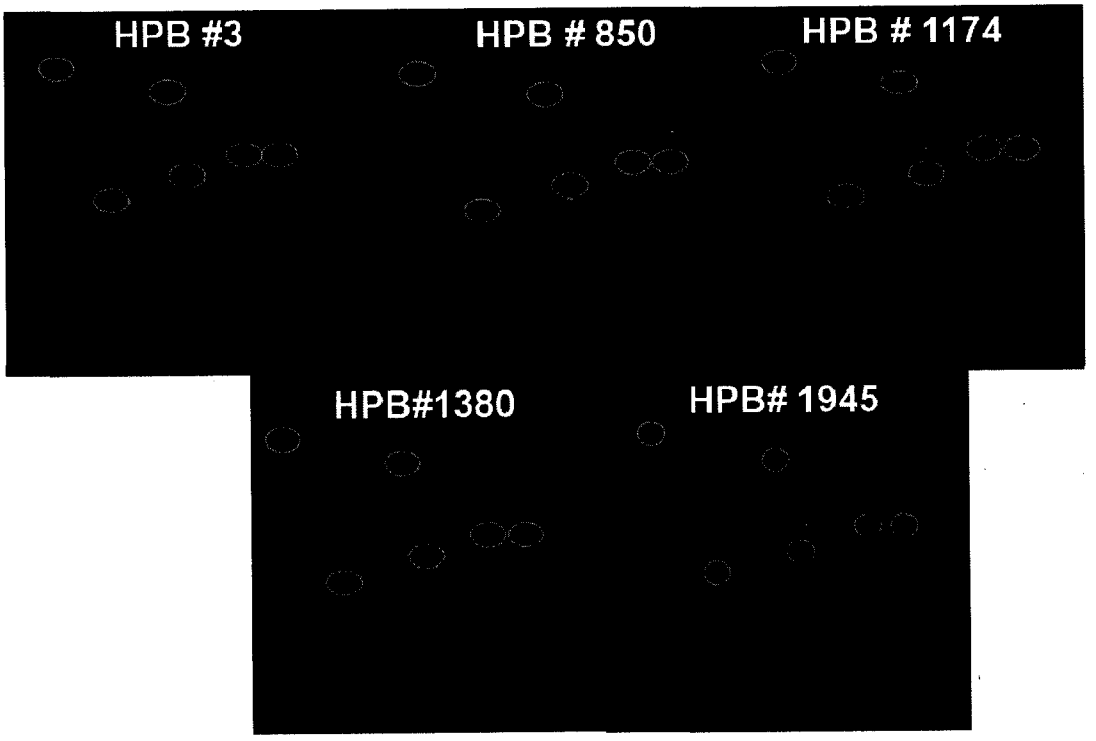
[2] Sequence homologies were obtained using ORF Finder (NCBI) and Protein BLASTX (NCBI) on the target sequences.

[3] C= Clinical Isolate; F= Food Isolate.

[4] A value of zero indicates that no hybridization was observed between the particular target and probe.

Figure 6. Scanned Images of *Listeria* Mixed Genome Array Version 2 Post-Hybridization to Isolates HPB# 3, 850, 1174, 1380 and 1945

Array spots that are circled in orange showed a consistently high % net intensity between isolates, while spots that are circled in purple showed differing net intensities upon analysis.


HPB #3
HPB # 850
HPB # 1174
HPB#1380
HPB# 1945

Several targets were selected for further investigation after their sequences were obtained and compared to nucleotide sequences in the NCBI Blast database. These targets included: 001-G23, similar to cystathionine β-lyase; 002-C16, a putative cell surface protein; 002-E8, a reverse transcriptase from retron ec67; and 003-F3, glycosyltransferase family 65. The target sequence 001-G23 appeared to be absent or significantly different in the serotype 4b isolates hybridized to the array when compared to the data of other serotypes. The target sequence 002-C16 appeared to be present in all of the isolates that were hybridized to the array, but showed weaker hybridization to a serotype 1/2b food isolate and a 4b food isolate. The target sequence 002-E8 was only present in two of the isolates that were hybridized to the array, a 3a clinical isolate and a 4b food isolate. Target sequence 003-F3 appeared to be present in some isolates based on array hybridization data, but the sequence appeared to be absent from a 4b clinical isolate, a 1/2b food isolate and showed extremely weak hybridization to a 3a clinical isolate.

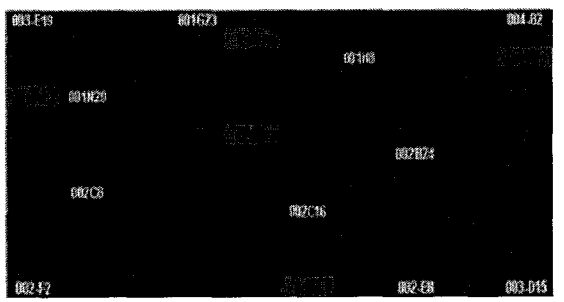
Array Version 3

The third version of the array contained 300 features in two subarrays (Figure 7), including 11 features from the array version 2 which showed either consistently high levels of hybridization or significantly different levels of hybridization between *L. monocytogenes* isolates. Spot Report Alien DNA (72) was not used in this array due to the problems that were occurring with cross-hybridization. Also, because the objective was not to obtain quantitative results for gene expression, but to compare hybridization patterns between isolates, the use of the Spot Report Alien DNA (72) was considered unnecessary. A complete list of features which showed different hybridization tendencies between

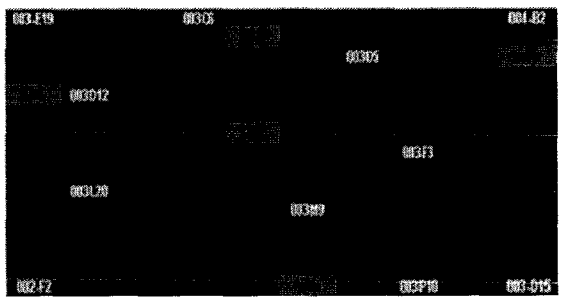
isolates of *L. monocytogenes* and the results of the BLAST searches using the obtained sequences of these features are shown in Table 8.

Figure 7. *Listeria* Mixed-Genome Array Version 3

Negative control spots can be seen in dark blue and purple. PCR products from the mixed-genome library are in light blue for subarray 1 and in purple for subarray 2. Targets with yellow text were also present in Version 2 of the array.



Subarray 1



Subarray 2

Table 8. Summary of Average Net Intensities of the Trimmed Means and Sequence Identities of Selected Features from *Listeria* Mixed Genome Array Version 3

Probe ID	BLAST Search	Average Net intensity of the Trimmed Mean			
HPB# (serotype, source)		**HPB# 1380 (1/2b, F)**	**HPB# 850 (4b, C)**	**HPB# 1945 (4b, F)**	**HPB# 1174 (4b, F)**
004 O10	Similar to homoserine O-acetyltransferase	-134.2[1]	51.0	141.3	29.8
004 O14	Similar to 6-phospho-β-glucosidase	661.6	63.3	254.5	914.2
004 O15	Putative aspartate aminotransferase	47.0	20.1	238.1	354.2
004 O17	Conserved hypothetical protein and acetyltransferase, GNAT family	117.3	-46.2	237.7	450.4
004 O20	Similar to A/G-specific adenine glycosylase	-161.1	-5.37	45.3	350.7
005 B17	PAP2 family protein	68.8	140.7	-42.3	412.3
005 C18	Similar to *B. subtiliss* SpoVG protein	ND[2]	-173.4	86.7	-25.2
005 E8	Cell division protein FtsW/RodA/SpoVE family	-87.3	-99.1	121.4	199.8
005 E10	Highly similar to *B. subtilis* YqfA protein	395.2	-19.5	464.0	1094.5
005 E23	Putative lipoprotein	-59.0	160.6	157.5	602.1
005 F5	Oxidoreductase, aldo/keto reductase family	37.5	-157.9	252.6	92.3
005 G12	*Iap* gene for protein p60	-61.0	21.7	254.6	131.0
005 I9	Peptidoglycan lytic protein p45	1.8	109.8	393.1	407.4

[1] A negative value indicates that hybridization did not occur between a particular isolate and the probe on the array. A higher positive value indicates a stronger signal net intensity in the array.

[2] ND = No data was available

Dot-blot Hybridization of Selected Targets from Mixed-Genome Array

To confirm the hybridization patterns seen in mixed-genome array experiments, dot-blot hybridizations were performed. This technique was used to confirm the presence or absence of selected target sequences. In this case, the fragmented genomic DNA was blotted on to a nylon membrane and the PCR products of the target sequences from the array were DIG-labelled and used as probes. A PCR product from one of the key virulence factors in *L. monocytogenes*, *hlyA*, was used as a positive control. The genomic DNA from all of the isolates showed a strong hybridization to this target. The results of the dot-blot hybridizations are shown in Table 9 and an image of a typical dot-blot in Figure 8. It should be noted that target 003-C6 did not show hybridization with any of the isolates in the dot-blot hybridizations. However, in the array data, hybridization occurred to HPB# 17 and weak hybridization to isolates HPB# 3, 850 and 2863. A similar situation was observed for target 002-C8. In the dot-blot hybridizations, a signal was only observed for strains HPB# 3, 17 and 2863, whereas in the array data, hybridization also occurred in isolates HPB#1380 and 1945 in addition to HPB# 3, 17 and 2863. For the target 002-C16, strong hybridization to HPB#17 occurred in array, but was not observed in the dot-blot hybridizations.

Table 9. Dot-blot Hybridization Results Using DIG-labelled PCR Products of Selected Targets of Interest from *Listeria* Mixed-Genome Array Version 2

	Probe ID						
HPB# (serotype, C/F[1])	***hlyA***	**002-C8**	**002-C16**	**002-E8**	**003-C6**	**003-D5**	**003-F3**
3 (4b, C)	+	+ (+)[4]	+ (+)	- (-)	- (+)	+ (+)	- (+)
17 (4c, ATCC)	+	+ (+)	- (+)	- (-)	- (+)	- (-)	- (+)
850 (4b, C)	+	-[2] (+)	+ (+)	- (-)	- (+)	+ (+)	- (-)
2863 (3a, C)	+	+ (+)	++[3] (+)	++ (+)	- (+)	+ (+)	- (+)
1380 (1/2b, F)	+	- (-)	++ (+)	- (+)	- (-)	++ (+)	- (-)
1945 (4b, F)	+	- (+)	++ (+)	++ (+)	- (-)	++ (+)	++ (+)
2184 (4a, C)	+	-	-	-	-	++	-
1174 (4b, F)	+	-	+	-	-	+	++

[1]C= Clinical isolate; F= Food isolate.

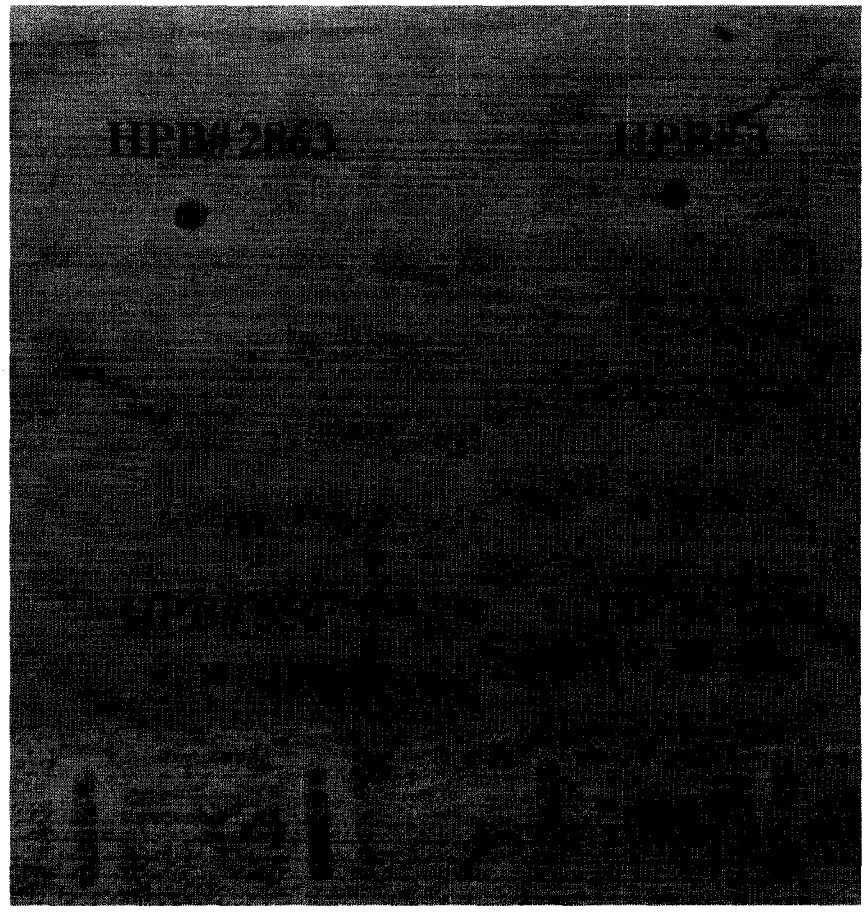
[2] Indicates a negative signal.

[3] ++: Signal was approximately 2-fold more intense

[4] Symbols in parentheses indicate if hybridization occurred in the *Listeria* mixed-genome array version 2. (+) indicates that hybridization occurred, (-) indicates that hybridization was not observed. Note that strains HPB# 2184 and 1174 were not tested in the *Listeria* mixed-genome array version 2.

Figure 8. A Typical Dot-Blot Hybridization Image Showing Hybridization of the Probe 003-D5 to the Genomic DNA of *L. monocytogenes* Isolates HPB#3, 850, 1380 and 2863

Hybridization of the *hlyA* positive control to the four isolates is shown



hlyA

PCR Screening of *L. monocytogenes* Isolates for Selected Target Sequences

A total of 45 isolates of *L. monocytogenes* (Table 5) were screened for the presence of target sequences that showed variability in hybridization tendencies between different isolates of the microorganism. Primers were designed for targets 001-G23, 002-E8, 002-C16 and 003-F3. While the target sequences 001-G23, showing nucleotide similarity to a cystathionine β-lyase, and 002-C16, to a putative cell surface protein, did not appear to hybridize with some *L. monocytogenes* isolates in the mixed-genome array, the PCR results and subsequent sequence analysis, indicated that the sequence was present in all of the screened isolates of *L. monocytogenes*. Figures 9A and D show representative PCR results from some of the isolates screened for the presence of target sequence 001-G23 and 002-C16, respectively.

When *L. monocytogenes* strains were screened for the presence of target sequence 002-E8, showing nucleotide similarity to a reverse transcriptase from retron ec67, the results of the PCR screening were consistent with the mixed-genome array data. The sequence was present in only 4 of the selected isolates: a serotype 4b food isolate, a serotype 1/2a clinical isolate, a serotype 4e clinical isolate and a serotype 3a clinical isolate. There was no correlation between the presence or absence of this target sequence and the source or serotype of the isolate. Representative PCR results from some of the isolates screened for the 002-E8 sequence is shown in Figure 9B.

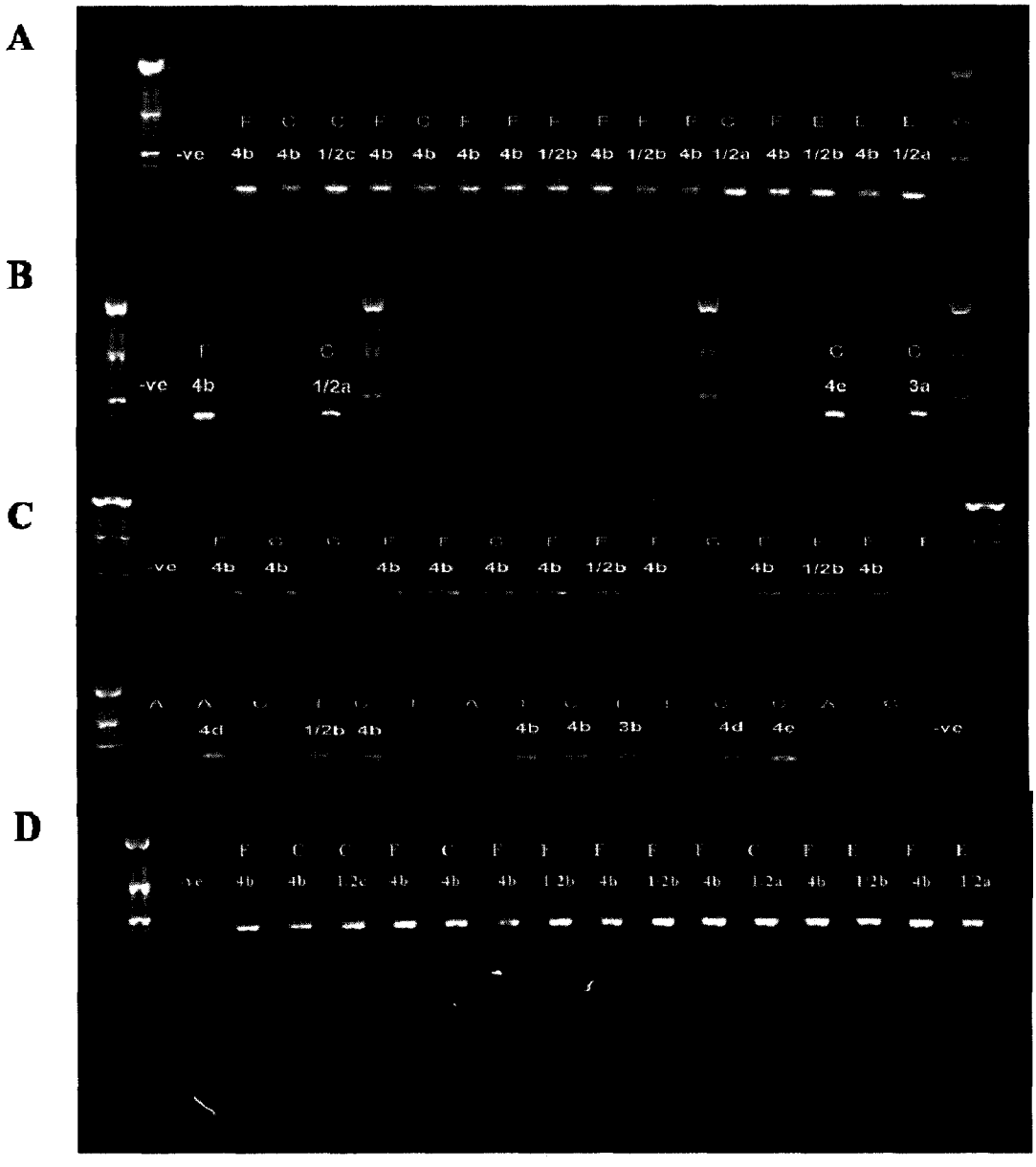
The most interesting results of the PCR screening was with target sequence 003-F3, showing nucleotide similarity to a glycosyltransferase family 65 (GT65). While the array data showed weak hybridization to a serotype 4b clinical isolate, a serotype 4b food isolate

and very weak hybridization to a serotype 3a clinical isolate, the PCR screening and subsequent sequence determination provided a more comprehensive analysis regarding the presence or absence of this particular sequence. Selected representative results of the PCR screening can be seen in Figure 9C. The most noticeable difference in the PCR results is the complete absence or presence of a very faint band for isolates of serotypes 1/2a, 1/2c or 3a. The sequences of the PCR products from all 45 screened isolates were analysed. The sequences obtained from serotypes 1/2b, 3b, 4a, 4b and 4e all showed 99-100% identity to the GT65 sequence from the *L. monocyotgenes* F2365 serotype 4b genome. On the other hand, the sequences obtained from the PCR products of the 1/2a, 1/2c and 3a serotypes all showed 99-100% identity to the gene *lmo2121*, annotated as being similar to a maltose phosphorylase, based on the *L. monocytogenes* EGD-e serotype 1/2a genome. The PCR products of these isolates showed only 89% identity to the GT65 sequence from the *L. monocytogenes* F2365 genome.

Figure 9. Results of PCR Screening of Selected *L. monocytogenes* Isolates for Feature Sequences (A) 001-G23, (B) 002-E8 (C) 003-F3 and (D) 002-C16 Identified from Differential Hybridization Tendencies of Features in *Listeria* Mixed-Genome Array Version 2

The serotype and source of the isolates are written above the PCR product. In 9C, the isolates with no PCR product or a fainter band are circled in red. A 100 bp DNA marker (150 ηg) (Roche Applied Science) can be seen in the first and last lanes.

C= Clinical isolate; F= Food isolate; E= Environmental isolate.

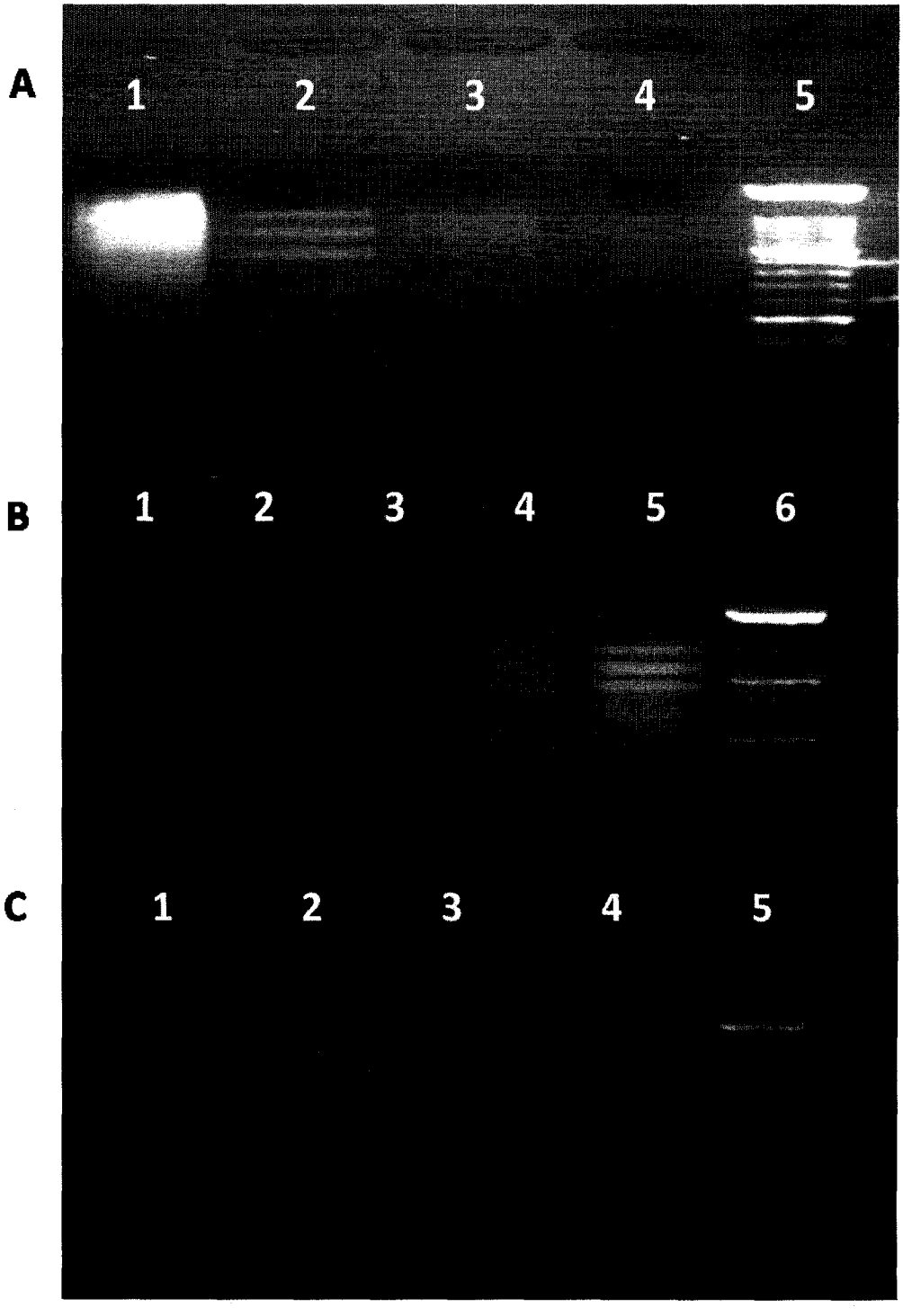

A
-ve 4b 4b 1/2c 4b 4b 4b 4b 1/2b 4b 1/2b 4b 1/2a 4b 1/2b 4b 1/2a
B
-ve 4b 1/2a 4e 3a
C
-ve 4b 4b 4b 4b 4b 4b 1/2b 4b 4b 1/2b 4b
4d 1/2b 4b 4b 4b 3b 4d 4e -ve
D
ve 4b 4b 1 2c 4b 4b 4b 1 2b 4b 1 2b 4b 1 2a 4b 1 2b 4b 1 2a

Suppressive Subtractive Hybridization

Initially, the genomes of a *L. innocua* isolate and a *L. monocytogenes* isolate were compared using suppressive subtractive hybridization (SSH). As seen in Figure 10A, many differences between the two genomes could be found using this method. The genomes of several *L. monocytogenes* isolates were also compared. Figure 10B and 10C shows the results of the comparison between *L. monocytogenes* serotypes 4a vs 4b and 3a vs 4b, respectively. PCR products representing genomic sequences that were different from one another were not observed in the intraspecies comparison. The hybridization conditions were altered to be less stringent in an attempt to find genomic sequence differences using SSH, however, the results remained negative. Due to time constraints, the SSH experiments were not further pursued.

Figure 10. Suppressive Subtractive Hybridization Results for Interspecies Comparison Between *L. monocytogenes* and *L. innocua* (A) and Intraspecies Comparison Between *L. monocytogenes* Isolates (B) and (C).

In Lanes 1 of 10A, B and C the SSH patterns of HPB# 583 (*L. innocua*) compared to HPB# 3 (*L. monocytogenes* 4b), HPB# 1631 (*L. monocytogenes* 4a) compared to HPB# 3 (*L. monocytogenes* 4b) and HPB# 2863 (*L. monocytogenes* 3a) compared to HPB# 3 (*L. monocytogenes* 4b) are shown respectively. Lanes 2 shows the SSH pattern for the *E. coli* positive controls, while Lanes 3 and 4 show the unsubtracted *E. coli* control and the negative control. Note Lane 3 in 10B is blank, and the controls are in Lanes 4 and 5. In Lane 5 for 10A and C and Lane 6 for 10B the 100 bp ladder (150ηg) is shown.


A
1
2
3
4
5
B
1
2
3
4
5
6
C
1
2
3
4
5

DISCUSSION

Listeria Mixed Genome Array Version 1

Version 1 of the mixed-genome array was not designed with the objective of identifying genomic features that may play a role in the tropism of the microorganism, but to optimize experimental conditions for hybridization, sample preparation and subsequent analysis. Using version 1 of the array, it was noted that at a hybridization temperature of 55°C (with the ramp-down temperature starting at 75°C), excessive amounts of air bubbles formed between the probe and the glass slide with the array. The formation of these air bubbles was compromising the results obtained during scanning of the array. In addition, due to the elevated starting temperature prior to final hybridization at 55°C, hybridization of the probe to the array was extremely weak. The hybridization temperature was lowered to 50°C and 45°C in an attempt to reduce bubble formation. However, at these lower temperatures, the hybridization conditions were not stringent enough and there was a greater amount of background and non-specific hybridization, preventing downstream analysis. To overcome these two issues, the organic solvent formamide was added to the hybridization buffer. Casey *et al.* (13) showed that increasing concentrations of formamide will linearly decrease the denaturation temperature of double-stranded DNA. For each percent of formamide added, the hybridization temperature can be decreased by 0.72°C. In our experiments, we added formamide to a final concentration of 40% to the hybridization buffer and after an initial ramp-down from 65°C, the hybridization conditions were set at 45°C. Under these conditions, the hybridizations were successful.

Cross-hybridization between the *L. monocytogenes* genomic DNA probe and the SpotReport® Alien® 7 and Alien® 3 cDNA PCR products from Stratagene was occurring

during hybridizations in Array Versions 1 and 2, respectively. For this reason, these external controls were removed from Array Version 3. These types of controls are often used in expression arrays where it is important to quantify the level of gene expression. They enable the user to assess the quality of their mRNA, quantitate dye ratios and act as negative hybridization controls (72).

Listeria Mixed Genome Array Version 2 Combined with PCR Screening Uncovers Multiple Genome Sequences Which Could Potentially Play a Role in the Tropism of *L. monocytogenes*

The second version of the *Listeria* mixed-genome array provided many genomic sequences that warranted further investigation both within this study and for studies in the future. Many features, shown in Table 7, gave differential hybridization tendencies with certain serovars of *L. monocytogenes*. Several of these features (001-G23, 002-E8, 002-C16 and 003-F3) were selected for confirmation using dot-blot and PCR-based methods. All products were sequenced on both the forward and reverse strands.

The target sequence 001-G23 was 99% identical at the nucleotide level to a cystathionine β-lyase. This enzyme is involved in sulphur metabolism, nitrogen metabolism, cysteine metabolism, methionine metabolism and selenoamino acid metabolism pathways in *L. monocytogenes* (40). In other microorganisms such as the Lactobacilli, the enzyme is also involved in sulphur metabolism, where the compound methanethiol is eventually released during the metabolic pathway. Interestingly, methanethiol is associated with desirable flavour in the production of Cheddar cheese (45, 79). The presence of this enzyme in *L. monocytogenes* may contribute to its ability to

survive and compete during the cheese manufacturing process with the other microorganisms essential for proper production of the cheese. Cysathionine β-lyase, encoded by *metC*, also plays a role in the virulence of *Salmonella enterica* serovar Typhimurium (26). When *metC* was inactivated, the virulence of the bacterium was attenuated in a mouse model. It was shown that the enzyme was a key player in the methionine and other sulphur-containing amino acid biosynthetic pathways. The authors concluded that the attenuation of the microorganism's ability to synthesize sulphur-containing amino acids, in an environment such as the serum of a mouse, where low concentrations of these amino acids are found, can result in the decreased virulence of the microorganism (26). A study into the effects of attenuating *metC* in *L. monocytogenes* would provide more information as to its potential role(s) in survival in foods, such as cheese, and to its potential role in the virulence of the microorganism. In the PCR screening of 45 isolates (Figure 9A), the sequence was present in all of the isolates and showed 99-100% identity to the cystathionine β-lyase sequence upon comparison. This work suggests that the presence of this gene sequence is not serotype-dependant. Given that this gene appears to be present in all of the isolates screened, it is unlikely that the presence of the sequence is playing a role in determining the particular niche of a clinical, food or environmental strain. However, this gene could be essential for the survival of all pathogenic *L. monocytogenes*.

The target sequence 002-C16, which showed nucleotide identity to a putative cell surface protein in *L. monocytogenes,* had weak hybridization to serotypes 4b and 1/2b food isolates and stronger hybridization to the clinical isolates of serotype 4c and 3a. Although the hybridization tendencies of different serovars of *L. monocytogenes* to this feature

appeared to be different in the microarray analysis, the PCR screening showed that the sequence was present in all 45 isolates. Sequence analysis showed 99-100% identity to the putative cell surface protein. Recently, new cell surface proteins, InlI and InlJ have been identified in *L. monocytogenes* (65). Both of these proteins contained the LPXTG motif, suggesting that they are in fact, cell surface proteins. While a deletion mutant of *inlJ* showed attenuated virulence in mice after intravenous and oral inoculation, the *inlI* deletion mutant did not show attenuated virulence in the mouse model (65). The 002-C16 sequence did not have identity to *inlI* or *inlJ*, but since 002-C16 shows identity to a putative cell surface protein, it may play a role in virulence in the host environment, as was the case for *inlJ*. The sequence was present in all of the *L. monocytogenes* isolates that were screened using PCR (BLAST searches did not show any sequence identities in non-pathogenic *Listeria* published genomes such as *L. innocua* or *L. welshimeri*), indicating that the presence of this genene sequence is not more common in certain *L. monocytogenes* isolates than others. If this genome sequence was to act as a virulence factor or if it enhanced adaptation to a particular environment for certain *L. monocytogenes* strains, it is most likely occurring at the transcriptional level and not at the genome level as all of the isolates that were screened possessed the gene sequence.

The next feature of interest in the *Listeria* mixed-genome array was feature 002-E8. It showed weak hybridization to a serotype 3a clinical isolate and stronger hybridization to a serotype 4b food isolate. Of the 45 isolates screened, only four contained the retron sequence (Figure 9B). There was no correlation between those isolates possessing the sequence and their source or serotype. The sequences of the PCR products using the primers for 002-E8 showed 99-100% identity to a reverse transcriptase enzyme from retron

ec67. Retron DNA is usually associated with prophage DNA and is found in a wide variety of bacterial genomes (43). Retrons synthesize an unusual type of satellite DNA referred to as msDNA. MsDNA is a complex of DNA, RNA and protein, and is produced in hundreds of copies per cell, but its function remains unknown. However, there is some evidence that suggests that the presence of retron DNA within a genome can result in the generation of multiple repeated copies of msDNA in the genome, resulting in a higher frequency of spontaneous mutations. While it has long been known that reverse transcription has effects on the genomes of eukaryotes, it is now being proposed that perhaps they are playing a role in the changes within prokaryotic genomes (43). Further investigations into the role of the retrons in the *L. monocytogenes* genome may provide insight into their potential function in this microorganism. Disruption or deletion of these genomic sequences could provide an indication as to whether these retrons enhance or decrease the microorganism's ability to infect a host or survive in different environments.

The final feature of the *Listeria* mixed-genome array that was investigated was feature 003-F3. This feature showed 99% identity to a glycosyltransferase family 65 sequence in the *L. monocytogenes* F2365 (serotype 4b) genome. The PCR screening and subsequent sequence analysis in the 45 isolates revealed that the sequence was not present in seroype 1/2a, 1/2c and 3a isolates (Figure 9C). Interestingly, while these serotypes can cause clinical disease, they are most commonly isolated from food sources. In these serovars, if a PCR product was obtained, the sequence showed 99-100% identity to a maltose phosphorylase sequence in the *L. monocytogenes* EGD (serotype 1/2a). In all other serotypes, the PCR product showed a 99-100% identity to the glycosyltransferase family 65 sequence.

Glycosyltransferases are a large family of enzymes which catalyze the transfer of a sugar residue to an acceptor molecule during both degradation and biosynthesis of polysaccharides, glycoproteins and glycolipids (31). They are involved in pathways ranging from sugar metabolism to cell wall biosynthesis. Since this genome sequence appeared to be absent in *L. monocytogenes* serotypes 1/2a, 1/2c and 3a, it was selected for further investigation for its potential role in the tropism of the microorganism. These results and a more in-depth discussion about the glycosyltransferases are included in Chapter 3.

Several sequences of interest were also identified in *Listeria* mixed-genome array version 3 (Table 8). Due to time constraints, the sequences identified in the array were not confirmed for their presence or absence by dot-blot hybridization or PCR. However, the data obtained in the array provides a starting point for future studies on the microorganism.

Dot-Blot Hybridization is a Useful Technique to Confirm the Presence or Absence of Genome Sequences

To confirm the presence or absence of some of the genome features that were identified using the mixed genome array, dot-blot hybridization was used. While this method was not used to quantify the intensity of the hybridization signals, it was useful to determine the presence or absence of hybridization between a particular feature and the genomic DNA of selected *L. monocytogenes* isolates. In this case, the fragmented genomic DNA that was used as a probe in the microarray experiments was spotted onto the nylon membrane and became the feature. The feature of interest from the microarray analysis was labelled with DIG and became the probe for this experiment. This allowed for faster throughput when screening for the presence or absence of a particular genome sequence in

several isolates of *L. monocytogenes*. For this reason, two more isolates of *L. monocytogenes*, HPB# 1174 (serotype 4b, food isolate) and HPB# 2184 (serotype 4a, clinical isolate) were screened for the presence or absence of sequences in dot-blot hybridizations. The *hlyA* positive control showed hybridization to all of the isolates that were tested. The majority of the dot-blot hybridization results were consistent with the hybridization tendencies seen in the microarray data. However, some of the features that showed weak hybridization in the microarray did not show any detectable hybridization in the dot-blot.

The dot-blot hybridization technique is useful in confirming array results when the hybridization of the feature to the probe is strong and provides a high net intensity signal in the microarray analysis. However, when the hybridization was weaker in the microarray, the dot-blot hybridization technique was unable to confirm the presence or absence of the nucleic acid sequences. This could be due to the stringent conditions in the dot-blot technique, insufficient incorporation of the DIG reagent into the probes as compared to the Cy3 incorporation in the genomic DNA used in the microarray, or the detection method for a signal in the dot-blot as compared to the microarray. In the microarray analysis, the signal intensity is detected by laser excitation of the Cy3 dye. This method is extremely sensitive and the laser power and sensitivity can be increased to obtain maximum excitation and detection of the fluorophore. However, in the dot-blot analysis, hybridization was colormetrically detected with the naked eye and therefore was not as sensitive as the microarray detection. In summary, the dot-blot technique used in this study was useful in screening *L. monocytogenes* isolates showing strong hybridization signals in the microarray, but was not powerful enough to detect those with weaker hybridization signals.

PCR Screening of *L. monocytogenes* Isolates is Important for Confirmation of Results Obtained in the Microarray Analysis

To confirm the data indicating the presence or absence of particular genome sequences obtained using mixed-genome arrays, PCR was used to screen for the presence or absence of the sequence. Sequence analysis of the PCR product was also used to determine if the nucleotide sequences for the particular feature were the same for all of the isolates screened.

While the microarray analysis and dot-blot hybridization indicated that features 001-G23, 002-C16 and 002-E8 were absent from certain isolates of *L. monocytogenes*, PCR screening indicated that they were present (Figure 9A, B and D). When the nucleotide sequences of all 45 isolates were compared, they all showed 99-100% identity to the PCR product used as a feature in the microarray. Feature 003-F3, which had 100% identity to a glycosyltransferase family 65 in the *L. monocytogenes* F2365 genome, did not hybridize to a serotype 1/2b food isolate or a 4b clinical isolate in the microarray experiments. However, the PCR screening and subsequent sequence analysis indicated that the sequence was present in all serotype 3b, 4b, 4d, 4e and 1/2b isolates, but not in isolates of serotype 1/2a, 1/2c or 3a (Figure 9C).

The results of the PCR screening for features identified as present or absent in the microarray experiments confirm the importance of using other methodologies to confirm data obtained using microarray-based technologies. In expression microarray studies, RT-PCR is often used as a means of confirming differences observed in gene expression. PCR was used in this study to confirm the microarray results because it is a fast and reliable

method and the primers for the feature sequences in the microarray were relatively straightforward to design. The PCR experiments were able to amplify genome sequences in isolates that appeared to have a negative hybridization in both the dot-blot and microarray experiments. This may be due, in part, to the kinetics of the hybridization reaction. In the PCR reactions, hybridization was occurring between a primer of approximately 20 bp and the genomic DNA. In contrast, in the microarray and dot-blot experiments, hybridization was occurring between sonicated genomic DNA fragments ranging in size from approximately 200-1200 bp and the PCR product feature which ranged in size from 500-1300 bp. Hybridization of small DNA fragments is more kinetically favourable than hybridization of large fragments. This is a potential explanation as to why the PCR reaction was able to amplify the genome sequences that appeared to be absent in the microarray analysis.

Sonicated genomic DNA was used in the microarray and dot-blot hybridization experiments. During sonication, the genomic DNA is randomly broken into small fragments. It is possible that during the sonication process, the genomic DNA was fragmented in such a way that the region that would have hybridized to an array target was fragmented into several pieces, therefore rendering hybridization unfavourable.

In this work, PCR screening of multiple serovars of *L. monocytogenes* for specific genomic sequences was essential for drawing accurate conclusions regarding their presence or absence in clinical, food or environmental strains.

Chapter 3: Investigation into the Potential Biological Role of Glycosyltransferase family 65 in the Tropism of *L. monocytogenes* for a Specific Niche

The information in this chapter was presented in poster format at the 2007 Federal Food Safety and Nutrition Research Meeting (Winnipeg, MB) and the 2007 Health Canada Science Forum (Ottawa, ON). The information was presented in oral format at the 2007 International Association of Food Protection 94th Annual Meeting (Lake Buena Vista, FL).

A manuscript of this work is being prepared and submitted for publication to a peer-reviewed journal.

INTRODUCTION

Glycosyltransferases

Glycosyltransferases are a family of enzymes responsible for the transfer of sugar (glycosyl) residues to an acceptor during degradation and biosynthesis of polysaccharides, glycoproteins and glycolipids (31). In *Listeria monocytogenes*, the cell wall contains large amounts of teichoic acid, covalently linked to peptidoglycan with diversity of the teichoic acid arising from glycosidic substitutions. For example, the teichoic acid of serogroups 1/2 and 3 is composed of a polyribitol phosphate with N-acetylglucosamine (Glc-NAc) with rhamnose substitutions on the ribitol sugar. On the other hand, *L. monocytogenes* serogroup 4 isolates contain galactose and/or glucose substituents on the Glc-NAc component of the teichoic acid (62). Serogroup classifications of *L. monocytogenes* are based on the flagellar and somatic surface antigens which are mainly carbohydrate based with an association to the teichoic acid (62). Glycosyltransferases are known for their role in catalyzing the polymerization step of cell wall biosynthesis, and are an accessible drug target for bacteria that are resistant to antibiotics (46). Glycosyltransferase family 65 (GT65) is annotated as an EC 2.4.1.8 enzyme by the Kyoto Encyclopedia of Genes and Genome (KEGG), that acts as a maltose phosphorylase in starch and sugar metabolism (40). In our experiments, we used PCR to screen *L. monocytogenes* isolates for the presence of *lmo2121* (similar to maltose phosphorylase) and the GT65 sequences. *Lmo2121* encodes for a gene in *L. monocytogenes* EGD that is annotated as similar to a maltose phosphorylase. The sequence of this gene is similar to the GT65 consensus sequence. Since GT65 was annotated as a maltose phosphorylase, fermentation of maltose by *L. monocytogenes* isolates was also

investigated. The susceptibility of isolates to cell-wall synthesis inhibiting antibiotics was also observed, to investigate if GT65 was potentially involved in cell-wall biosynthesis.

Promadej *et al.* (62) found a gene, *gtcA*, which was essential in the serotype 4 groups of *L. monocytogenes* for the addition of the correct amounts of glucose and galactose onto the teichoic acid of the cell wall. The cell wall of Gram-positive bacteria is very complex with many steps in the degradation and synthesis processes, including the generation of teichoic acids specific to the particular serotype of *L. monocytogenes*. Cell wall and TA modifications in *L. monocytogenes* are thought to play a major role in the pathogenesis and perhaps even a role in environmental adaptation (62). It is known that the glycosyltransferases play a role in cell wall biosynthesis and sugar metabolism (46), but the specific function of GT65, first discovered in our study to be present only in some serotypes of *L. monocytogenes,* remains unknown.

Antibiotic Susceptibility of *L. monocytogenes*

In order to investigate the role of GT65 in cell wall biosynthesis, isolates of *L. monocytogenes* were grown in the presence of the cell wall synthesis inhibiting antibiotics ampicillin and vancomycin. Unlike some other microorganisms, antimicrobial resistance in *L. monocytogenes* does not appear to be a problem at the present time (57). In 2003, Safdar and Armstrong (66) performed antimicrobial susceptibility tests on 84 clinical isolates of *L. monocytogenes* using penicillin, ampicillin, erythromycin, tetracycline, gentamicin and clindamycin. With the exception of clindamycin, all of the isolates were equally susceptible. However, resistance to clindamycin was observed in most isolates, indicating

that resistance to this antimicrobial could become a problem in the future and that perhaps this particular antimicrobial should not be used for the treatment of listeriosis (66).

To treat a case of listeriosis, a cocktail of ampicillin and gentamicin is usually used, and as a secondary treatment, the patient may be prescribed erythromycin, vancomycin or fluoroquinolones (57). In our study, we chose to use ampicillin and vancomycin treatments to study their effect on the growth of 45 *L. monocytogenes* isolates. These antibiotics were chosen because of their use in clinical applications and their mechanism of action on the biosynthesis of the cell wall, creating a potential link to the action of the GT65 enzyme.

Homologous Recombination and the Vector pAUL-A

To gain insight into the role of the GT65 enzyme in *L. monocytogenes*, an isogenic deletion mutant for this gene was created. Using homologous recombination, with the vector pAUL-GT (derived from pAUL-A) containing an insert from sequences surrounding the GT65 gene and a few nucleotides from the 5' and 3' ends of the gene, the GT65 sequence was removed from the genome of *L. monocytogenes* HPB#3 (serotype 4b).

To clone DNA fragments in *Listeria*, special cloning vectors must be used (68). In our study, we chose to use the pE194-based shuttle vector pAUL-A to create the GT65 deletion mutant. This vector was kindly provided by Dr. Trinad Chakraborty from the Institut für Medizinische Mikrobiologie in Giessen, Germany. The pAUL-A vector has been used in creating mutants for many different genes in *L. monocytogenes,* with one of the most notable being the *prfA* gene. In 1992, Chakraborty *et al.* (15), created the pAUL-A shuttle vector and used it to construct an insertion mutant in the *prfA* gene. It was this study

that highlighted for the first time the essential role of *prfA* in the regulation of virulence genes in *L. monocytogenes*.

MATERIALS AND METHODS

Amplification of *lmo2121*-Similar to Maltose Phosphorylase

Primers were designed for PCR amplification for part of the *lmo2121* (Similar to maltose phosphorylase) sequence found in the 1/2a, 1/2c and 3a serotypes of *L. monocytogenes*. Since this sequence is highly similar to the lmof2365_2155 (GT65) sequence, the primers were designed for the regions of the genes that varied the most between the two sequences. The primer sequences and properties can be seen in Table 10. The boil preps of the 45 *L. monocytogenes* strains described in Table 5 were used as a DNA template for the PCR reactions. The PCR conditions were as follows: 2 min at 95°C followed by 35 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 min. After the 35 cycles, an extra extension of 72°C for 10 min was added. Samples were held at 4°C until analysis by agarose gel electrophoresis. Each 100 µL PCR reaction contained 1 µL of boil prep DNA, ddH_2O up to 100 µL, 1X PCR buffer for Promega GoTaq® Flexi DNA Polymerase (Fisher Scientific), 2 mM $MgCl_2$ (Fisher Scientific), 0.2 mM dNTPS (Roche Applied Science), 1 µM of primers (Sigma-Genosys Inc) and 5 Units of Promega GoTaq® Flexi DNA Polymerase (Fisher Scientific). PCR products were analyzed on a 1% agarose gel in 1X TBE buffer, run at 100 V for 30 minutes. All PCR products were sent to DNA Landmarks Inc. (St-Jean-sur-Richelieu, QC) for sequencing. The ABI trace files were analysed using Chromas (Technelysium Pty Ltd) software and the consensus sequences were generated using Tera Term Software. Homologies to the consensus sequences were determined by using the Nucleotide MegaBLAST algorithm and searching the Nucleotide Collection and Reference Genomic Sequences databases to ensure that the correct sequences had been amplified.

Table 10. Primers Used to Amplify *lmo2121*-Similar Maltose Phosphorylase

Primer Name	Tm (°C)	Sequence
lmo2121 F	60.00	5'-TTGGTCCTGTTACACCGTGA-3'
lmo2121 R	59.97	5'-ATAAGACCAATGCGGCAAAC-3'

Growth of *L. monocytogenes* on Purple Maltose Agar

The utilization of maltose by *L. monocytogenes* was investigated by observing its growth on Purple Agar Plates containing of maltose. Maltose fermentation of 45 isolates (Table 5) and the GT65 deletion mutant were observed. Each isolate was transferred from frozen stock to a TA plate and grown at 30°C for 48 h. A colony was picked from this plate and streaked onto one-quarter of a purple agar maltose plate and incubated at 30°C for 48 h. Four isolates were streaked per plate, with each one occupying one-quarter of the plate. The plates were visually examined after 24 and 48 h for maltose usage. The assessment was qualitative, with a positive maltose usage changing the agar from purple to yellow due to a change in pH from the sugar fermentation. If the isolate was negative for maltose fermentation, the agar remained purple after 48 h.

Growth of *L. monocytogenes* in Antibiotic Media

To investigate the differences in *L. monocytogenes* isolates with and without a GT65 gene, the isolates were grown in antibiotic broth. Two cell wall biosynthesis inhibiting antibiotics were selected (ampicillin and vancomycin) for their inhibition of peptidoglycan synthesis.

The 45 isolates and their properties selected for antibiotic screening can be seen in Table 5. The GT65 deletion mutant was also tested for its susceptibility to ampicillin and vancomycin, using the same conditions. A single colony was picked from a TA plate and grown overnight in five mL of Mueller-Hinton Broth at 37°C. Mueller-Hinton Broth containing various concentrations of ampicillin (0.25, 0.1, 0.085 and 0.055 μg/mL) or vancomycin (1, 0.75, 0.5 and 0.25 μg/mL) was inoculated with the overnight growth of *L.*

monocytogenes to a final concentration of 1×10^5 CFU/mL. Cells were grown for 18 h at 37°C in the antibiotic medium. After 18 h, an OD_{620} measurement was taken and converted into a CFU/mL value to measure cell growth. Duplicate experiments were performed for each strain, the average OD_{620} was calculated and then converted into a CFU/mL value.

Creation of a Glycosyl Transferase Deletion Mutant

A deletion mutant of the GT65 gene sequence in *L. monocytogenes* EGD was created via homologous recombination. The insert was created via overlap PCR and contained the first few base pairs of the 5' end of the gene and the last few base pairs of the 3' end of the gene, as well as approximately 400 bp of DNA corresponding to the sequence directly upstream and downstream of the GT65 sequence. The primer sequences and characteristics can be seen in Table 11. Figure 11 describes the generation of the insert by overlap PCR.

The first PCR reactions were carried out separately and generated fragments that were approximately 400 bp in size, named GTdel1 and GTdel2. To generate the final PCR product, GTdel3, to be used as the insert for homologous recombination, GTdel1 and GTdel2 were used in an overlap PCR. Approximately 200 ηg of GTdel1 and GTdel2 were combined and used as DNA template and self-primers for the PCR reaction to create the GTdel3 product. The 100 µL PCR reaction contained ddH_20 up to 100 µL, 1X PCR buffer for Platinum® Taq DNA Polymerase (Invitrogen Canada Inc.), 1.5 mM $MgCl_2$ (Invitrogen Canada Inc.), 0.2 mM dNTPS (Roche Applied Science) and 5U of Platinum® Taq DNA Polymerase (Invitrogen Canada Inc.). The PCR conditions were as follows: 2 min at 95°C followed by 15 cycles of 95°C for 30 seconds, 55°C for 30 seconds and 72°C for 1 min.

Following the initial 15 cycles, 1 μM each of primers GTdelA and GTdelD were added followed by an additional 35 cycles of PCR as described above and a final extension at 72 °C for 10 min. The PCR products were held at 4°C until analysis by agarose gel electrophoresis. The PCR products were analyzed on a 1% agarose gel in 1X TBE buffer, run at 100V for 30 min.

The 846 bp fragment corresponding to the PCR product GTdel3 was excised from the gel and the DNA was extracted using the QIAEX II Agarose Gel Extraction Kit (Qiagen). The manufacturer's protocol was followed and the purified DNA was eluted in ddH_20.

GTdel3 was subcloned into the pCR®4-TOPO vector (Invitrogen Canada Inc.) and the vector containing the insert was transformed into *E. coli* One Shot® TOP10 chemically competent cells, as per the manufacturers protocol. Transformants were incubated at 37°C for 1 h in SOC medium and then grown on LB plates containing 50 μg/mL of kanamycin for 24 h. Clones were screened for presence of the GTdel3 insert by PCR using M13 as well as GTdelA and D primers. Positive clones were selected and the TOPO vector with GTdel3 insert was isolated. The insert was removed via digestion with BamHI and HindIII and ligated into the pre-digested pAUL-A vector overnight at 16°C. The pAUL-GT vector was used for electroporation into electrocompetent *L. monocytogenes* (HPB#3, serotype 4b) at settings of 25 μF, 2.5 kV, 200 Ω. The transformants were incubated in 1 mL of SOC medium at 30°C for 3 h then spread onto LB agar plates containing 5 μg/mL erythromycin. The plates were incubated at 30°C for 3-6 days, with small colonies appearing after 3 days. Colonies were selected and grown in 5 mL of LB broth with 5 μg/mL of erythromycin at 30°C overnight in a shaking incubator (Sheldon Manufacturing Inc., Cornelius, OR, USA),

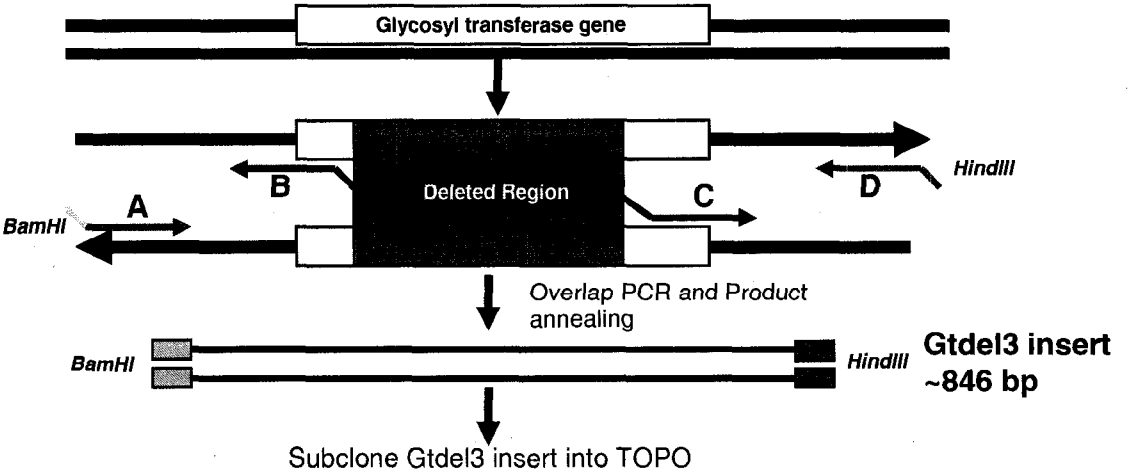
set to 500 rpm , and a stab culture of each colony was created on an LB agar plate with 5 µg/mL of erythromycin. The plasmid was isolated from the broth cultures and screened for the presence of the insert using M13 forward and reverse primers. Clones containing the recombinant insert were streaked from the stab culture onto an LB agar plate containing 5 µg/mL of erythromycin and incubated overnight at 42°C (the non-permissive temperature for replication of the plasmid). This step was repeated twice, with a pure colony grown at 42°C. The pure colony obtained after three passages at 42°C was inoculated into 20 mL of LB broth and incubated in a shaking incubator (Sheldon Manufacturing Inc.) set at 100 rpm, overnight at 30°C. A total of 20 µL of overnight culture was inoculated into 40 mL of LB broth and incubated with shaking at 100 rpm overnight at 30°C. The overnight culture was once again diluted, 40 µL of culture into 40 mL of LB broth, and then incubated overnight at 42°C in a shaking incubator (Sheldon Manufacturing Inc.), set to 100 rpm. The cells from the overnight culture were spread onto LB agar and incubated at 37°C. Colonies were screened for erythromycin resistance and those that were susceptible to the antibiotic were considered to be deletion mutants of GT65. The genomic DNA was extracted (see Chapter 2 methods) and analyzed to confirm that the GT65 gene had been deleted by using primers flanking the deleted region and primers targeting the deleted region.

Table 11. Primers Used to Create the GTdel3 PCR Product for the Generation of the GT65 Deletion Mutant

Primer Name	Sequence (5'-3')	Tm (°C)
GTdelA	[1]GGATCCGTGCAATATTTGTTCC	61
GTdelB	[2]CATAGTACGACGTAACTAAATGAAAAATG	60.5
GTdelC	[2]GTCGTACTATGTGCCATACTAAAACC	59.1
GTdelD	[3]AAGCTTAAATAACGTCAGGCGTGT	61.9

[1] BamHI recognition site.
[2] Complementary sequences on GTdelB and GTdelC primers.
[3] HindIII recognition site.

Figure 11. Generation of a GT65 Deletion Mutant Using the pAUL-A Vector and Homologous Recombination



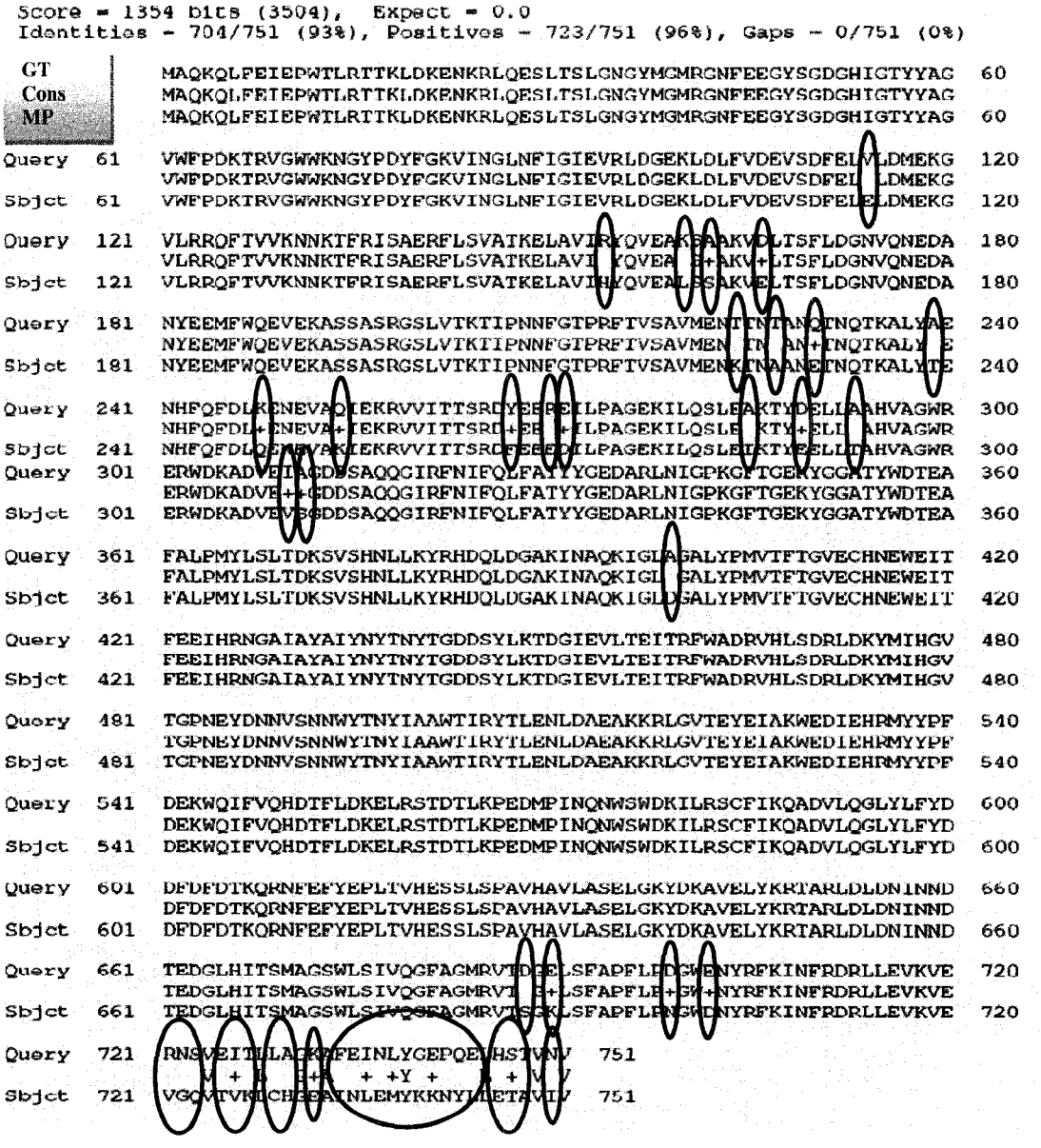
Subclone Gtdel3 insert into TOPO
Cut out Gtdel3 from TOPO vector and clone into pAUL-A
Transform pAUL-GTdel3 insert into *E. coli* TOP10
Extract plasmid and transform into *L. monocytogenes*

RESULTS

Comparison of Gene and Amino Acid Sequences of *lmof2365_2155* (Glycosyltransferase family 65) and *lmo2121* (Similar to maltose phosphorylase)

The coding sequence of the gene encoding GT65 from the *L. monocytogenes* F2365 serotype 4b genome and *lmo2121* from the *L. monocytogenes* EGD serotype 1/2a genome were compared using the BLAST from NCBI. Both genes are 2.2 kbp in size and are encoded on the reverse strand. *lmo2121* starts at base pair 2,201,719 and ends at base pair 2,203,980 in the *L. monocytogenes* EGD genome, while the GT65 (*lmof2365_2155*) gene starts at base pair 2,191,682 and ends at base pair 2,193,943 in the *L. monocytogenes* F2365 genome. The coding sequences of the two genes showed 92% identity. The first third and the 3' end of the coding sequence showed the greatest and least variability, respectively. When the amino acid sequences of the two genes were compared, they showed a 93% identity. The majority of the differences are located within the last 30 amino acids at the C-terminus end of the amino acid sequence. However, there are other significant amino acid differences in the first half of the sequence. The alignment of the amino acid sequences are shown in Figure 12.

Figure 12. Amino Acid Sequence Alignment of GT65 (lmof2365_2155) and lmo2121
Differences in the amino acid sequence are circled in red.
GT = lmf2365_2155 from *L. monocytogenes* genome F2365, serotype 4b; MP = lmo2121 from *L. monocytogenes* genome EGD, serotype 1/2a; Cons = amino acid consensus sequence.


Score = 1354 bits (3504), Expect = 0.0
Identities = 704/751 (93%), Positives = 723/751 (96%), Gaps = 0/751 (0%)
GT
Cons
MP
MAQKQLFEIEPWTLRTTKLDKENKRLQESLTSLGNGYMGMRGNFEEGYSGDGHIGTYYAG 60
MAQKQLFEIEPWTLRTTKLDKENKRLQESLTSLGNGYMGMRGNFEEGYSGDGHIGTYYAG
MAQKQLFEIEPWTLRTTKLDKENKRLQESLTSLGNGYMGMRGNFEEGYSGDGHIGTYYAG 60
Query 61 VWFPDKTRVGWWKNGYPDYFGKVINGLNFIGIEVRLDGEKLDLFVDEVSDFELVLDMEKG 120
VWFPDKTRVGWWKNGYPDYFGKVINGLNFIGIEVRLDGEKLDLFVDEVSDFEL LDMEKG
Sbjct 61 VWFPDKTRVGWWKNGYPDYFGKVINGLNFIGIEVRLDGEKLDLFVDEVSDFELELDMEKG 120
Query 121 VLRRQFTVVKNNKTFRISAERFLSVATKELAVIRYQVEAKSAAKVELTSFLDGNVQNEDA 180
VLRRQFTVVKNNKTFRISAERFLSVATKELAVI YQVEA +AKV+LTSFLDGNVQNEDA
Sbjct 121 VLRRQFTVVKNNKTFRISAERFLSVATKELAVIHYQVEALSAKVELTSFLDGNVQNEDA 180
Query 181 NYEEMFWQEVEKASSASRGSLVTKTIPNNFGTPRFTVSAVMENTTNTANQTNQTKALYAE 240
NYEEMFWQEVEKASSASRGSLVTKTIPNNFGTPRFTVSAVMEN TN AN+TNQTKALY E
Sbjct 181 NYEEMFWQEVEKASSASRGSLVTKTIPNNFGTPRFTVSAVMENKTNAANETNQTKALYIE 240
Query 241 NHFQFDLRCNEVAQIEKRVVITTSRDYEEPEILPAGEKILQSLEAKTYDELLAAHVAGWR 300
NHFQFDL+CNEVA+IEKRVVITTSRD+EEP+ILPAGEKILQSLE KTY+ELL AHVAGWR
Sbjct 241 NHFQFDLQCNEVARIEKRVVITTSRDFEEPDILPAGEKILQSLEIKTYEELLTAHVAGWR 300
Query 301 ERWDKADVEIACDDSAQQGIRFNIFQLFATYYGEDARLNIGPKGFTGEKYGGATYWDTEA 360
ERWDKADVE+ CDDSAQQGIRFNIFQLFATYYGEDARLNIGPKGFTGEKYGGATYWDTEA
Sbjct 301 ERWDKADVEVTCDDSAQQGIRFNIFQLFATYYGEDARLNIGPKGFTGEKYGGATYWDTEA 360
Query 361 FALPMYLSLTDKSVSHNLLKYRHDQLDGAKINAQKIGLAGALYPMVTFTGVECHNEWEIT 420
FALPMYLSLTDKSVSHNLLKYRHDQLDGAKINAQKIGL GALYPMVTFTGVECHNEWEIT
Sbjct 361 FALPMYLSLTDKSVSHNLLKYRHDQLDGAKINAQKIGLIGALYPMVTFTGVECHNEWEIT 420
Query 421 FEEIHRNGAIAYAIYNYTNYTGDDSYLKTDGIEVLTEITRFWADRVHLSDRLDKYMIHGV 480
FEEIHRNGAIAYAIYNYTNYTGDDSYLKTDGIEVLTEITRFWADRVHLSDRLDKYMIHGV
Sbjct 421 FEEIHRNGAIAYAIYNYTNYTGDDSYLKTDGIEVLTEITRFWADRVHLSDRLDKYMIHGV 480
Query 481 TGPNEYDNNVSNNWYTNYIAAWTIRYTLENLDAEAKKRLGVTEYEIAKWEDIEHRMYYPF 540
TGPNEYDNNVSNNWYTNYIAAWTIRYTLENLDAEAKKRLGVTEYEIAKWEDIEHRMYYPF
Sbjct 481 TGPNEYDNNVSNNWYTNYIAAWTIRYTLENLDAEAKKRLGVTEYEIAKWEDIEHRMYYPF 540
Query 541 DEKWQIFVQHDTFLDKELRSTDTLKPEDMPINQNWSWDKILRSCFIKQADVLQGLYLFYD 600
DEKWQIFVQHDTFLDKELRSTDTLKPEDMPINQNWSWDKILRSCFIKQADVLQGLYLFYD
Sbjct 541 DEKWQIFVQHDTFLDKELRSTDTLKPEDMPINQNWSWDKILRSCFIKQADVLQGLYLFYD 600
Query 601 DFDFDTKQRNFEFYEPLTVHESSLSPAVHAVLASELGKYDKAVELYKRTARLDLDNINND 660
DFDFDTKQRNFEFYEPLTVHESSLSPAVHAVLASELGKYDKAVELYKRTARLDLDNINND
Sbjct 601 DFDFDTKQRNFEFYEPLTVHESSLSPAVHAVLASELGKYDKAVELYKRTARLDLDNINND 660
Query 661 TEDGLHITSMAGSWLSIVQGFAGMRVTDDELSFAPFLPDGWENYRFKINFRDRLLEVKVE 720
TEDGLHITSMAGSWLSIVQGFAGMRVT +LSFAPFLP+GW+NYRFKINFRDRLLEVKVE
Sbjct 661 TEDGLHITSMAGSWLSIVQGFAGMRVTSGKLSFAPFLPNGWDNYRFKINFRDRLLEVKVE 720
Query 721 RNSVEITLLAGKAFEINLYGEPQEVHSTVNV 751
V + L +R + +Y + + V V
Sbjct 721 VGQVTVKCHEGRTNLEMYKKNYLETAVIV 751

PCR Screening of *L. monocytogenes* isolates for the presence of *lmo2121*

Forty-five isolates of *L. monocytogenes* were screened by PCR for the presence of the *lmo2121* gene. The PCR results were all positive, with every isolate having a product approximately 850 bp in size. However, the sequence analysis of the PCR products indicated that the *lmo2121* sequence was only present in the serotype 1/2a, 1/2c and 3a isolates. The PCR products of all other serotypes had a 99-100% sequence identity to the GT65 gene sequence from the genome of *L. monocytogenes* F2365, serotype 4b, and only 89% identity to the *lmo2121* gene sequence from the genome of *L. monocytogenes* EGD, serotype 1/2a. These results are summarized in Table 12.

Table 12. Sequence Identities of 45 PCR Products Generated Using Primers Designed for *lmo2121*

HPB#	Serotype	Source	BLAST Alignment of *lmo2121* PCR Product	Lineage
2	4b	F	GT65	II
3	4b	C	GT65	II
22	4b	F	GT65	II
850	4b	C	GT65	II
774	4b	F	GT65	II
1174	4b	F	GT65	II
1945	4b	F	GT65	II
2142	4b	F	GT65	II
2905	4b	F	GT65	II
426	4b	E	GT65	II
842	4b	C	GT65	II
1015	4b	F	GT65	II
1026	4b	C	GT65	II
3724	4b	E	GT65	II
3763	4b	E	GT65	II
1380	1/2b	F	GT65	II
2129	1/2b	F	GT65	II
425	1/2b	E	GT65	II
394	1/2b	F	GT65	II
3830	1/2b	E	GT65	II
1031	3b	F	GT65	II
19	4d	A	GT65	II
1088	4d	C	GT65	II
2182	4e	C	GT65	II
11	1/2c	A	*lmo2121*	I
12	1/2c	C	*lmo2121*	I
3768	1/2c, 3c[1]	E	*lmo2121*	I
3812	1/2c	E	*lmo2121*	I
3826	1/2c	E	*lmo2121*	I
3883	1/2c	E	*lmo2121*	I
2411	1/2a	C	*lmo2121*	I
427	1/2a	E	*lmo2121*	I
337	1/2a	C	*lmo2121*	I
977	1/2a	F	*lmo2121*	I
1012	1/2a	A	*lmo2121*	I
1045	1/2a	F	*lmo2121*	I
2443	1/2a	A	*lmo2121*	I
3726	1/2a	E	*lmo2121*	I
3731	1/2a, 3a[1]	E	*lmo2121*	I
3744	1/2a	E	*lmo2121*	I
3838	1/2a	E	*lmo2121*	I
3893	1/2a	E	*lmo2121*	I
3943	1/2a, 3a[1]	E	*lmo2121*	I
3949	1/2a, 3a[1]	E	*lmo2121*	I
2863	3a	C	*lmo2121*	I

[1] Serotype based on multiplex PCR as described by Doumith *et al.* (23).

Screening of *L. monocytogenes* Isolates for Maltose Fermentation

The 45 isolates listed in Table 5 and the GT65 deletion mutant were screened for maltose fermentation. The screening of the isolates on the maltose purple agar indicated that there was no significant difference in maltose usage between the isolates or the GT65 deletion mutant (Figure 13).

Comparison of Genome Sequences Surrounding the GT65 and *lmo2121* Genes

The genome sequences surrounding the GT65 and *lmo2121* genes were compared using three of the available genome sequences in the J. Craig Venter Institute (JCVI) database. The *L. monocytogenes* isolates used were EGD (serotype 1/2a), F2365 (serotype 4b) and F6854 (serotype 1/2a). Figure 14 shows a schematic of the genome organizations surrounding the GT65 and the *lmo2121* sequence. Interestingly the F6854 isolate contained a GT65 gene, unlike the data obtained in our screening of 45 isolates, which showed that none of the serotype 1/2a, 1/2c and 3a isolates contained this sequence. At the 5' end of the GT65 gene sequence in isolate F2365 (serotype 4b), there are sequences encoding a hypothetical protein and a maltodextrin ABC transporter. At the 5' end of the GT65 sequence in the F6854 (serotype 1/2a) strain there are sequences encoding a maltose utilization protein (MalA) and a maltodextrin ABC transporter protein. At the 5' end of the *lmo2121* sequence in the EGD (serotype 1/2a) isolate is a gene encoding a protein similar to MalA and a gene encoding a protein similar to the maltodextrin ABC transporter. Adjacent to the 3' end of the genes of interest in both EGD and F6854 are sequences encoding hypothetical and putative proteins. Adjacent to the

3' end of the glycosyltransferase gene in isolate F2365 is a gene involved in teichoic acid biosynthesis and a gene encoding a hypothetical protein.

Figure 13. Maltose Fermentation in Selected *L. monocytogenes* Isolates
(A) Agar color change from purple to yellow due to maltose fermentation is shown for 13 *L. monocytogenes* isolates. The negative control remains purple. (B) Comparison of maltose fermentation between HPB#3 and the GT65 deletion mutant.

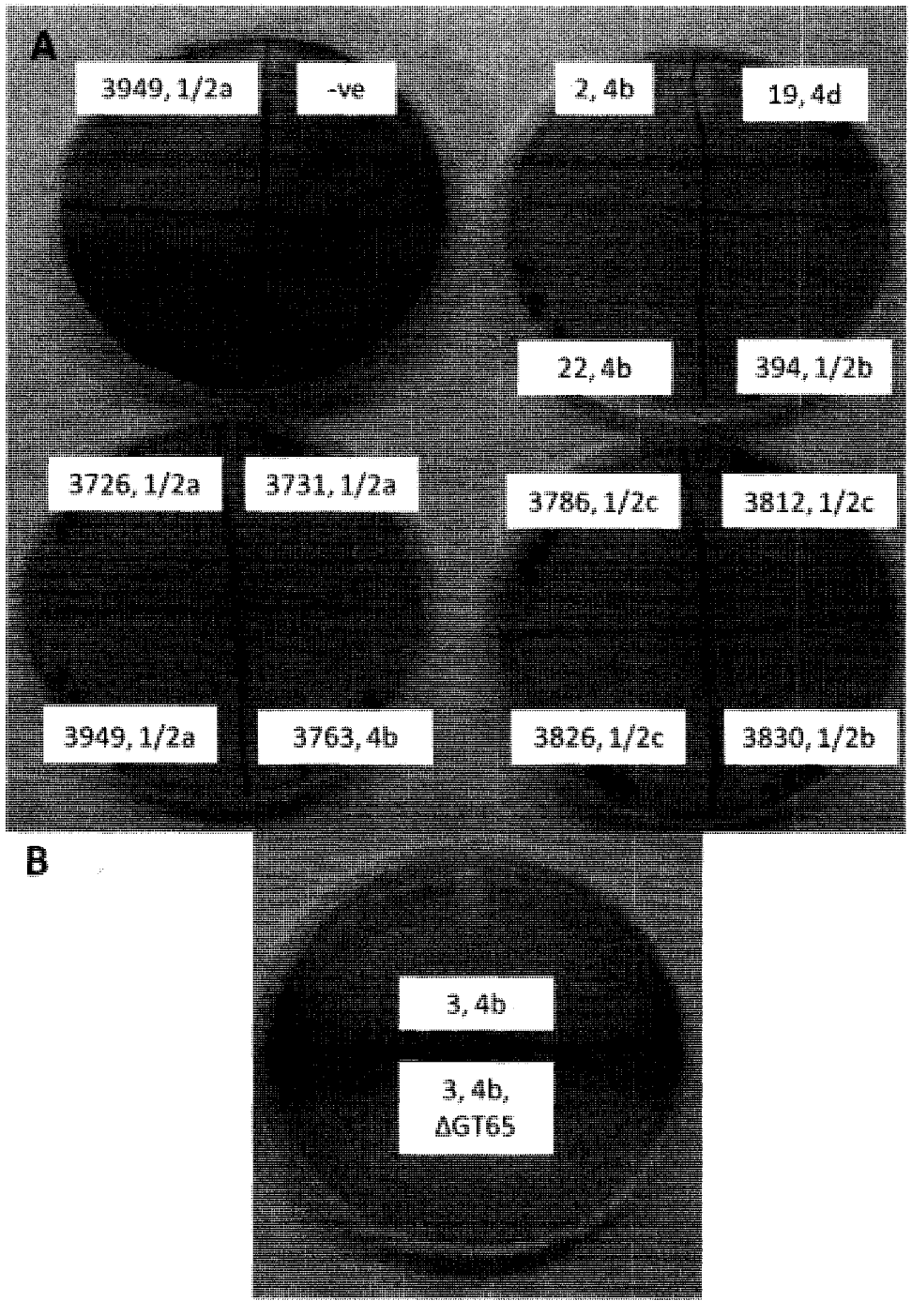
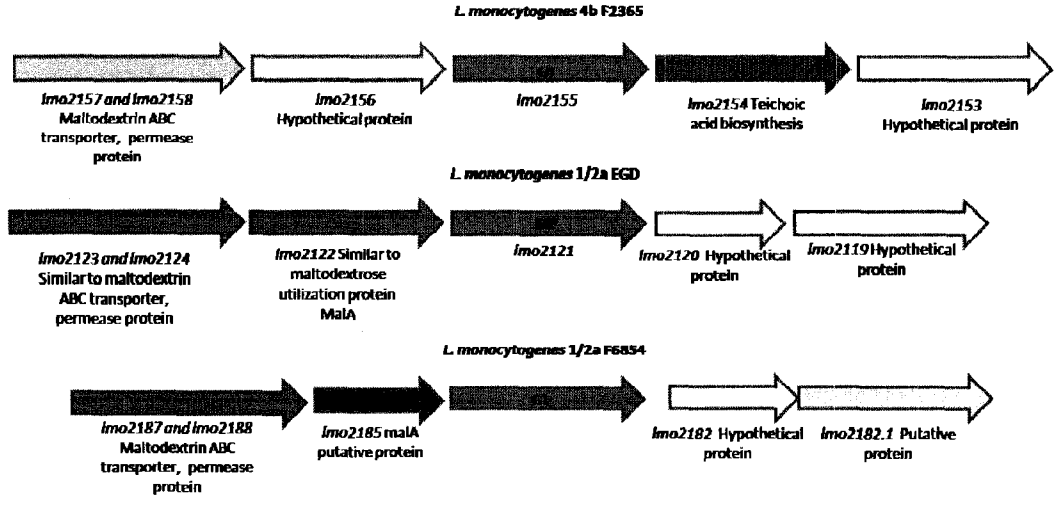

A
3949, 1/2a
-ve
2, 4b
19, 4d
22, 4b
394, 1/2b
3726, 1/2a
3731, 1/2a
3786, 1/2c
3812, 1/2c
3949, 1/2a
3763, 4b
3826, 1/2c
3830, 1/2b
B
3, 4b
3, 4b, ΔGT65

Figure 14. Schematic of Genomic Sequences Surrounding the GT65 and *lmo2121* (Similar to Maltose Phosphorylase) Genes in Three Different Strains of *L. monocytogenes*

Figure adapted from data obtained using the Genome Region Comparison function in the J. Craig Venter Institute Web Database (formerly known as The Institute for Genomics Research)



GT= glycosyltransferase family 65
MP= similar to maltose phosphorylase

Susceptibility of Selected *L. monocytogenes* Isolates and the GT65 Deletion Mutant to Ampicillin and Vancomycin

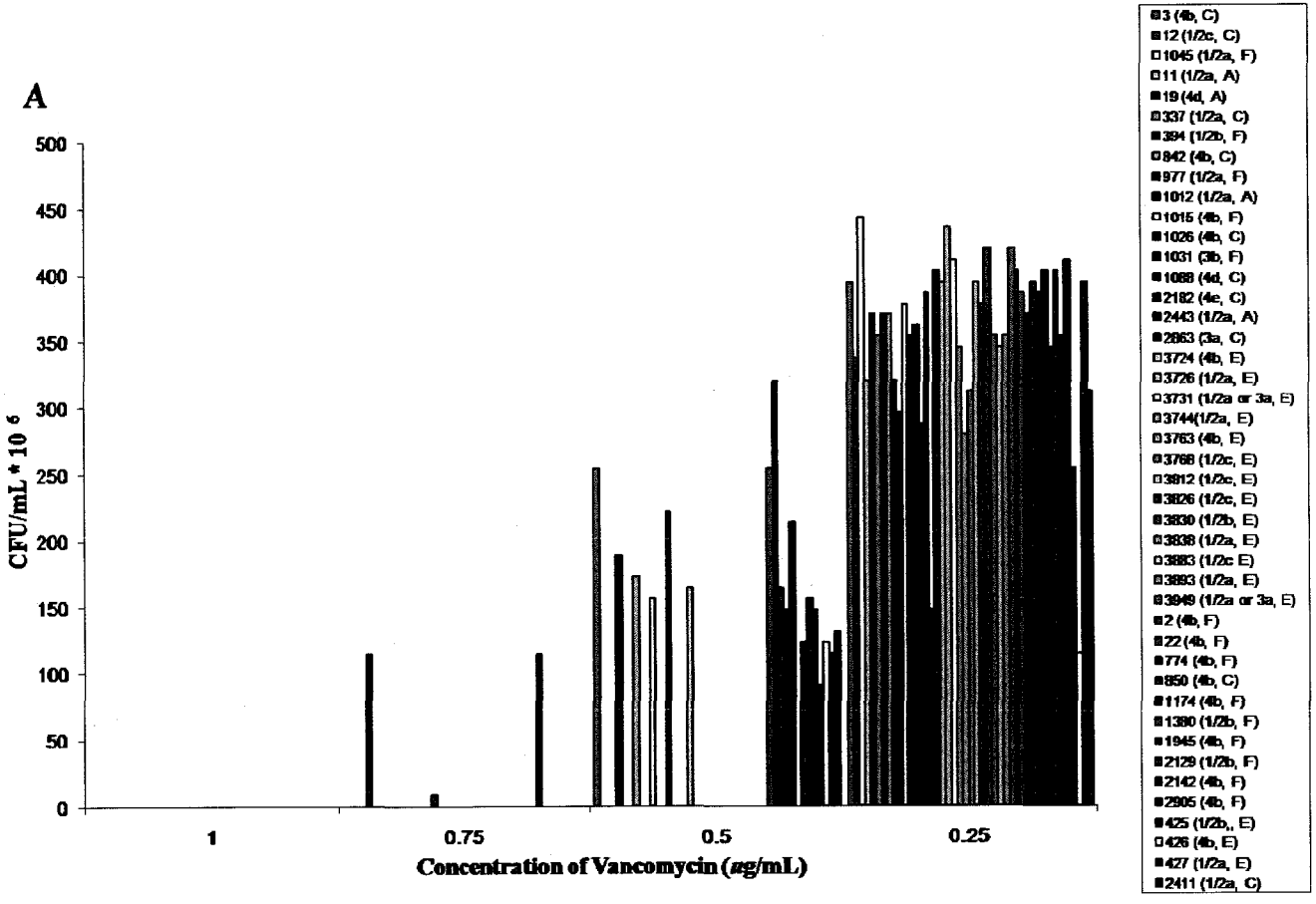
The 45 isolates of *L. monocytogenes* listed in Table 5 were analyzed for their susceptibility to the cell wall biosynthesis inhibiting antibiotics ampicillin and vancomycin. Figures 15A and16A illustrate the average growth of each of the 45 isolates in vancomycin and ampicillin antibiotic medium. To obtain information about the effect of serotype on the growth of the microorganism in the presence of the antibiotics, these data were grouped according to the serotype of the isolate. The data from serotypes 1/2a, 1/2c and 3a was grouped, and all other serotypes were grouped as separate data. As seen in Figure 16B, in the ampicillin medium, there was a significant difference in the growth of the serotype 1/2a, 1/2c and 3a isolates as compared to the other serotypes. When the concentration of ampicillin was 0.25 μg/mL, none of the serotype 1/2a, 1/2c or 3a isolates were able to grow, but some isolates, the majority of which were serotype 4b, were able to grow in the presence of this concentration of ampicillin. At lower concentrations of ampicillin, serotype 1/2a, 1/2c and 3a isolates were able to grow, but the concentration of cells measured was consistently lower than the growth observed in other serotypes.

In the vancomycin medium (Figure 15 A and B), none of the isolates showed growth at the highest concentration of 1μg/mL. At concentrations of 0.75 μg/mL and 0.50 μg/mL, serotype 1/2a, 1/2c and 3a isolates did not grow to the same cell concentration as isolates of other serotypes. However, at a vancomycin concentration of 0.25 μg/mL, there was no significant difference in growth between isolates belonging to the 1/2a, 1/2c and 3a serotype group as compared to isolates of other serotypes.

The GT65 deletion mutant showed a significantly reduced ability to grow in vancomycin at a concentration of 0.5 μg/mL (Figure 17A). The ability of the mutant to grow at the lowest concentration of vancomycin (0.25 μg/mL) was not significantly different than its isogenic wild-type, HPB#3. In the ampicillin susceptibility tests, the GT65 deletion mutant showed a significantly reduced ability to grow at ampicillin concentrations of 0.25, 0.1 and 0.085 μg/mL, when compared to the wild-type strain (Figure 17B). At the lowest concentration of ampicillin (0.055 μg/mL), there was no significant difference in growth between the deletion mutant and the wild-type strain.

Figure 15. (A) Average Growth of 45 *L. monocytogenes* Isolates and (B) Average Growth of Serotypes 1/2a, 1/2c and 3a as Compared to Other Serotypes* in Different Concentrations of Vancomycin

Averages were calculated from three replicates of the experiment.
* Other serotypes include 1/2b, 3b, 4b, 4d and 4e.
Significant data is indicated by a * ($p<0.05$).
C=Clinical Isolate; F=Food Isolate; E=Environmental Isolate; A=Animal Isolate.


A
500
450
400
350
300
250
200
150
100
50
0
CFU/mL * 10 6
1
0.75
0.5
0.25
Concentration of Vancomycin (µg/mL)
3 (4b, C)
12 (1/2c, C)
1045 (1/2a, F)
11 (1/2a, A)
19 (4d, A)
337 (1/2a, C)
394 (1/2b, F)
842 (4b, C)
977 (1/2a, F)
1012 (1/2a, A)
1015 (4b, F)
1026 (4b, C)
1031 (3b, F)
1088 (4d, C)
2182 (4e, C)
2443 (1/2a, A)
2863 (3a, C)
3724 (4b, E)
3726 (1/2a, E)
3731 (1/2a or 3a, E)
3744(1/2a, E)
3763 (4b, E)
3768 (1/2c, E)
3812 (1/2c, E)
3826 (1/2c, E)
3830 (1/2b, E)
3838 (1/2a, E)
3883 (1/2c E)
3893 (1/2a, E)
3949 (1/2a or 3a, E)
2 (4b, F)
22 (4b, F)
774 (4b, F)
850 (4b, C)
1174 (4b, F)
1380 (1/2b, F)
1945 (4b, F)
2129 (1/2b, F)
2142 (4b, F)
2905 (4b, F)
425 (1/2b,, E)
426 (4b, E)
427 (1/2a, E)
2411 (1/2a, C)

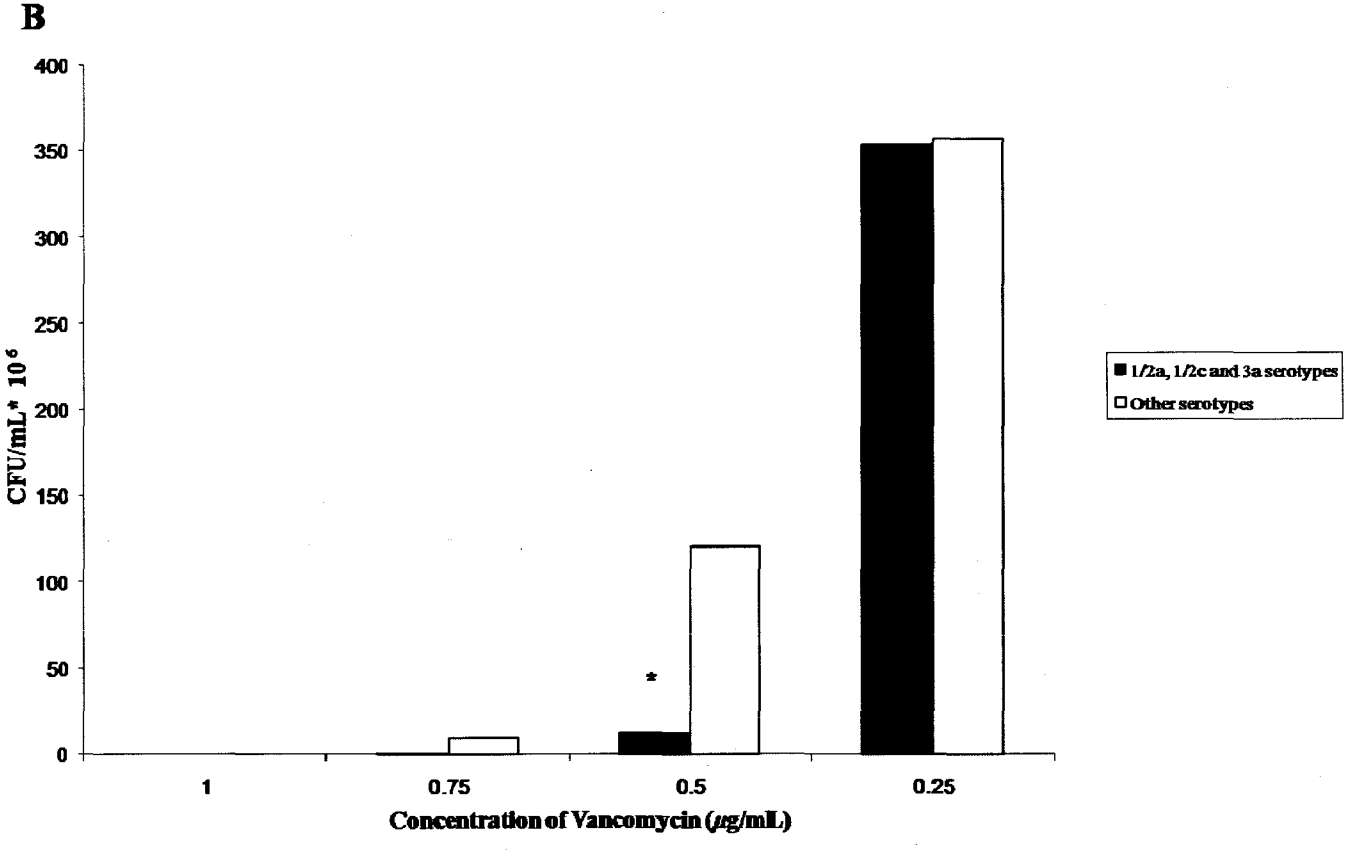
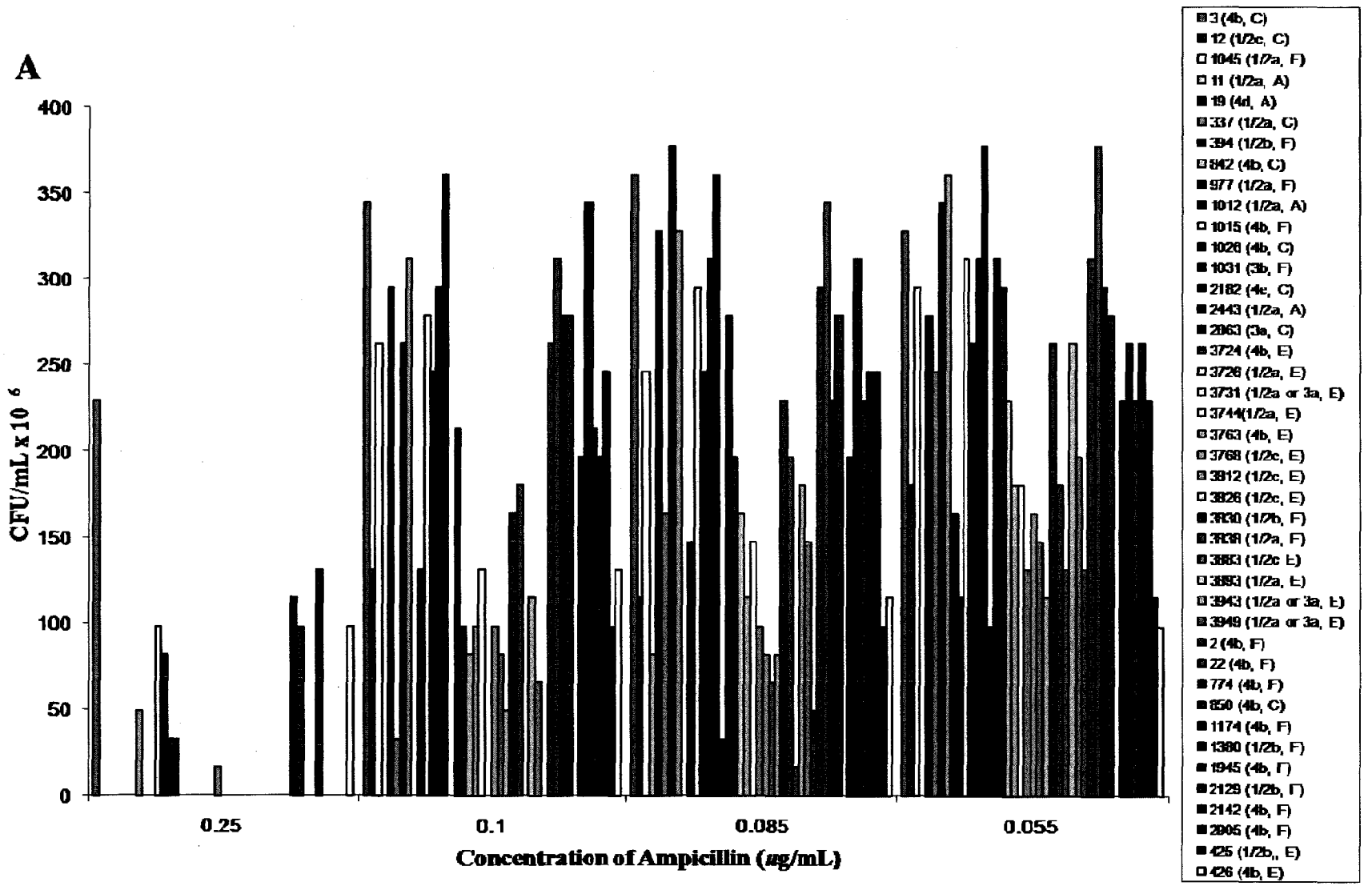

B
400
350
300
250
200
150
100
50
0
CFU/mL* 10⁶
*
1
0.75
0.5
0.25
Concentration of Vancomycin (μg/mL)
1/2a, 1/2c and 3a serotypes
Other serotypes

Figure 16. (A) Average Growth of 45 *L. monocytogenes* Isolates and (B) Average Growth of Serotypes 1/2a, 1/2c and 3a as Compared to Other Serotypes* in Different Concentrations of Ampicillin

Averages were calculated from three replicates of the experiment.
* Other serotypes included 1/2b, 3b, 4b, 4d and 4e.
Significant data is indicated by a * ($p<0.05$).
C=Clinical Isolate; F=Food Isolate; E=Environmental Isolate; A=Animal Isolate.


A
CFU/mL x 10 6
400
350
300
250
200
150
100
50
0
0.25
0.1
0.085
0.055
Concentration of Ampicillin (μg/mL)
3 (4b, C)
12 (1/2c, C)
1045 (1/2a, F)
11 (1/2a, A)
19 (4d, A)
337 (1/2a, C)
394 (1/2b, F)
842 (4b, C)
977 (1/2a, F)
1012 (1/2a, A)
1015 (4b, F)
1026 (4b, C)
1031 (3b, F)
2182 (4c, C)
2443 (1/2a, A)
2963 (3a, C)
3724 (4b, E)
3726 (1/2a, E)
3731 (1/2a or 3a, E)
3744(1/2a, E)
3763 (4b, E)
3768 (1/2c, E)
3812 (1/2c, E)
3826 (1/2c, E)
3830 (1/2b, F)
3839 (1/2a, F)
3853 (1/2c E)
3883 (1/2a, E)
3943 (1/2a or 3a, E)
3949 (1/2a or 3a, E)
2 (4b, F)
22 (4b, F)
774 (4b, F)
850 (4b, C)
1174 (4b, F)
1380 (1/2b, F)
1945 (4b, F)
2129 (1/2b, F)
2142 (4b, F)
2905 (4b, F)
425 (1/2b,, E)
426 (4b, E)

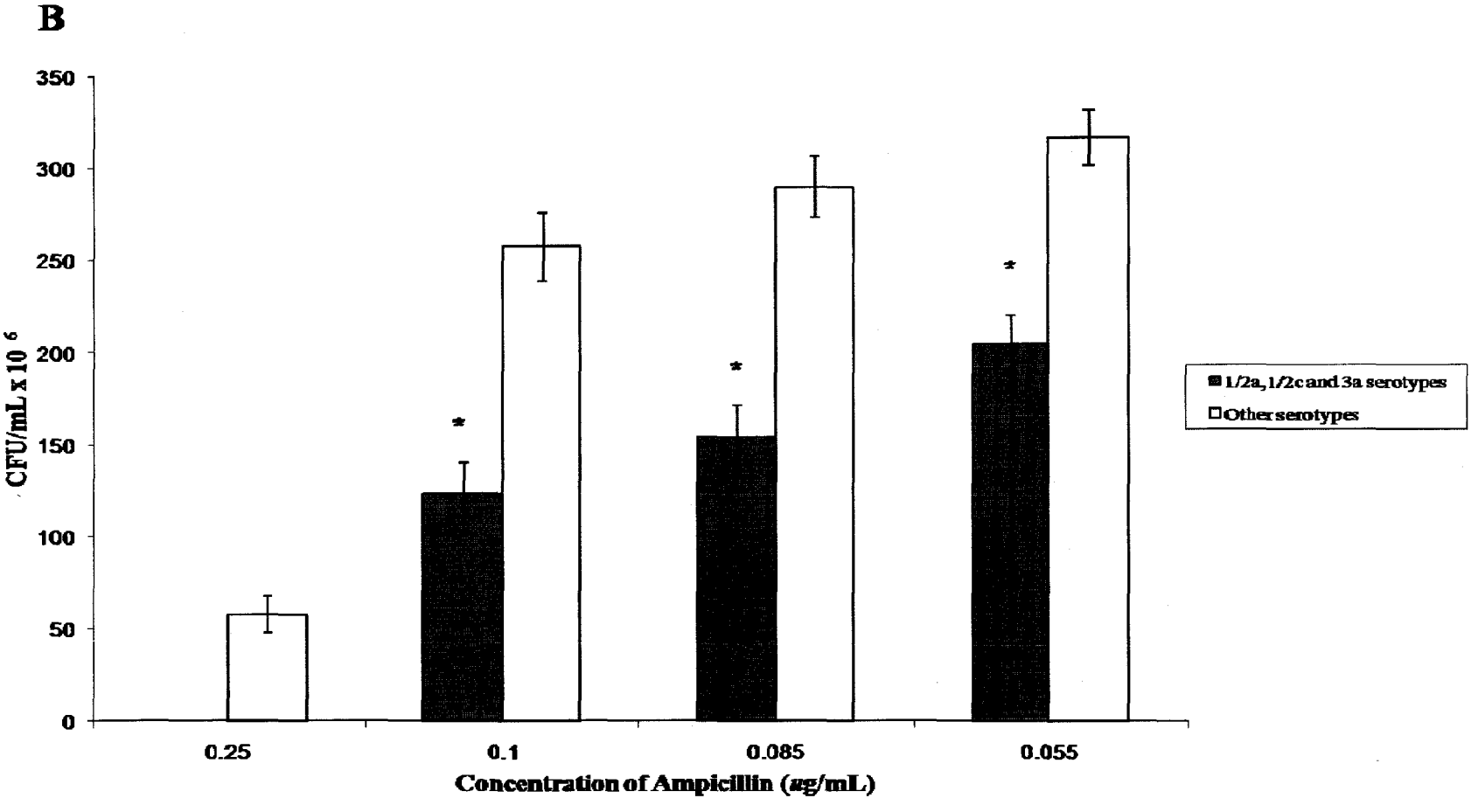
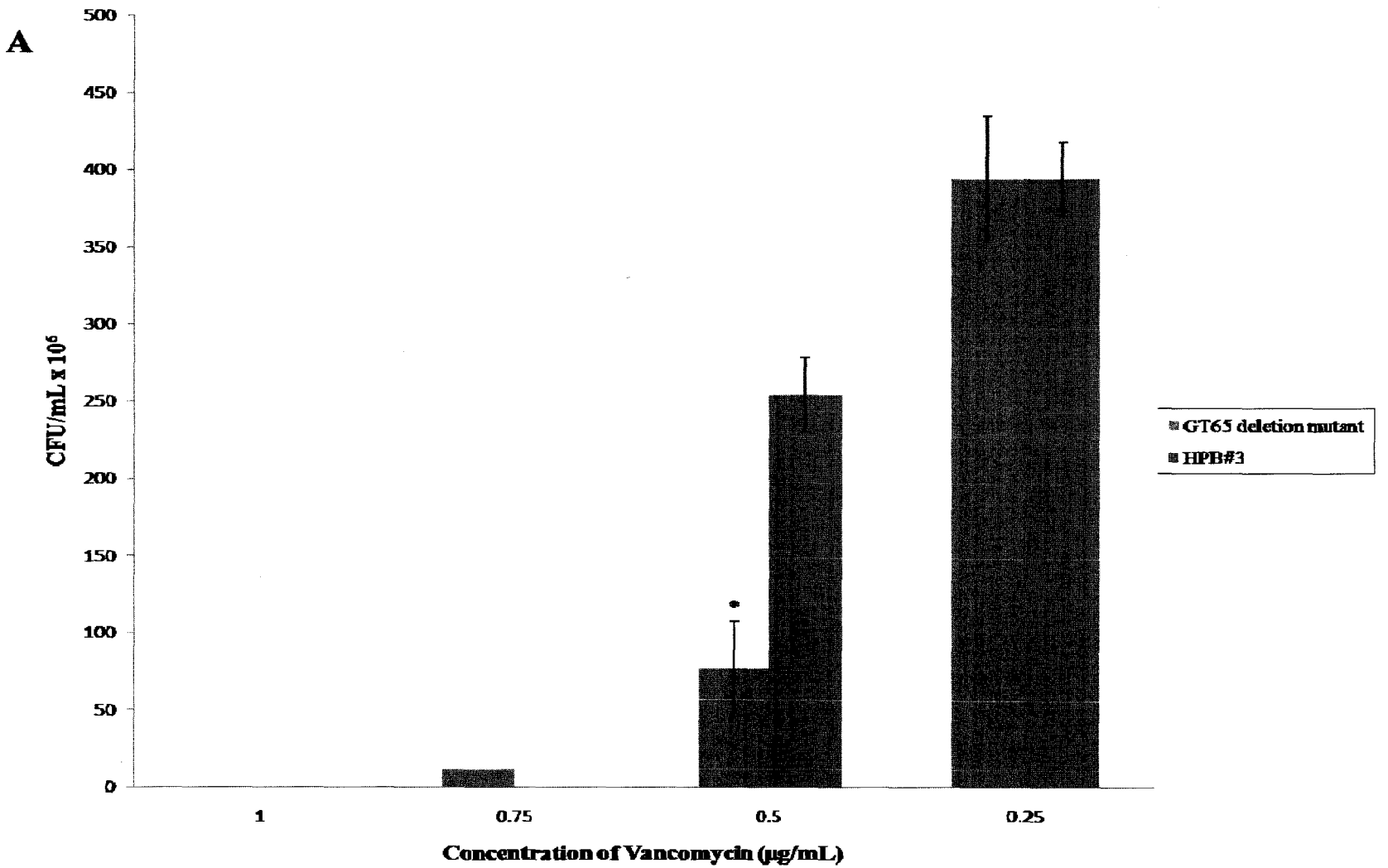
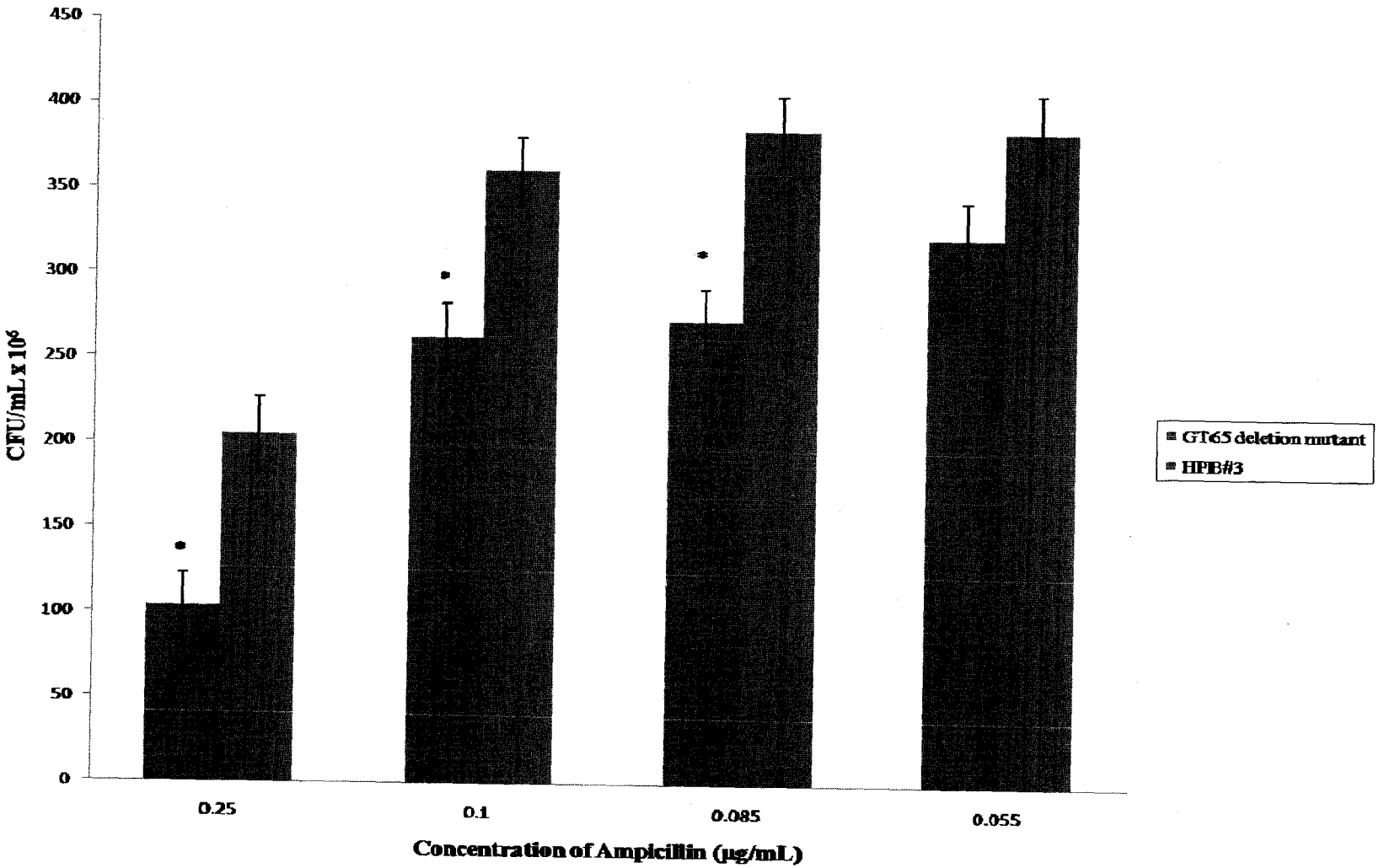

B
350
300
250
200
150
100
50
0
CFU/mL x 10 6
*
*
*
0.25
0.1
0.085
0.055
Concentration of Ampicillin (µg/mL)
1/2a, 1/2c and 3a serotypes
Other serotypes

Figure 17. Average Growth Comparisons of the GT65 Deletion Mutant to the Isogenic Wild-type HPB#3 in Different Concentrations of Vancomycin (A) and Ampicillin (B).
Averages were calculated from three replicate experiments.
Significant differences are indicated by an * ($p<0.05$).


A
500
450
400
350
300
250
200
150
100
50
0
CFU/mL x 10^6
1
0.75
0.5
0.25
Concentration of Vancomycin (µg/mL)
GT65 deletion mutant
HPB#3
*

B


450
400
350
300
250
200
150
100
50
0
CFU/mL x 10^6
0.25
0.1
0.085
0.055
Concentration of Ampicillin (µg/mL)
GT65 deletion mutant
HPB#3

DISCUSSION

Nucleotide and Amino Acid Sequence Comparisons of lmo2121 Similar to Maltose Phosphorylase and lmof2365_2155 Glycosyltransferase family 65

At the nucleotide level, the sequences of *lmo2121* and *lmof2365_2155* showed a 92% identity and, at the amino acid level, the sequences showed 93% homology. At the C terminus of the amino acid sequence (i.e., 5' end of the gene sequence), there were a large number of amino acid differences between the two sequences. Some of these differences included a change in one amino acid for another with similar charge and property (i.e., at position 726 the GT65 sequence encodes for a valine, whereas the maltose phosphorylase encodes for an isoleucine). Therefore, this non-synonymous mutation may be considered a neutral mutation, where the mutation to a chemically similar amino acid may not result in an altered function or property of the protein. However, other amino acid substitutions may result in the replacement of one amino acid for another with a different charge or property. For example, at position 750, there is an arginine in the GT65 sequence but in the maltose phosphorylase sequence there is an isoleucine in the same position. Arginine is a polar, hydrophilic amino acid whereas isoleucine is non-polar and hydrophobic. These examples indicate that amino acid differences between Lmo2121 and Lmof2365_2155 could cause differences in the physical and chemical properties (and perhaps the substrate specificity and function) of the protein which they encode, if these substitutions are occurring within the catalytic and binding domains of the enzyme.

Comparisons of Genome Sequences Surrounding the Glycosyltransferase family 65 Genes and *lmo2121* Similar to Maltose Phosphorylase

The genome regions surrounding the GT65 and *lmo2121* genes were compared in three different isolates of *L. monocytogenes* (F2365, serovar 4b; EGD, serovar 1/2a; and F6854, serovar 1/2a). At the 5' end of the genomes, there was an operon encoding a maltodextrin ABC transporter. In the EGD genome sequence, the operon was annotated as "similar to a maltodextrin ABC transporter", indicating that there was most likely differences in the nucleotide sequence from the other two isolates. In isolates EGD and F6854, both serotype 1/2a, there is a gene encoding a protein similar to a maltodextrin utilization protein and a maltodextrin utilization protein, respectively, directly upstream of the *lmo2121* and the GT65 genes. These genes are all related to the metabolism of maltose and may be regulated together as one operon during starch and sugar metabolism. On the other hand, the gene directly adjacent to the 5' end of the GT65 gene in isolate F2365 (serotype 4b) is annotated as a hypothetical protein, of which the function is yet to be defined. Directly downstream of the GT65 gene in strain F2365 is a gene for teichoic acid biosynthesis, whereas in the EGD and F6854 isolates, the genes in the same positions are annotated as hypothetical proteins with an unknown function. This comparison of the genome regions suggest that the *lmo2121* and the glycosyltransferase genes in EGD and F6854, respectively, may be involved in sugar, particularly maltose metabolism. On the other hand, the genes surrounding the GT65 gene in the F2365 isolate do not suggest such an obvious role in maltose metabolism, but rather a role in generating components of the cell wall. Teichoic acid is a component of the cell wall of *L. monocytogenes* and many other Gram-positive bacteria and is an essential component of microorganism viability. In

L. monocytogenes, the cell wall contains a polyribitol phosphate type of teichoic acid, with serotype diversity arising from glycosidic substitutions of the ribitol phosphate units. It is also believed that different patterns of teichoic acid glycosylation may be involved in the pathogenesis and adaptive physiology of the microorganism (62). The potential role(s) of GT65 in cell wall biosynthesis will be discussed later in the Chapter.

Strain F2365 (serotype 4b) belongs to the lineage II family of *L. monocytogenes*, while the EGD and F6854 isolates (serotype 1/2a) both belong to lineage I. In our study, as described by Hain *et al* (35), we refer to serotypes 1/2a, 1/2c, 3a and 3c as belonging to lineage I and serotypes 4b, 4d, 4e, 1/2b, 3b and 7 as belonging to lineage II. Lineage I isolates are isolated more frequently from foods, while lineage II isolates are more often implicated in human cases of listeriosis. Molecular phylogeny and evolutionary data suggests that the two lineages differ in their evolutionary history, with horizontal gene transfer and positive selection contributing to the evolution of *L. monocytogenes*. It has also been observed that genetic exchange between the two lineages is a rare occurrence (55). While some strain specific and serotype-specific genes between serotype 4b and 1/2a isolates were found in a study by Nelson *et al.* (54), the *L. monocytogenes* genome was found to be stable, with most differences occurring from phage events, transposable elements, strain or serotype unique genes scattered throughout the genome, or SNPs in known virulence genes and islands of unknown hypothetical proteins. It is believed that *L. monocytogenes* and *L. innocua* both arose from a common ancestor and *L. monocytogenes* later acquired virulence genes to become a pathogenic microorganism. *L. ivanovii*, the animal pathogen, is not as genetically closely related to *L. monocytogenes* as *L. innocua*, and is believed to have acquired its virulence traits in a separate, genomic-based event (70).

As the genome of *L. monocytogenes* evolved over time into separate lineages, it is possible the *lmo2121* sequence involved in maltose metabolism in lineage I isolates, which predominate in food isolates, has evolved to become a functionally different enzyme potentially involved in cell wall biosynthesis in the lineage II isolates, which are more prevalent in clinical cases. It is unlikely that *lmo2121* and the *lmof2365_2155* sequences were inserted into the genome in separate phage or gene acquisition events, as the sequences surrounding them do share some similarity and the genes themselves are both exactly 2262 bp in length. This suggests an evolution of a new function for the GT65 gene via SNPs and positive selection, potentially resulting in a GT65 gene that is providing lineage II isolates with a competitive advantage in the host environment.

PCR Screening for *lmo2121* Reveals its Presence in 1/2a, 1/2c and 3a Serotypes and Absence in Other Serotypes of *L. monocytogenes*

In all 45 of the isolates screened by this method, those of serotype 1/2a, 1/2c and 3a contained the *lmo2121* sequence, while the sequence analysis of all the other serotypes revealed that their PCR product showed a higher (99-100%) identity to the GT65 sequence than the *lmo2121* sequence, to which they showed only 89% identity. This is a strong indication that the *lmo2121* sequence is only present in *L. monocytogenes* isolates of serotype 1/2a, 1/2c and 3a. While the GT65 sequence in other serotypes is very similar (i.e. those in lineage II), there are enough significant differences at the nucleotide and amino acid levels from the *lmo2121* sequence, to indicate that the proteins encoded by these genes are performing different functions and could potentially play a role in the environmental adaptation or tropism to a host environment of the microorganism.

Maltose Fermentation is Similar for all Isolates of *L. monocytogenes* Tested, Including the Glycosyltransferase Family 65 Deletion Mutant

To investigate if there were differences in maltose fermentation in *L. monocytogenes* isolates possessing the *lmo2121* sequence versus those with the GT65 sequence, isolates were streaked onto maltose agar and the fermentation of the sugar (or lack thereof) was observed. After 24 and 48 h, based on visual observation, maltose fermentation appeared to be occurring at an equal rate for all of the isolates. The GT65 deletion mutant also showed the same rate of maltose usage as its wild type counter-part, HPB#3 (serotype 4b). In the KEGG database, the description of the GT65 is as a maltose phosphorylase involved in the starch and sucrose metabolism pathway (40). The maltose phosphorylase enzyme converts maltose, in the presence of phosphate, into D-glucose and β-D-glucose-1-phosphate. The *L. monocytogenes* genes *lmo2121* and *lmof2365_2155* are both listed as members of the maltose phosphorylase enzyme group (E.C. 2.4.1.8). However, the amino acid and nucleotide sequences are not completely identical, and several of the amino acid substitutions could potentially play a role in altering the function of the protein. Further, the GT65 deletion mutant was able to utilize maltose as efficiently as the wild-type isolate. This indicates that the GT65 enzyme encoded by the gene *lmof2365_2155* is not involved with maltose utilization and may be involved in other cellular processes. Alternatively, *L. monocytogenes* may possess a redundant gene for maltose metabolism that was activated in the deletion mutant in order for the organism to replicate and survive. Given that the GT65 deletion mutant was able to successfully

metabolize maltose, further investigations into the role of this enzyme would be necessary to reveal its true function in the ecology of the microorganism.

Deletion Mutants of Glycosyltransferase Family 65 Show a Reduced Ability to Multiply in the Presence of Ampicillin and Vancomycin

The effects of the cell wall biosynthesis inhibiting antibiotics ampicillin and vancomycin were observed for a set of 45 *L. monocytogenes* strains (Table 5), including the GT65 deletion mutant. In the presence of ampicillin, serotypes 1/2a, 1/2c and 3a and the GT65 deletion mutant, which do not contain the GT65 genome sequence, showed a significantly reduced ability to survive as compared to all other serotypes (Figures 16B and 17B). This was seen throughout a range of ampicillin concentrations, from 0.085 μg/mL up to 0.25 μg/mL. Similar results were obtained in the vancomycin assay, except that a significant difference in growth was only observed in the growth of the 1/2a, 1/2c and 3a serotypes and the GT65 deletion mutant at a concentration of 0.5 μg/mL (Figures 15A and 17A). The GT65 deletion mutant and *L. monocytogenes* serovars 1/2a, 1/2c and 3a all showed a significantly reduced ability to thrive in ampicillin and vancomycin media.

Ampicillin and vancomycin are both inhibitors of cell wall biosynthesis. The inability of the GT65 deletion mutant and isolates of serotype 1/2a, 1/2c and 3a to thrive in these media indicates that the GT65 enzyme could potentially be involved in cell wall metabolism and biosynthesis, or even function as as a penicillin-binding protein.

Acriflavine is an ingredient in UVM broth which is used in the pre-enrichment step for the isolation of *Listeria* from food and environmental sources (58). The role of acriflavine is to inhibit the growth Gram-positive microorganisms. While *Listeria* is Gram-

positive, the presence of esculin in the medium promotes its growth, while acriflavine inhibits the growth of the other Gram-positive bacteria that may be present in the sample. The same 45 isolates used in the vancomycin and ampicillin assays, as well as the GT65 deletion mutant, were grown in Mueller-Hinton broth with various concentrations of acriflavine. Interestingly, isolates of serotype 1/2b and 4b, as well as the GT65 mutant, were not able to grow as quickly as compared to serotypes 1/2a, 1/2c and 3a (Appendix A). This is an important observation, as it may account for the fact that serotypes 1/2a, 1/2c and 3a are more commonly isolated from food. If the serotype 4b and 1/2b isolates do not grow as rapidly in the pre-enrichment step and the sample contains multiple serotypes, the colonies of these serotypes would be much smaller, or not visible after 48 h of growth, and the larger colonies of serotypes 1/2a, 1/2c or 3a would be selected for further characterization. Furthermore, *L. innocua* is known to grow more rapidly than *L. monocytogenes* and thus serotype 4b and 1/2b strains and potentially others possessing the GT65 sequence and protein, may be excluded or outcompeted from standard isolation procedures in samples containing both *L. innocua* and *L. monocytogenes* (18). Further investigation of this phenomenon may be warranted to prevent any incorrect assumptions regarding the serotype prevalence of *L. monocytogenes* found in food and environmental sources.

The Glycosyltransferase family 65 Gene is a Potentially Important Element for the Tropism of *L. monocytogenes*

Glycosyltransferases are a group of enzymes which catalyze the transfer of a sugar molecule from a sugar donor molecule onto a saccharide or a non-saccharide acceptor.

They are involved in a variety of cellular functions including sugar and cell-wall metabolism, host-pathogen interactions and protein glycosylation (19). The great diversity of the glycosyltransferases makes it extremely challenging in the post-genomic era to elucidate the function of the many ORFs with similarities to the glycosyltransferase family. Furthermore, glycosyltransferases are known to be extremely difficult to classify biochemically in the laboratory and therefore, bioinformatic annotations for thousands of putative glycosyltransferases are assigned based on only a handful of enzymes whose specific biological activity and 3D structure is known (19). Glycosyltransferases, as well as most other enzymes, are classified based on their donor, acceptor and product specificity as recommended by the International Union of Biochemistry and Molecular Biology (IUBMB) (52). However, this method requires complete characterisation of the donor, acceptor and product of the enzyme, which cannot be truly elucidated unless the biochemistry of the purified enzyme has been studied. Additionally, many glycosyltransferases act on several different substrates, and the IUBMB Enzyme Commission (EC) numbers are unable to accommodate such diversity in one enzyme, and do not reflect the intrinsic enzyme structure or mechanisms (19).

Under the IUBMB EC system, EC 2 are the tranferases, EC 2.4 are the glycosyltransferases and EC 2.4.1 are the hexosyltransferases. Under this system, GT65 is classified as an EC 2.4.1.8 maltose phosphorylase (40). In the CAZY database (available at www.cazy.org), all the glycosyltransferases are categorized by amino acid sequence similarities. However, assigning a function to putative GT sequences based on sequence homology remains problematic, since there are many instances of closely related sequences with different catalytic activity (9). The CAZY database has assigned the GT65 family as

an EC 2.4.1.- enzyme, with a known activity of a GDP-Fuc: protein O-α-fucosyltransferase (19, 20). However, the assignment of this function to the enzyme is based only on data obtained in eukaryotes, and there is no information regarding its role in a prokaryotic system. In eukaryotes, fucosylation is an integral part of many biological and pathological processes, including cell adhesion, inflammation, tumor metastasis and fertilization (47). On the other hand, fucosylation in prokaryotes seems to be a less common occurrence, but has been implicated in a wide range of biological processes that include molecular mimicry, colonization, adhesion and modulation of the host immune response (47). Putative fucosyltransferases have been found in several bacteria including *E. coli*, *Salmonella enterica* serovar Typhi, *Y. pestis*, *Y. enterocolitica* and *H. pylori*. Of this group, *H. pylori* is the only microorganism whose fucosyltransferases have been functionally characterized. Fucosyltransferases (FucT) in *H. pylori* are involved in the final synthesis steps of Lewis blood structures on the LPS of the bacterium. While the function of these fucosylated LPS molecules is not completely understood, it has been proposed that they may be involved in molecular mimicry of the carbohydrates found in the stomach of the host and thereby aid the bacterium in establishing a long-term chronic infection (3, 47). Since the GT65 enzyme family has an assigned function of a FucT in the CAZY database, we compared the nucleotide and amino acid sequences of the *L. monocytogenes* GT65 sequence and the FucT sequences of *H. pylori*. The sequences showed no significant similarity at either the amino acid or nucleotide levels (data not shown). Given this result, it is likely that the GT65 sequence in *L. monocytogenes* is not acting as a fucosyltransferase. This result further emphasizes that the classification of the glycosyltransferase families is complicated, with much of the annotation of putative glycosyltransferases being inferred

from a handful of the enzymes whose biochemical function has been studied. While the sequence comparisons between the *L. monocytogenes* GT65 and *H. pylori* FucT show no significant similarity, further studies into the mechanism of GT65 are necessary to elucidate the function of this enzyme in *L. monocytogenes* and its potential role in the tropism of the microorganism.

The NCBI Entrez Gene function (www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene), was used to show that *lmof2365_2155* (GT65) had nucleotide sequence similarities to the conserved central catalytic domain, the N-terminal domain and the C-terminal domain of a glycosyl hydrolase family 65. Maltose phosphorylases and vacuolar acid trehalases are included in this family. Given that *L. monocytogenes* is an intracellular pathogen that enters the host cell through a phagosome and then proceeds to survive within and then degrade the phagosome, the vacuolar acid trehalase was of particular interest to this research. This particular enzyme appears to be well studied in the yeast *Saccharomyces cerevisiae*. In *S. cerevisiae*, the hydrolysis of trehalose is dependent on two enzymes: the cytoplasmic neutral trehalase (Nth1) and the vacuolar acid trehalase (Ath1). Limited work has been performed on Ath1, but it is known that the enzyme is most active in a low pH environment such as a vacuole, suggesting that this is the area of localization for the enzyme. However, Ath1 hydrolyses extracellular trehalose and it remains unclear how the enzyme accesses the trehalose and subsequently localizes to the vacuole (37). The most significant role of trehalose is to provide membrane protection for the microorganism under stressful conditions such as heat, drying and nutrient depletion (37). While there is little known about the role of trehalose and vacuolar trehalases in *L. monocytogenes*, the GT65 sequence shows similarity to the catalytic

domains of the vacuolar trehalase. If GT65 is acting as a vacuolar trehalase, it may be assisting *L. monocytogenes* in obtaining trehalose from outside the vacuole to use as an energy source within the vacuole of the host cell, thus aiding the microorganism in survival and subsequent proliferation within the host cell. While a more detailed study is required to test this hypothesis, it may provide insight into the tropism of certain *L. monocytogenes* serotypes (particularly 4b) for the human environment, i.e., intracellular survival. The absence of the GT65 sequence from 1/2a, 1/2c and 3a serotypes may put them at a slight disadvantage as compared to other serotypes, particularly 4b and 1/2b, in the phagosome of the host cell. *L. monocytogenes* isolates which do not possess the GT65 gene clearly still survive within the host vacuole, but perhaps those isolates which do possess the gene have an advantage for survival in this niche, as the majority of outbreaks are associated with serovar 4b (57).

The GT65 deletion mutant and isolates of serotypes 1/2a, 1/2c and 3a which do not have the GT65 sequence showed a reduced ability to survive in the presence of the cell wall biosynthesis inhibiting antibiotics ampicillin and vancomycin. This indicates that GT65 is potentially playing a role in cell wall biosynthesis and turnover. In *L. monocytogenes*, the penicillin-binding proteins (PBPs) are essential enzymes in bacterial cell wall peptidoglycan synthesis. They are multifaceted, having both glycosyltransferase and transpeptidase activity (81). The function of the glycosyltransferase portion of the penicillin-binding proteins is to catalyze the elongation of the glycan chain, while the transpeptidases catalyze the cross-linking between peptides of two adjacent glycan chains (81). In a study by Zawadzka-Skomial *et al.* (81), the *lmo2229* gene in *L. monocytogenes* EGD encoding PBP4 was deleted. This deletion did not have a significant effect on growth

rate, peptidoglycan composition or sensitivity to β-lactam antibiotics, but increased resistance to moenomycin was observed (81). Moenomycin is a natural antibiotic which inhibits cell wall biosynthesis via binding of the transglycosylases which are involved in formation of the peptidoglycan chain (74). When the PBP4 gene was deleted in *L. monocytogenes* EGD, the moenomycin did not have its usual substrate to act upon, and therefore the *lmo2229* deletion mutant was more resistant to the antibiotic. While GT65 has glycosyltransferase activity like the PBPs, it is not believed to possess the transpeptidase domain. This indicates that the GT65 is not likely acting as the most recently studied PBPs. Most PBPs possess a short cytoplasmic region at the N-terminus, followed by a transmembrane region. Class A PBPs possess a glycosyltransferase and transpeptidase domains, while Class B penicillin binding proteins have only the transpeptidase domain (49). While the GT65 enzyme does not appear to be acting as a PBP as its domains are significantly different from both the Class A and B PBPs, it is potentially involved in cell wall biosynthesis, as the deletion mutant became more susceptible to both ampicillin and vancomycin treatments as compared to the wild-type.

GT65 may be involved in generating new peptidoglycan and cell wall turnover during the synthesis of the *L. monocytogenes* cell wall. While not essential to cell wall synthesis since the deletion mutant was able to survive, it is possible that with the presence of this enzyme, *L. monocytogenes* is perhaps able to generate new cell wall components more efficiently and is able to counteract the inhibiting action of ampicillin and vancomycin, especially at the lower concentrations of these antibiotics. Ampicillin prevents peptidoglycan cross-linking, while vancomycin prevents elongation of the peptidoglycan chain (16). In the GT65 deletion mutant and the isolates without this sequence, it is

possible that cell wall turnover was slower, therefore rendering the cell wall more susceptible to the action of ampicillin and vancomycin. Peptidoglycan synthesis can be divided into three stages. The first stage is the synthesis of UDP-*N*-acetylmuramyl-pentapeptide (UDP-MurNAc-pentapide) and UDP-*N*-acetyl-glucosamine (UDP-GlcNAc) and takes place within the cytosol (69). In the second stage, precursor lipid intermediates are made. Lipid I is formed from the transfer of the phospho-MurNAc-pentapeptide portion of the UDP-MurNAc-pentapeptide to the membranes bound acceptor bactoprenol. Lipid II is created from the transfer of GlcNAc from UDP-GlcNAc to lipid I (69). The antibiotic vancomycin inhibits transglycosylase activity at this second stage, preventing elongation of the peptidoglycan pentapeptide and subsequent cross-linking (16). The third step of peptidoglycan synthesis, which occurs on the outside of the cytoplasmic membrane, is the incorporation of the newly-formed peptide units into the growing chain of peptidoglycan. This step uses the transglycosylation and transpeptidation functions of the PBPs (69). Ampicillin inhibits peptidoglycan synthesis by binding to the transpeptidase region of the PBP, preventing cross-linking of the peptidoglycan (16). In the GT65 deletion mutants and serotype 1/2a, 1/2c and 3a isolates, the cell wall turnover and synthesis may have been occurring more slowly, allowing greater opportunity for the antibiotics to bind to their substrates and prevent peptidoglycan synthesis.

Recently, it has been proposed that *L. monocytogenes* can evade host innate immune defences in the GI tract and macrophages via deacetylation of its peptidoglycan (8). The innate immune system recognizes microbes using pattern recognition molecules and since the peptidoglycan is a major component of the majority of bacteria, especially the Gram-positive bacteria, it acts as an excellent target for the pattern recognition molecules,

namely the peptidoglycan recognition proteins (PGRPs). Mammalian PGRPs are expressed in the skin, bone marrow, intestinal tract and liver (34). The PGRPs in mammals can be bactericidal, and function as microbial scavengers and neutrophil recruiters. Their bactericidal effect, especially towards Gram-positive bacteria, comes from their ability to bind to the bacterial peptidoglycan and presumably inhibit peptidoglycan synthesis, resulting in cell death (34). In *L. monocytogenes* the gene *pdgA* encodes for a peptidoglycan N-deacetylase, which partially N-deacetylates the *N*-acetylglucosamine residues of pepidoglycan into glucosamine. By altering the peptidoglycan, the PGRPs do not recognize the cell wall of *L. monocytogenes* and the microorganism is able to survive the gastrointestinal tract and uptake by macrophages, evade the activity of host lysozyme and down regulate the inflammatory response of the host (8). While GT65 does not show any nucleotide similarity to *pdgA*, it may be playing a role in cell wall synthesis and turnover. With greater cell wall turnover, N-deacetylation of the peptidoglycan may occur more readily, allowing for more successful evasion of the host immune system. The presence of GT65 in *L. monocytogenes* serotypes 4b, 1/2b and 3b may in part account for the success of these serotypes in causing human illness and their predominance in clinical cases versus serotypes such as 1/2a, 1/2c and 3a, which do not possess this gene.

The above proposed roles for GT65 in *L. monocytogenes* are based on biological data obtained in experiments described herein and information found in the literature. In order to completely understand the function of GT65 and its potential role in the tropism of the microorganism, further studies are warranted. It should be noted, however, while GT65 is annotated as a maltose phosphorylase, our experiments indicate that the enzyme may be involved in cell wall metabolism. The glycosyltransferases are extremely difficult to

classify biochemically and in the post-genomic era functions are being assigned to ORFs with similarity to different glycosyltransferase families based on sequence information alone. In addition, the CAZY database has classified lmof2365_2155 as a glycoside hydrolase. Further explanation can be found in Appendices B and C. Therefore, this body of work emphasizes the need to experimentally investigate the role of these glycosyltransferases, or any putative protein, before any conclusions can be made.

Chapter 4: Conclusions and Future Directions

In this project, the use of comparative genomics indicate that there are many differences at the genomic level that may have a role in the adaptation of different serovars of *L. monocytogenes* to host, food and/or environmental niches. Mixed-genome microarrays, combined with dot-blot hybridization and PCR were used to confirm these genomic differences. The second and third version of the array uncovered genome sequences that can be the starting point of future experiments to further aid in understanding the tropism of *L. monocytogenes*. Our mixed-genome arrays have not covered the entire genome of the microorganism and future arrays made from our mixed-genome library will certainly yield genomic differences that could be involved in the tropism of *L. monocytogenes* for host, food and environmental niches.

The second version of the mixed-genome array showed many differences in the genome sequences of the different serovars that were investigated. The array provided a starting point for future investigation into the role of many different genome sequences such as putative cell surface proteins, transcriptional regulators, retrons and genes involved in metabolic processes. While the array was useful in detecting the genome sequences that were present in some serovars and absent in others, it was apparent from our experimental data that other means of confirming the data obtained from the array was necessary. In some cases, the array would indicate that a genome sequence was absent from certain serovars, but upon further investigation via dot-blot hybridization and PCR, it was discovered that the genome sequence was present. Mixed-genome microarrays are a useful tool to uncover genomic differences among *L. monocytogenes*, but confirmation of the data obtained from the array using other methods is essential.

Suppressive subtractive hybridization was able to elucidate genomic differences between *L. monocytogenes* and *L. innocua*, but was unable to detect differences between different serovars of *L. monocytogenes*. In future, this method could be explored more thoroughly, as it has been used successfully to uncover genomic differences in other microorganisms. The time constraints in this study prevented further investigation into this particular methodology.

The use of comparative genomics uncovered a gene in certain serovars of *L. monocytogenes* that is possibly involved in the tropism of the microorganism and potentially provides it with an advantage for survival in the host environment. Glycosyltransferase family 65 (GT65) was present in serovars 1/2b, 4b, 4d, 4e and 4a, but absent from serovars 1/2a, 1/2c and 3a. Interestingly, serovars 1/2b and 4b are more prevalent in outbreaks of human listeriosis, while serovars 1/2a, 1/2c and 3a are more commonly isolated from foods. GT65 is classified as an EC 2.4.1.8 maltose phosphorylase, involved in the starch and sugar metabolism of *L. monocytogenes* (40). However, our experimental evidence pointed to a role for this enzyme in cell-wall metabolism and biosynthesis. There are a number of potential mechanisms by which this enzyme could be providing an advantage to *L. monocytogenes* in the host environment. For example, it is possible that GT65 is involved in processes to aid the microorganism in gut colonization, adhesion, molecular mimicry, survival in the phagolysosome, cell wall turnover or evasion of the host innate immune system.

To better understand the role of the GT65 gene in the tropism of *L. monocytogenes*, the GT65 deletion mutant could be used and compared to its wild-type strain in vacuolar survival studies, cell invasion assays, as well as animal-model infection and pathogenesis

studies. To better understand the substrate specificity of this enzyme, biochemical studies of the purified enzyme could be performed to elucidate the biochemical reaction(s) performed by the enzyme. Currently, the annotation of this enzyme is based only on the amino acid sequence, and as mentioned in the discussion (Chapter 3), there are many inherent difficulties in the classification of the glycosyltransferase families which can be overcome by doing more in-depth biochemical studies of the purified enzymes.

The growth of different serovars of *L. monocytogenes* in the presence of acrivflavine (an ingredient in UVM broth which is commonly used to isolate the microorganism from foods) indicated that after 24h of incubation, serovars 1/2b and 4b were not able to grow as quickly as serovars 1/2a, 1/2c and 3a. This could lead to a bias in the serotypes that are isolated from food sources, especially if the sample contains a combination of the serotypes and the 1/2b and 4b serotypes are outcompeted by the 1/2a, 1/2c and 3a serotypes. In the future, further studies into this problem may be warranted in order to prevent a serotype bias when isolating the microorganism from food and environmental samples.

In summary, comparative genomic studies of *L. monocytogenes* uncovered many genomic differences between strains of this microorganism. Genomic differences were uncovered between serovars of *L. monocytogenes* and between strains of the same serotype from different sources, i.e., clinical, food or environmental. One of the genomic sequences, GT65, was found in serovars which predominate in human illness and was absent from serovars most commonly isolated from foods. The deletion of this gene resulted in a reduced ability of *L. monocytogenes* to survive in the presence of the cell wall biosynthesis inhibiting antibiotics ampicillin and vancomycin. Our study discovered that the GT65 gene,

present in lineage II but absent from lineage I isolates, may be somehow involved in the tropism of certain isolates for the host environment. Further studies of this genomic marker could reveal its exact role in providing *L. monocytogenes* serovars 4b and 1/2b an enhanced ability to cause invasive listeriosis in humans. This could lead to a better understanding of why some strains of this microorganism are more prevalent in human illness.

REFERENCES

1. **Aarts, H. J., L. E. Hakemulder, and A. M. Van Hoef.** 1999. Genomic typing of *Listeria monocytogenes* strains by automated laser fluorescence analysis of amplified fragment length polymorphism fingerprint patterns. Int J Food Microbiol **49:**95-102.

2. **Akopyants, N. S., A. Fradkov, L. Diatchenko, J. E. Hill, P. D. Siebert, S. A. Lukyanov, E. D. Sverdlov, and D. E. Berg.** 1998. PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori.* Proc Natl Acad Sci U S A **95:**13108-13113.

3. **Appelmelk, B. J., R. Negrini, A.P. Moran and E.J. Kuipers.** 1997. Molecular mimicry between *Helicobacter pylori* and the host. Trends Microbiol **5:**70-73.

4. **Arvanitidou, M., A. Papa, T.C. Constantinidis *et al.*** 1997. The occurence of *Listeria* spp. and *Salmonella* spp. in surface waters. Microbiol Res. **152:**395-397.

5. **Bernagozzi, M., F. Bianucci, R. Sacchetti and P. Bisbini.** 1994. Study of the prevalence of *Listeria* spp. in surface water. Zentralbl Hyg Umweltmed **196:**237-244.

6. **Bibb, W. F., B. G. Gellin, R. Weaver, B. Schwartz, B. D. Plikaytis, M. W. Reeves, R. W. Pinner, and C. V. Broome.** 1990. Analysis of clinical and food-borne isolates of *Listeria monocytogenes* in the United States by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. Appl Environ Microbiol **56:**2133-2141.

7. **Bogush, M. L., T. V. Velikodvorskaya, Y. B. Lebedev, L. G. Nikolaev, S. A. Lukyanov, A. F. Fradkov, B. K. Pliyev, M. N. Boichenko, G. N. Usatova, A. A. Vorobiev, G. L. Andersen, and E. D. Sverdlov.** 1999. Identification and localization of differences between *Escherichia coli* and *Salmonella typhimurium* genomes by suppressive subtractive hybridization. Mol Gen Genet **262:**721-9.

8. **Boneca, I. G., O. Dussurget, D. Cabanes, M. A. Nahori, S. Sousa, M. Lecuit, E. Psylinakis, V. Bouriotis, J. P. Hugot, M. Giovannini, A. Coyle, J. Bertin, A. Namane, J. C. Rousselle, N. Cayet, M. C. Prevost, V. Balloy, M. Chignard, D. J. Philpott, P. Cossart, and S. E. Girardin.** 2007. A critical role for peptidoglycan N-deacetylation in *Listeria* evasion from the host innate immune system. Proc Natl Acad Sci U S A **104:**997-1002.

9. **Breton, C., L. Snajdrova, C. Jeanneau, J. Koca, and A. Imberty.** 2006. Structures and mechanisms of glycosyltransferases. Glycobiology **16:**29R-37R.

10. **Brosch, R., J. Chen, and J. B. Luchansky.** 1994. Pulsed-field fingerprinting of listeriae: identification of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. Appl Environ Microbiol **60:**2584-2592.

11. **Buchrieser, C., C. Rusniok, F. Kunst, P. Cossart, and P. Glaser.** 2003. Comparison of the genome sequences of *Listeria monocytogenes* and *Listeria innocua*: clues for evolution and pathogenicity. FEMS Immunol Med Microbiol **35:**207-213.

12. **Call, D. R., M. K. Borucki, and T. E. Besser.** 2003. Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of *Listeria monocytogenes*. J Clin Microbiol **41:**632-9.

13. **Casey, J., and N. Davidson.** 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. Nucleic Acids Res **4:**1539-52.

14. **CDC** Accession date: February 13 2008. January 2008, posting date. Listeriosis General Information. [Online.] http://www.cdc.gov/nczved/dfbmd/disease_listing/listeriosis_gi.html

15. **Chakraborty, T., M. Leimeister-Wachter, E. Domann, M. Hartl, W. Goebel, T. Nichterlein, and S. Notermans.** 1992. Coordinate regulation of virulence genes in *Listeria monocytogenes* requires the product of the prfA gene. J Bacteriol **174:**568-574.

16. **Chambers, H.** 2003. Beta-Lactam Antibiotics & Other Inhibitors of Cell Wall Synthesis., p 754-768. *In* B. Katzung (ed.), Basic & Clinical Pharmacology, Eighth ed. Lange Medical Books/McGraw-Hill Medical Publishing Divsion, New York.

17. **Colburn, K. G., C.A. Kaysner, C. Abeyta Jr and M.M. Wekell.** 1990. *Listeria* species in a California coast estuarine environment. Appl Environ Microbiol **56:**2007-2011.

18. **Cornu, M., M. Kalmokoff, and J. P. Flandrois.** 2002. Modelling the competitive growth of *Listeria monocytogenes* and *Listeria innocua* in enrichment broths. Int J Food Microbiol **73:**261-274.

19. **Coutinho, P. M., E. Deleury, G. J. Davies, and B. Henrissat.** 2003. An evolving hierarchical family classification for glycosyltransferases. J Mol Biol **328:**307-317.

20. **Coutinho, P. M., and B. Henrissat.** 1999. Carbohydrate-active enzymes: an integrated database approach., p. 3-12. *In* H. J. Gilbert, G. Davies, B. Henrissat, and B. Svensson (ed.), Recent Advances in Carbohydrate Bioengineering. The Royal Society of Chemistry, Cambridge.

21. **Diatchenko, L., Y. F. Lau, A. P. Campbell, A. Chenchik, F. Moqadam, B. Huang, S. Lukyanov, K. Lukyanov, N. Gurskaya, E. D. Sverdlov, and P. D. Siebert.** 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. Proc Natl Acad Sci U S A **93:**6025-6030.

22. **Dorrell, N., S. J. Hinchliffe, and B. W. Wren.** 2005. Comparative phylogenomics of pathogenic bacteria by microarray analysis. Curr Opin Microbiol **8:**620-626.

23. **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.

24. **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.

25. **Dziejman, M., E. Balon, D. Boyd, C. M. Fraser, J. F. Heidelberg, and J. J. Mekalanos.** 2002. Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc Natl Acad Sci U S A **99:**1556-15561.

26. **Ejim, L. J., V. M. D'Costa, N. H. Elowe, J. C. Loredo-Osti, D. Malo, and G. D. Wright.** 2004. Cystathionine beta-lyase is important for virulence of *Salmonella enterica* serovar Typhimurium. Infect Immun **72:**3310-3314.

27. **Emmerth, M., W. Goebel, S. I. Miller, and C. J. Hueck.** 1999. Genomic subtraction identifies *Salmonella typhimurium* prophages, F-related plasmid sequences, and a novel fimbrial operon, *stf*, which are absent in *Salmonella typhi*. J Bacteriol **181:**5652-5661.

28. **Farber, J. M., and P. I. Peterkin.** 1991. *Listeria monocytogenes*, a food-borne pathogen. Microbiol Rev **55:**476-511.

29. **Fenlon, D. R., T. Stewart, and W. Donachie.** 1995. The incidence, numbers and types of *Listeria monocytogenes* isolated from farm bulk tank milks. Lett Appl Microbiol **20:**57-60.

30. **Galbraith, E. A., D. A. Antonopoulos, and B. A. White.** 2004. Suppressive subtractive hybridization as a tool for identifying genetic diversity in an environmental metagenome: the rumen as a model. Environ Microbiol **6:**928-37.

31. **Gasteiger E., G. A., Hoogland C., Ivanyi I., Appel R.D., Bairoch A.** Accession date: February 16 2008. April 2003, posting date. ExPASy: the proteomics server for in-depth protein knowledge and analysis. [Online.] http://expasy.org.

32. **Graves, L. M., B. Swaminathan, M. W. Reeves, S. B. Hunter, R. E. Weaver, B. D. Plikaytis, and A. Schuchat.** 1994. Comparison of ribotyping and multilocus enzyme electrophoresis for subtyping of *Listeria monocytogenes* isolates. J Clin Microbiol **32:**2936-2943.

33. **Gray, M. L., and A. H. Killinger.** 1966. *Listeria monocytogenes* and listeric infections. Bacteriol. Rev. **30:**309-382.

34. **Guan, R., and R. A. Mariuzza.** 2007. Peptidoglycan recognition proteins of the innate immune system. Trends Microbiol **15:**127-134.

35. **Hain, T., S. S. Chatterjee, R. Ghai, C. T. Kuenne, A. Billion, C. Steinweg, E. Domann, U. Karst, L. Jansch, J. Wehland, W. Eisenreich, A. Bacher, B. Joseph, J. Schar, J. Kreft, J. Klumpp, M. J. Loessner, J. Dorscht, K. Neuhaus, T. M. Fuchs, S. Scherer, M. Doumith, C. Jacquet, P. Martin, P. Cossart, C. Rusniock, P. Glaser, C. Buchrieser, W. Goebel, and T. Chakraborty.** 2007. Pathogenomics of *Listeria* spp. Int J Med Microbiol **297:**541-557.

36. **Holt, J. J.** 1994. Bergey's Manual of Determinative Bacteriology, 9th ed. p 550., William and Wilkins, Baltimore, MD.

37. **Huang, J., F. Reggiori, and D. J. Klionsky.** 2007. The transmembrane domain of acid trehalase mediates ubiquitin-independent multivesicular body pathway sorting. Mol Biol Cell **18:**2511-2524.

38. **Israel, D. A., N. Salama, C. N. Arnold, S. F. Moss, T. Ando, H. P. Wirth, K. T. Tham, M. Camorlinga, M. J. Blaser, S. Falkow, and R. M. Peek, Jr.** 2001. *Helicobacter pylori* strain-specific differences in genetic content, identified by microarray, influence host inflammatory responses. J Clin Invest **107:**611-620.

39. **Johnson, J., K. Jinneman, G. Stelma, B. G. Smith, D. Lye, J. Messer, J. Ulaszek, L. Evsen, S. Gendel, R. W. Bennett, B. Swaminathan, J. Pruckler, A. Steigerwalt, S. Kathariou, S. Yildirim, D. Volokhov, A. Rasooly, V. Chizhikov, M. Wiedmann, E. Fortes, R. E. Duvall, and A. D. Hitchins.** 2004. Natural atypical *Listeria innocua* strains with *Listeria monocytogenes* pathogenicity island 1 genes. Appl Environ Microbiol **70:**4256-4266.

40. **Kanehisa, M., and S. Goto.** Accession date: March 3 2008. January 1995, posting date. KEGG: Kyoto Encyclopedia of Genes and Genomes. Kanehisa Laboratories.[Online]. http://www.genome.jp/kegg/.

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

42. **Kuhn, M., and W. Goebel.** 2007. Molecular Virulence Determinants of *Listeria monocytogenes*. p 111-156., *In* E. T. Ryser and E. H. Marth (ed.), *Listeria*, Listeriosis and Food Safety, 3rd ed. CRC Press Taylor & Francis Group.

43. **Lampson, B. C., M. Inouye, and S. Inouye.** 2005. Retrons, msDNA, and the bacterial genome. Cytogenet Genome Res **110:**491-499.

44. **Lawrence, L. M., and A. Gilmour.** 1995. Characterization of *Listeria monocytogenes* isolated from poultry products and from the poultry-processing environment by random amplification of polymorphic DNA and multilocus enzyme electrophoresis. Appl Environ Microbiol **61:**2139-2144.

45. **Lee, W. J., D. S. Banavara, J. E. Hughes, J. K. Christiansen, J. L. Steele, J. R. Broadbent, and S. A. Rankin.** 2007. Role of cystathionine beta-lyase in catabolism of amino acids to sulfur volatiles by genetic variants of *Lactobacillus helveticus* CNRZ 32. Appl Environ Microbiol **73:**3034-3039.

46. **Lovering, A. L., L. H. de Castro, D. Lim, and N. C. Strynadka.** 2007. Structural insight into the transglycosylation step of bacterial cell-wall biosynthesis. Science **315:**1402-1405.

47. **Ma, B., J.L. Simala-Grant and D.E. Taylor.** 2006. Fucosylation in prokaryotes and eukaryotes. Glycobiology **16:**158R-184R.

48. **MacGowen, A. P., D. S. Reeves, and J. McLauchlin.** 1990. Antibiotic resistance of *Listeria monocytogenes*. Lancet **336:**513-514.

49. **Macheboeuf, P., C. Contreras-Martel, V. Job, O. Dideberg, and A. Dessen.** 2006. Penicillin binding proteins: key players in bacterial cell cycle and drug resistance processes. FEMS Microbiol Rev **30:**673-691.

50. **Mavrodi, D. V., O. V. Mavrodi, B. B. McSpadden-Gardener, B. B. Landa, D. M. Weller, and L. S. Thomashow.** 2002. Identification of differences in genome content among phlD-positive *Pseudomonas fluorescens* strains by using PCR-based subtractive hybridization. Appl Environ Microbiol **68:**5170-5176.

51. **McLauchin, J., and J. C. Low.** 1994. Primary cutaneous listeriosis in adults: an occupational disease of veterinarians and farmers. Vet. Record **24:**615-617.

52. **Moss, G.** 2000. Recommendation of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the Nomenclature and Classification of Enzymes by the Reactions they Catalyze. Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB).

53. **Murray, E. G. D., R. A. Webb, and M. B. R. Swann.** 1926. A disease of rabbits characterised by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus *Bacterium monocytogenes* (n. sp.). J Pathol Bacteriol **29:**407-439.

54. **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.

55. **Nightingale, K. K., K. Windham, and M. Wiedmann.** 2005. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. J Bacteriol **187:**5537-5551.

56. **O'Donnell, E. T.** 1995. The incidence of *Salmonella* and *Listeria* in raw milk from farm bulk tanks in England and Wales. J. Soc. Dairy Technol. **48:**25-29.

57. **Pagotto, F., N. Corneau, and J. Farber.** 2006. *Listeria monocytogenes* Infections. p. 313-340., *In* H. P. Riemann and D. O. Cliver (ed.), Foodborne Infections and Intoxications, 3rd ed. Elsevier Science Inc.

58. **Pagotto, F., E. Daley, J. Farber, and D. Warburton.** 2001. MFHPB-30 Isolation of *Listeria monocytogenes* from all food and envionmental samples, vol. 2. Bureau of Microbial Hazards, Food Directorate, Health Canada. Avaialble online: http://www.hc-sc.gc.ca/fn-an/res-rech/analy-meth/microbio/volume2/mfhpb30-01-eng.php.

59. **Pagotto, F., L. K. Ng, C. Clark, J. Farber, and C. P. H. L. N.** 2006. Canadian listeriosis reference service. Foodborne Pathog Dis **3:**132-137.

60. **Painter, J., and L. Slutsker.** 2007. Listeriosis in Humans. p. 85-110., *In* E. T. a. M. Ryser, E. H. (ed.), *Listeria*, Listeriosis and Food Safety, 3rd ed. CRC Press Taylor & Francis Group.

61. **Pohl, M. A., M. Wiedmann, and K. K. Nightingale.** 2006. Associations among *Listeria monocytogenes* genotypes and distinct clinical manifestations of listeriosis in cattle. Am J Vet Res **67:**616-626.

62. **Promadej, N., F. Fiedler, P. Cossart, S. Dramsi, and S. Kathariou.** 1999. Cell wall teichoic acid glycosylation in *Listeria monocytogenes* serotype 4b requires *gtcA*, a novel, serogroup-specific gene. J Bacteriol **181:**418-425.

63. **Ramaswamy, V., V. M. Cresence, J. S. Rejitha, U. Lekshmi, K. S. Dharsana, P. S.P., and H. M. Vijila.** 2007. *Listeria*-Review of Epidemiology and Pathogenesis. J Microbiol Immunol Infect **40:**4-13.

64. **Rocourt, J., and C. Buchrieser.** 2007. The Genus *Listeria* and *Listeria monocytogenes*: Phylogenetic Position, Taxonomy and Identification. p. 1-20., *In* E. T. Ryser and E. H. Marth (ed.), *Listeria*, Listeriosis and Food Safety, 3rd ed. CRC Press Taylor & Francis Group.

65. **Sabet, C., M. Lecuit, D. Cabanes, P. Cossart, and H. Bierne.** 2005. LPXTG protein InlJ, a newly identified internalin involved in *Listeria monocytogenes* virulence. Infect Immun **73:**6912-6922.

66. **Safdar, A., and D. Armstrong.** 2003. Antimicrobial activities against 84 *Listeria monocytogenes* isolates from patients with systemic listeriosis at a comprehensive cancer center (1955-1997). J Clin Microbiol **41:**483-485.

67. **Saunders, N. A., A. Underwood, A. M. Kearns, and G. Hallas.** 2004. A virulence-associated gene microarray: a tool for investigation of the evolution and pathogenic potential of *Staphylococcus aureus*. Microbiol **150:**3763-3771.

68. **Schaferkordt, S., E. Domann, and T. Chakraborty.** 1998. Molecular Approaches for the Study of *Listeria*, p. 421-431., *In* P. Williams, J. Ketley, and G. Salmond (ed.), Methods in Microbiology, vol. 27. Academic Press Ltd.

69. **Scheffers, D. J., and M. G. Pinho.** 2005. Bacterial cell wall synthesis: new insights from localization studies. Microbiol Mol Biol Rev **69:**585-607.

70. **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.

71. **Skovgaard, N., and C. A. Morgen.** 1988. Detection of *Listeria* spp. in faeces from animals, in feeds, and in raw foods of animal origin. Int J of Food Microbiol **6:**229-242.

72. **Stratagene.** 2006. SpotReport® Alien® cDNA Array Validation System Instruction Manual, Revision #050062f ed. Stratagene.

73. **Taboada, E. N., R. R. Acedillo, C. D. Carrillo, W. A. Findlay, D. T. Medeiros, O. L. Mykytczuk, M. J. Roberts, C. A. Valencia, J. M. Farber, and J. H. E. Nash.** 2004. Large-Scale Comparative Genomics Meta-Analysis of *Campylobacter jejuni* Isolates Reveals Low Level Genome Plasticity. J Clin Microbiol **42:**4566-4576.

74. **Taylor, J. G., X. Li, M. Oberthur, W. Zhu, and D. E. Kahne.** 2006. The total synthesis of moenomycin A. J Am Chem Soc **128:**15084-15085.

75. **Van Renterghem, B., F. Huysman, R. Rygole, and W. Verstraete.** 1991. Detection and prevalence of *Listeria monocytogenes* in the agricultural ecosystem. J Appl Bacteriol **71:**211-217.

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

77. **Vazquez-Boland, J. A., M. Kuhn, P. Berche, T. Chakraborty, G. Dominguez-Bernal, W. Goebel, B. Gonzalez-Zorn, J. Wehland, and J. Kreft.** 2001. *Listeria* pathogenesis and molecular virulence determinants. Clin Microbiol Rev **14:**584-640.

78. **Watkins, J., and K. P. Sleath.** 1981. Isolation and enumeration of *Listeria monocytogenes* from sewage, silage sludge and river water. J of Appl Bacteriol **50:**1-9.

79. **Weimer, B., K. Seefeldt, and B. Dias.** 1999. Sulfur metabolism in bacteria associated with cheese. Antonie Van Leeuwenhoek **76:**247-261.

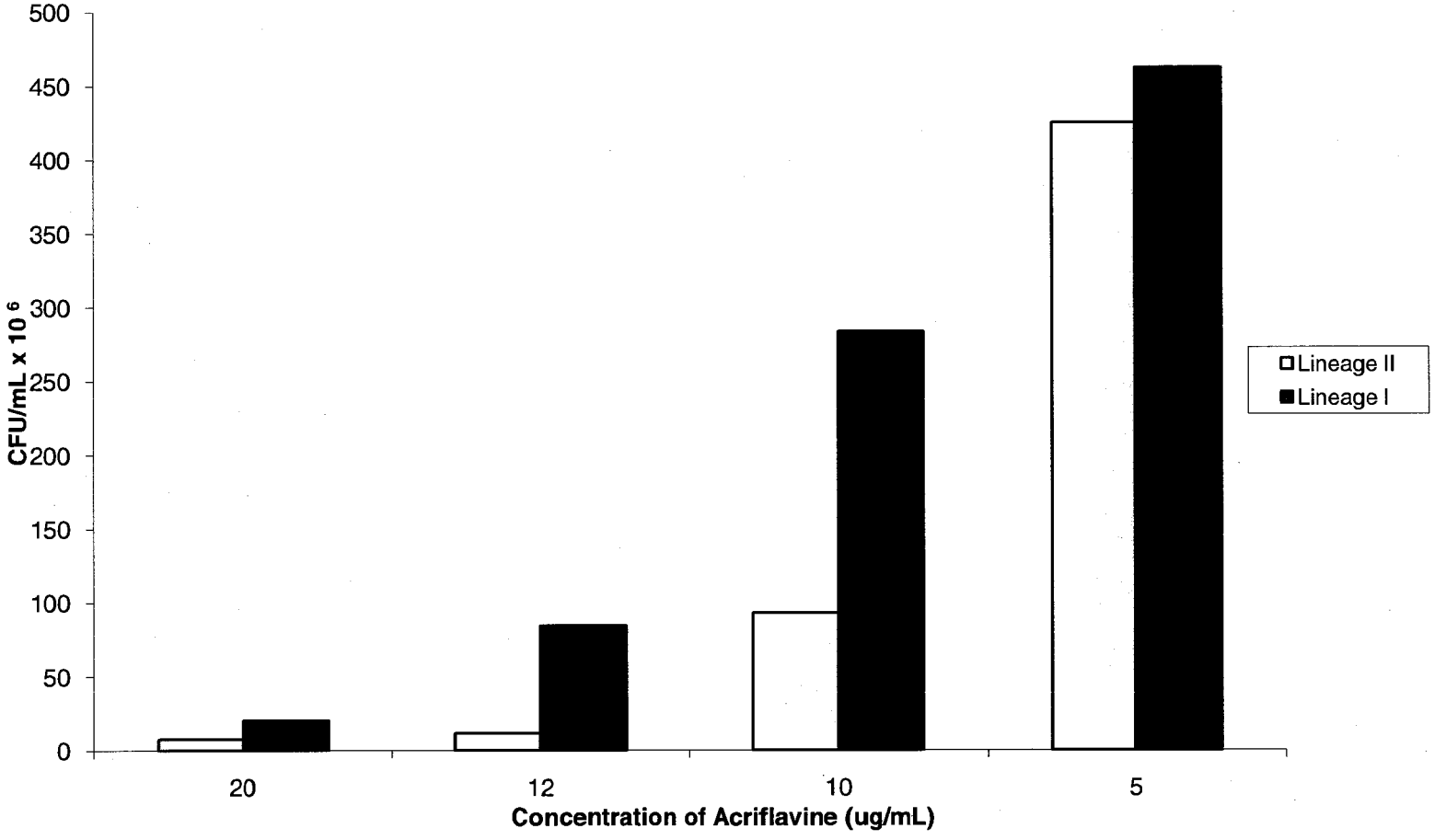
80. **Wuenscher, M. D., S. Kohler, A. Bubert, U. Gerike, and W. Goebel.** 1993. The *iap* gene of *Listeria monocytogenes* is essential for cell viability, and its gene product, p60, has bacteriolytic activity. J Bacteriol **175:**3491-3501.

81. **Zawadzka-Skomial, J., Z. Markiewicz, M. Nguyen-Disteche, B. Devreese, J. M. Frere, and M. Terrak.** 2006. Characterization of the bifunctional glycosyltransferase/acyltransferase penicillin-binding protein 4 of *Listeria monocytogenes*. J Bacteriol **188:**1875-1881.

82. **Zenewicz, L. A., and H. Shen.** 2007. Innate and adaptive immune responses to *Listeria monocytogenes*: a short overview. Microbes Infect. **9:** 1208-1215

Appendix A: Average Growth of 45 Lineage I and Lineage II Isolates of *L. monocytogenes* in Acriflavine

Averages were calculated from one experimental trial.


500
450
400
350
300
250
200
150
100
50
0
CFU/mL x 10 6
20
12
10
5
Concentration of Acriflavine (ug/mL)
Lineage II
Lineage I

Appendix B. Misannotation of *lmof2365_2155* in *L. monocytogenes*

At the time this study was performed, the gene *lmof2365_2155* was annotated in the *L. monocytogenes* genome F2365 (serovar 4b) as a glycosyltransferase family 65, with an IUBMB EC number 2.4.1.8. In the KEGG, MetaCyc and BRENDA Gene Ontology Databases EC number 2.4.1.8 is listed as a glycosyltransferase. In our analyses of the amino acid sequence, it was immediately observed that prokaryotic organisms did not have a glycosyltransferase belonging to family 65 (see discussion, page 103). Serendipitously, one of the thesis defence members, an expert in prokaryotic glycoconjugates, also pointed this out and suggested that it might have been an annotation error.

After the successful defence, it was investigated further, respecting the time limits for submission of corrected thesis. While *lmof2365_2155* is listed in the CAZY database under the glycoside hydrolase family 65, our laboratory received the following from the Microbial Genomes Curator of the National Center for Biological Information: "The protein in question comes from automatic annotation for WGS genomes. Unfortunately, it is not curated, nor corrected due to the constant revisions that occur for WGS genomes. Links from numerous sources point from this 'glycosyl transferase family 65' protein to this entry: glycoside hydrolase family 65 (http://www.cazy.org/fam/GH65.html), which suggests there was an error at some point in the annotation. We will correct the incorrect annotations for proteins derived from the complete genome set."

Because there were over 12 hits with *Listeria* genomes sequenced by TIGR, the Institute for Medical Microbiology (Giessen, Germany) and the *Listeria* European Consortium (Pasteur Institute, France) with the translated sequence found in this work, it was decided to describe it as a putative glycosyltransferase. However, cautious diligence

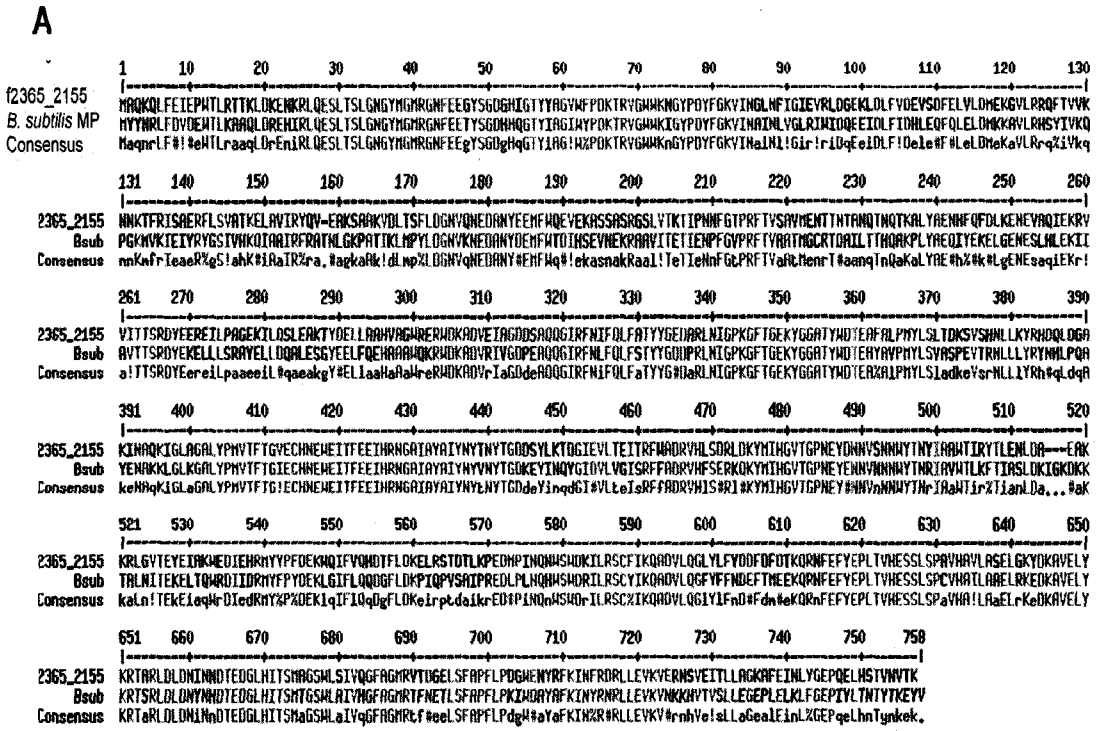
(see page 109) was noted to highlight that it may in fact be a hydrolase, and further work was required to ascertain its true function.

Bacillus subtilis, closely related to *L. monocytogenes*, has several glycoside hydrolase family 65 genes. When a BLAST search is performed on the amino acid sequence of lmof2365_2155, the first non-*Listeria* sequence showing homology to f2365_2155 is the yvdK gene of *B. subtilis*, followed by the maltose phosphorylase gene of *B. subtilis*. When these amino acid sequences were aligned with lmof2365_2155 they showed 60% homology, indicating that there is similarity between these genes (Appendix C).

The misannotation of *lmof2365_2155* further enforces the concept that in order to fully comprehend the function of an enzyme, substrate/donor and product specificity assays need to be performed on the purified form of the enzyme. The use of sequence data alone cannot reliably predict the role of an enzyme within a biological system.

Appendix C. Amino acid Sequence Alignments of lmof2365_2155 to Glycoside hydrolase Family 65 proteins in *B. subtilis*

(A) Alignment of lmof2365_2155 to *B. subtilis* maltose phosphorylase,
(B) Alignment of lmof2365_2155 to *B. subtilis* yvdK.


A
f2365_2155
B. subtilis MP
Consensus
B
f2365_2155
B. subtilis yvdk
Consensus