CONTRIBUTION OF SIGMA B TO STRESS RESPONSE AND VIRULENCE IN *LISTERIA MONOCYTOGENES* LINEAGES I, II, IIIA, AND IIIB

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

HALEY ANN FRANKS OLIVER

August 2009

CONTRIBUTION OF SIGMA B TO STRESS RESPONSE AND VIRULENCE IN *LISTERIA MONOCYTOGENES* LINEAGES I, II, IIIA, AND IIIB

HALEY ANN FRANKS OLIVER, PhD

Cornell University 2009

The widespread presence of *Listeria monocytogenes* in diverse environments, including those that are natural (i.e., non-agricultural), agricultural, and food-associated, suggests that these environments may serve as sources or reservoirs of *L. monocytogenes* that can be transmitted to various hosts, including humans. The vast majority of human listeriosis infections are recognized to occur through consumption of contaminated foods. Phylogenetic analysis of *L. monocytogenes* strains has identified three distinct lineages. While lineages I and II are both common among human clinical and food isolates, lineage I strains are overrepresented among clinical isolates, and lineage II strains are overrepresented among food and environmental isolates. Lineage III, which includes subgroups IIIA and IIIB, is rare and predominantly associated with animal disease. σ^B, encoded by *sigB*, is a sigma factor previously demonstrated to critically contribute to stress response and virulence in lineage II strains. We used transcriptomic and phenotypic analyses to characterize the role of σ^B in *L. monocytogenes* strains representing lineages I, II, IIIA, and IIIB. Whole-genome expression microarrays, phenotypic assays, and the guinea pig gastrointestinal model for listeriosis were used to characterize the role of σ^B stationary phase wildtype and isogenic Δ*sigB* mutants representing *L. monocytogenes* diversity. Our results indicate that the role of σ^B may differ among *L. monocytogenes* strains. The stationary phase transcriptome and σ^B regulon of *L. monocytogenes* 10403S was

also defined using RNA sequencing (RNA-Seq) with the Illumina Genome Analyzer. We found that 83% of all genes were transcribed in stationary phase and a total of 96 genes had significantly higher transcript levels in 10403S than in Δ*sigB*, indicating σ^B-dependent transcription of these genes. RNA-Seq analyses suggest that a total of 65 noncoding RNA molecules (ncRNAs) were transcribed in stationary phase. The RNA-Seq data also enabled annotation of putative operons and visualization of transcription start and stop sites. The results from these studies suggest that σ^B contributes to a complex network of transcriptional regulators which allows *L. monocytogenes* to survive stress and subsequently cause disease and RNA-Seq allows quantitative characterization of prokaryotic transcriptomes and is a new strategy for exploring transcriptional regulatory networks in bacteria.

BIOGRAPHICAL SKETCH

Haley Oliver began her Ph.D. program with Dr. Kathryn Boor at Cornell University in August 2004. Haley has been involved in numerous extracurricular and professional development activities including a summer internship with the FDA Center for Veterinary Medicine, Food Science Bioexplorations, Cornell Graduate School Iron Chef, Cornell University Infection and Pathobiology Program, College of Agriculture and Life Sciences Leadership Development Program, ASM Kadner Institute for Graduate Students and Postdoctoral Scientists in Preparation for Careers in Microbiology, and is an active member of IFT, ASM, and IAFP. As a part of her graduate research experience, Haley has contributed numerous hours to three undergraduate research assistants. Two of these students were participants in the Department of Food Science Summer Scholars Program and the distinguished Hughes Scholars Program and as a result, they are pursuing research as a career. While at Cornell, she has received American Society for Microbiogy Corporate Activities Program Graduate Student Travel Grants, the Eastern New York Branch ASM Graduate Travel Grant, the IFT Food Microbiology Division Graduate Fellowship, IFT Frito Lay Graduate Fellowships, Clinton DeWitt Smith Graduate Research Fellowships, The Kosi Award in Food Science, V. Duane Rath Graduate Research Fellowship, Cornell University Provost's Diversity Fellowship, IAFP Foundation Student Travel Scholarship, IFT GMA Graduate Fellowship, and the Olin Graduate Research Fellowship. Further, Haley was recently selected to serve as one of two Cornell University Ph.D. Marshals for the May 2009 Cornell University commencement ceremonies.

ACKNOWLEDGMENTS

While investigating graduate programs which offered studies in food safety in the fall of 2003, I came across the "Food Safety Lab at Cornell University" homepage. I pined over the email that I sent to Dr. Kathryn J. Boor detailing my research experience and graduate school goals in hope of learning more about the program. The effort was worth her response which invited me to visit the Department of Food Science to meet the Food Safety Lab. I express my sincerest thanks to Kathryn Boor and Martin Wiedmann as the opportunity to study under them and with all members of the Food Safety Lab has been the gift of a lifetime and I expect to continuously discover new appreciations for my experiences as time passes.

I also express thanks to Dr. Yrjo Grohn and Dr. Ray Glahn who served as my Epidemiology and Field-appointed committee members, respectively. Their areas of expertise and therefore outside perspectives challenged my work, my thought process, and rational and as a result has made me a better scientist.

Over the past five years, I have had the opportunity to work with an outstanding, yet ever-changing, team of colleagues at the FSL. While space prevents me from personally thanking each member, all of whom contributed positively to my experience at Cornell, I would like to highlight a few individuals who are not only excellent scientists, but lifetime friends. Dr. Kendra Nightingale is someone who everyone should strive to be—enthusiastic, hard-working, and a tried-and-true "fun-haver" (and did I mention an outstanding scientist?). Yeşim Soyer (my dear Turkish friend) and Soraya Chaturongakul (whose name I love to spell) are my far away forever-friends. I admire your strength and I thank you for the worldliness you brought to my life. I wish you all the best in your life charges to establish FSL Turkey

and FSL Thailand and hope I can replay all you've done for me. Esther Fortes keeps the FSL boat afloat, she is a grand friend, and I will never forget her encouragement at the "fitness center" where we have and will continue to spend quality time.

I thank my family for their patience, time, love, and support. Each member has listened to their less-than-fair share of food safety ranting and raving and I hope that you learned a little despite it being all-too-often focused on σ^B, *Listeria monocytogenes*, and gene regulation. My sister, Dink (Cora), has been an awesome part of my college life since she was 8. I loved her company in Laramie during undergraduate days, our visit to Ithaca in April 2004, our trek to Ithaca in a moving van, and I thank her for spending the summer of 2009 with me (which I hope to be the best visit ever!). Thanks to Matty for never failing to call when I needed someone to run interference on my mood and to Graham for all his hard work and sacrifice during the summer 2008. My mom has kept me grounded throughout my Ph.D. She has always pushed me to be better than my best and has taught me to work through difficult situations one step at a time. Finally, I thank Dusty Oliver for his patience, understanding, love, and selflessness. He allowed me the freedom to pursue my ambitions and has given more of himself than anyone person should. Thanks Bud!

Finally, I thank my dad for his encouragement and imparting his love for learning. I miss you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

L. monocytogenes	*Listeria monocytogenes*
L. innocua	*Listeria innocua*
L. seeligeri	*Listeria seeligeri*
L. ivanovii	*Listeria ivanovii*
MPN	most probable number
CFU	colony forming unit
MLEE	multilocus enzyme electrophoresis
PFGE	pulsed-field gel electrophoresis
MLST	multilocus sequence-based typing
spp.	species
EC	epidemic clone
RTE	ready-to-eat
HI	Hybridization Index
GEI	Gene Expression Index
RNA-Seq	RNA Sequencing
ncRNA	noncoding RNA
RACE-PCR	Rapid Amplification of cDNA Ends PCR
FDR	False Discovery Rate
HMM	Hidden Markov Model

CHAPTER ONE

Environmental Reservoir and Transmission into the Mammalian Host

ABSTRACT

The widespread presence of *Listeria monocytogenes* in various diverse environments, including those that are natural (i.e., non-agricultural), agricultural, and food-associated, suggests that these environments may serve as sources or reservoirs of *L. monocytogenes* that can be transmitted to various hosts, including humans. As the vast majority of human listeriosis infections are recognized to occur through consumption of contaminated foods, and as animal listeriosis infections also appear to be predominantly feedborne, development of effective intervention strategies for reducing the incidence of listeriosis among susceptible human and animal populations will require elucidation of specific routes of *L. monocytogenes* transmission among different ecosystems and compartments within food and feed production systems. Current knowledge of *L. monocytogenes* ecology is presented to provide insight into the primary sources that appear to contribute to its introduction into human food-associated environments and foods as well as its transmission among various compartments in food and agricultural production systems.

INTRODUCTION

Listeria species, including *L. monocytogenes*, are often described as ubiquitous in nature as they have been isolated from a diverse array of natural, man-made, agricultural, and food-associated environments [1-5]. The vast majority of human listeriosis infections (99%) are foodborne [6], while animal listeriosis infections appear to be predominantly feedborne. An emerging understanding of *L. monocytogenes* ecology has provided insight into the primary sources that appear to contribute to its introduction into human food-associated environments and foods as well as its transmission among various compartments in food production systems. As farm ruminants represent the mammalian hosts most commonly affected by clinical listeriosis, ruminant animal agricultural systems are likely to serve as reservoirs or sources of *L. monocytogenes* that are transmitted into the human food chain. Therefore, for a complete understanding of *L. monocytogenes* transmission to human food and humans, it is also important to understand the transmission of *L. monocytogenes* into other mammals, and particularly farm ruminants. This chapter reviews our knowledge of the ecology and transmission of *L. monocytogenes* in natural, non-agricultural environments, agricultural environments, and food-associated environments and addresses the potential contributions of these environments as sources or reservoirs of *L. monocytogenes* that can be transmitted to mammalian and particularly human hosts. We propose that, given the ability of *L. monocytogenes* to survive for prolonged time periods in many different environments, as well as its wide distribution and prevalence in different environments, it is likely that selective pressures associated with different environments play an important role in the evolution of this pathogen. We further propose that humans likely represent an accidental host for this environmental pathogen and that the true importance of most

virulence factors may lie in their role in enhancing *L. monocytogenes* survival in non-human host-associated environments [7].

METHODS FOR STUDYING *LISTERIA MONOCYTOGENES* TRANSMISSION

Studies on transmission of any pathogen are critically dependent on our ability to reliably detect, and ideally, quantitate, a given pathogen in different environments and hosts. Methodological capabilities for accurately recovering and characterizing the representative diversity of a pathogen present in a given environment determine the quality of the information obtained. Since limitations in detection and subtyping methods critically impact our ability to understand the ecology of *L. monocytogenes*, we will briefly review commonly used detection and subtyping methods for *L. monocytogenes*, including their relevant limitations. A number of recent reviews and book chapters provide more comprehensive coverage of *L. monocytogenes* detection and subtyping methods [8-11].

Detection methods for *L. monocytogenes.* Despite their wide distribution in nature, *L. monocytogenes* and other *Listeria* spp. usually occur in small numbers, within the context of large numbers of other microorganisms, in most natural habitats. Therefore, detection methods for *Listeria* spp. typically include a selective enrichment step to allow amplification of the small numbers of *Listeria* spp. initially present, followed by plating on selective and differential media to enable their detection. This strategy is prone to providing false negative results, particularly if the *Listeria* spp. present in the sample are injured prior to exposure to the selective medium. On the other hand, use of a non-selective pre-enrichment step may allow other microorganisms to overgrow *Listeria* spp. in a given sample, thus also yielding false negative results. While these issues have led to the development of a variety of

different enrichment media and procedures for *Listeria* spp., in general, a single enrichment procedure is unlikely to detect all *L. monocytogenes* that may be present in a given set of samples [11]. For example, *L. innocua* has been shown to out-compete *L. monocytogenes* during some enrichment procedures [12] and different enrichment media appear to favor recovery of different bacterial subtypes from the same sample [13, 14]. As a consequence, the use of a combination of different enrichment and plating procedures in parallel will provide the most sensitive detection of *Listeria* spp., however, this approach is usually cost-prohibitive and therefore not practical. Since many environments can contain multiple *Listeria* species and/or multiple *Listeria* strains [14-16] and because *L. monocytogenes* can be overgrown by other *Listeria* spp. during enrichment, *L. monocytogenes* prevalences reported for different environments likely underestimate the true prevalence and diversity of *L. monocytogenes* in a given sample.

Quantitative data on the presence of *Listeria* spp. and *L. monocytogenes* are important for understanding the ecology and transmission of *Listeria.* Most probable number (MPN) methods are generally used for quantification since *Listeria* populations in most environments are usually <100 CFU g^{-1} [11]. A paucity of quantitative data on *L. monocytogenes* loads in different environments [17] other than in food samples [18] exists due to the labor- and cost-intensive nature of MPN methods.

Subtyping methods for *L. monocytogenes.* Application of subtyping methods is critical to our ability to understand the ecology and transmission of *L. monocytogenes*. While serotyping is commonly used to characterize *L. monocytogenes*, it provides limited discriminatory power as only 13 *L. monocytogenes* serotypes can be differentiated with three serotypes (1/2a, 1/2b, and 4b) representing the vast majority of human listeriosis isolates. Multilocus enzyme electrophoresis (MLEE) and phage

typing provided initial insight into the population genetics and transmission of *L. monocytogenes.* Molecular subtyping methods, including ribotyping [19, 20], pulsed-field gel electrophoresis (PFGE) [21, 22], and more recently, multilocus sequence-based typing (MLST) [23, 24], have provided recent advances in our understanding of *L. monocytogenes* ecology and transmission and have been used in many studies on *L. monocytogenes* ecology reported since 1995. While all three molecular methods provide discriminatory power, PFGE was shown to be more discriminatory for *L. monocytogenes* than MLST or ribotyping. MLST and ribotyping have provided important and relevant subtyping information, however, including identification of epidemic clones and virulence attenuated subtypes [25-27].

Characterization of *L. monocytogenes* isolates from a variety of different hosts and environments by a variety of different subtyping methods, including initial MLEE work by Pifaretti et al. [28], has also shown that strains comprising the species *L. monocytogenes* represent at least three distinct genetic lineages. While different nomenclatures have been used to designate these *L. monocytogenes* lineages [2], the main lineages described in different studies appear to be identical, as supported by consistent grouping of specific *L. monocytogenes* serotypes into lineages [28, 29]. Based on the lineage designations used by most groups [25, 30, 31], lineage I predominantly includes serotypes 1/2b, 3b, 3c, and 4b strains, while lineage II primarily includes serotypes 1/2a, 1/2c, and 3a [29]. Interestingly, previous reports have shown that lineage I strains are significantly overrepresented among human clinical listeriosis cases as compared to their prevalence among animal listeriosis cases and contaminated foods [25, 29, 32]. On the other hand, lineage II strains show a significantly higher prevalence among food isolates and animal clinical cases than among human listeriosis cases [25, 33]. In addition, lineage I isolates appear to have significantly greater pathogenic potential, as determined by their ability to spread to

neighboring host cells in a cell culture plaque assay, when compared to lineage II isolates [25, 32]. Lineage III predominantly includes serotypes 4a and 4c, as well as some serotype 4b strains that are distinct from those grouped into lineage I [29]. Strains classified in lineage III appear to be associated with isolation from animals and are occasionally isolated from human listeriosis cases with clinical disease, but are rarely isolated from foods [25, 33]. Increasing evidence thus exists that *L. monocytogenes* strains represent multiple lineages that appear to differ in their abilities to be transmitted to humans, as also supported by recent subtype-specific mathematical modeling data, which indicate that the likelihood of human disease caused by *L. monocytogenes* classified into different lineages can differ by more than 2 logs [34].

LISTERIA MONOCYTOGENES IN THE NATURAL AND OTHER NON-AGRICULTURAL ENVIRONMENTS

While most studies on the presence of *L. monocytogenes* in different environments have focused on food-associated and farm environments [35, 36], multiple studies have reported that *L. monocytogenes* are common in natural and other non-agricultural environments, and can also survive for extended time periods in soil and water. Recent subtyping studies also indicate that at least some of the subtypes found in natural and other non-agricultural environments are also found among human listeriosis cases, indicating that these environments may represent a source of human pathogenic *L. monocytogenes* subtypes.

***Listeria monocytogenes* prevalence and load in natural and other non-agricultural environments.** Initial studies by Welshimer et al. [37], which investigated the presence of *L. monocytogenes* in soil and plant materials from agricultural and nonagricultural environments, reported that 6 of 7 nonagricultural sites were positive

for *L. monocytogenes* during the spring but not the fall. While this study has to be interpreted carefully, since all recognized *Listeria* spp. were classified as *L. monocytogenes* at that time, it provided initial evidence for the presence of *Listeria* spp. in natural environments. Later, Weiss and Seeliger [38] found *Listeria* spp. from plant samples collected from cornfields (9.7% of samples were positive), grain fields (13.3%), cultivated fields (12.5%), uncultivated fields (44%), meadows and pastures (15.5%), forests (21.3%), and wildlife feeding areas (23.1%) in southern Germany. While the original paper by Weis and Seeliger reported these numbers as *L. monocytogenes* prevalence, only 37 of 103 *Listeria* isolates elicited disease consistent with listeriosis in their mouse bioassay, suggesting that as many as 64% of their isolates may have been *Listeria* spp. other than *L. monocytogenes* or *L. ivanovii*. Fenlon et al. [39] did not find any *L. monocytogenes* in the soil samples associated with vegetable crops that they examined, but they isolated *L. monocytogenes* from soil collected from fields where cattle or sheep fed silage diets had been kept, indicating the importance of animals as sources of *L. monocytogenes* in soil and on plants. One of the most comprehensive, recent studies on *L. monocytogenes* and *Listeria* spp. prevalences in different environments was conducted by Sauders [40] who tested approximately 900 samples from each natural and urban environments (e.g., soil, water, and plant materials) in New York State over a 2 year period. In this study, *L. monocytogenes* prevalences were significantly higher in urban environments (7.5%) as compared to natural environments (1.4%). Additional recent studies have further confirmed that *L. monocytogenes* can be found in a number of different natural and non-agricultural environments, including surface waters [41, 42], estuarine environments [43], and sewage [44-50].

Overall, most studies on *Listeria* in non-farm-associated natural environments indicate that *L. monocytogenes* is found at a lower prevalence than other *Listeria* spp.

[40, 49]. Interestingly, MacGowan et al. [49] found that *L. seeligeri* was the *Listeria* spp. most frequently isolated from soils, while Sauders [40] found that *L. seeligeri* was the most common *Listeria* spp. isolated from both urban and natural environments. This is particularly intriguing since *L. seeligeri*, while not considered a mammalian pathogen, does contain the *Listeria* virulence gene cluster and is hemolytic, which suggests a potential role of at least some *Listeria* virulence genes for survival in selected environments.

***Listeria monocytogenes* growth and survival in natural and other non-agricultural environments.** Some of the first efforts to characterize *L. monocytogenes* as a naturally occurring, saprophytic organism took place in the late 1950s. H.J. Welshimer conducted a study on *L. monocytogenes'* survival and concluded that *L. monocytogenes* could survive for at least 295 days in certain types of soil under defined conditions [51]. Another study [52] also showed that *L. monocytogenes* was able to survive, and in some instances, multiply, in non-sterilized (i.e., natural) and sterilized soil and water at ambient winter temperatures ranging from –15° to +18°C. In this study, sterile and natural soil both supported *L. monocytogenes* survival and growth. For example, *L. monocytogenes* inoculated into a sterile soil suspension at approximately 10^5 CFU ml^{-1} increased up to 2.14 x 10^7 CFU ml^{-1} over a 154 day period. These data provide evidence of the ability of *L. monocytogenes* to survive and multiply in different niches in natural and other non-agricultural environments.

Subtype analysis of *L. monocytogenes* found in natural and non-agricultural environments. Early serotype analysis by Weis and Seeliger [38] of *L. monocytogenes* isolated from natural and agricultural fields found that serotypes 1/2b and 4b were the two most prevalent serotypes isolated from soil and plant samples, which provided initial evidence that human pathogenic *L. monocytogenes* may be

present in diverse environments, since serotypes 1/2b and 4b are commonly associated with human disease [53]. Ribotype analysis of 80 *L. monocytogenes* isolates collected from urban and natural environments by Sauders et al. [54] also found that a number of ribotypes identified among these isolates had previously been linked to human listeriosis cases, including outbreaks. Specifically, a number of isolates from urban sites and a single isolate from a natural environment were ribotype DUP-1038B, which represents a subtype classified into one of the three *L. monocytogenes* epidemic clones (ECII) [2] that has been associated with multiple human listeriosis outbreaks [33, 55]. To illustrate, DUP-1038B was the predominant ribotype isolated over more than a year from multiple sites in a single urban environment, indicating the persistence of this subtype in an urban environment. Overall, these data provide evidence that urban environments represent sources of human pathogenic *L. monocytogenes* strains. Interestingly, the vast majority of isolates (>90%) from natural environments were classified as *L. monocytogenes* lineage II. The lineage II classification is significantly less common among human isolates than classification into lineage I [54] and strains in lineage II appear less likely to cause human disease as compared to those in lineage I [34]. The ribotype most frequently isolated from natural environments (DUP-1039C) is also commonly found in farm environments and animal listeriosis cases, as well as infrequently among human cases, supporting the hypothesis that the natural environment represents a source of animal and human pathogenic *L. monocytogenes*, even though many of the strains found in natural environments may be less virulent for humans than strains found in other environments (e.g., the urban environment). Furthermore, the presence in natural environments of a ribotype that is commonly found in farm environments could also indicate that farms, farm environments, and farm animals may represent a source of *L.*

monocytogenes introduction into natural environments, e.g., via runoff from farms or animal movement.

LISTERIA MONOCYTOGENES IN AGRICULTURAL ENVIRONMENTS

Listeriosis was first observed in animals (i.e., rabbits) in 1926 [56] and has since been described in a number of domesticated and wild animals [57-59]. Most reported animal listeriosis cases have occurred in farm ruminants, including cattle, goats, and sheep, therefore, most available information on *L. monocytogenes* in agricultural environments focuses on the presence and ecology of *L. monocytogenes* in ruminants and on ruminant farms. Overall, *L. monocytogenes* prevalence in ruminants and on ruminant farms varies, but appears to be highest in animals fed silage (fermented plant material, such as grass, hay or chopped field corn [36]). Prevalence in farm environments and in fecal material of silage fed animals appears to, on average, exceed 20% [36] and often includes human disease-associated *L. monocytogenes* subtypes. Farm environments and farm animals may thus be an important source, and potentially a reservoir, of human pathogenic *L. monocytogenes*.

***Listeria monocytogenes* prevalence and load in agricultural environments.** A umber of studies [36, 38, 57, 60, 61] have reported that *L. monocytogenes* is commonly present throughout the agricultural environment, particularly in environments associated with ruminants, including farm soil, vegetation and water, as well as in animal feeds (especially silage), in fecal material, and on animal hides and external surfaces. The presence of *L. monocytogenes* in silage and further dispersal through fecal shedding in ruminant-associated agricultural environments likely have the greatest impacts on transmission of *L. monocytogenes*, both within animal populations as well as from animal populations and farm environments to humans. Fecal shedding is particularly likely to contribute to environmental dispersal of *L.*

monocytogenes, i.e., onto plant materials and fields that may be a source of animal feed or human food (e.g., vegetables). A number of studies have shown that the prevalence of fecal shedding in farm ruminants can range from a few percent of total animals [62, 63] to more than fifty percent [64], with a higher prevalence of fecal shedding in silage fed animals [65]. For example, a recent large study on *L. monocytogenes* prevalence and ecology in farm animals and farm environments found an average fecal prevalence of 20.2% among cattle [36]. While fecal shedding can occur in clinical listeriosis cases, most cattle with fecal samples that test positive for *L. monocytogenes* do not show listeriosis symptoms [66, 67]. On the other hand, small ruminants (e.g., sheep and goats) generally show a lower *L. monocytogenes* prevalence in fecal samples, most likely since silage feeding, and hence, exposure to *L. monocytogenes*, is less common on small ruminant farms. Quantitative data on *L. monocytogenes* levels in ruminant fecal samples are limited, however, Fenlon et al. [39] reported *L. monocytogenes* levels as high as 5.0×10^2 CFU g^{-1} among cattle fed silage, while fecal levels among grazing cattle were 0.4 CFU g^{-1}. Oral exposure to *L. monocytogenes* appears to be a critical risk factor for fecal shedding of this pathogen, but other external factors, such as stress, e.g., climate and feed changes, transport and changes in immunological state, such as pregnancy, may also enhance the likelihood of fecal shedding in ruminants [39, 66-68].

Similar to human listeriosis, animal listeriosis appears to be predominantly a feed-borne disease, with consumption of silage and particularly improperly fermented (and *L. monocytogenes* contaminated) silage, as a major risk factor for clinical listeriosis in ruminants [69] and for *L. monocytogenes* fecal shedding in ruminants [36] as well as *L. monocytogenes* presence in raw milk [70]. Silage is commonly used for feed in modern ruminant production, and particularly for dairy cattle feed, due to its year-round availability. While properly fermented silage has a pH of $\leq$ 4.5 which

helps to inhibit growth of spoilage microorganisms and pathogens, including *L. monocytogenes*, improperly fermented silage often has an elevated pH (e.g., > 5.5) which allows for the growth of spoilage organisms and pathogens [16]. Since crops used for silage may be contaminated with *L. monocytogenes* prior to harvest through a variety of pathways, including fecal deposition by wild [57] or farm animals, contaminated soil [39], or deposition of sewage sludge and manure, *L. monocytogenes* is commonly found in poorly fermented silage. As silage is an important source of *L. monocytogenes* infection in ruminants, a number of studies on *L. monocytogenes* prevalence and loads in silage have been conducted. Overall, *L. monocytogenes* has been isolated from up to 44% of silage samples tested [57] with widely ranging bacterial loads [71, 72] as high as 1.0×10^8 CFU g^{-1} reported [20]. Considering the high *L. monocytogenes* prevalence and densities that can be found on ruminant farms [36], these environments represent a likely point of introduction for *L. monocytogenes* into the human food chain through a variety of pathways, including through use of contaminated manure for fertilization of human food crops, consumption of animal products lacking a listeriocidal heat treatment (e.g., raw milk) and transmission of the organism via fomites into food processing environments, where *L. monocytogenes* may subsequently persist for extended time periods, thus enabling re-contamination of processed foods.

In addition to ruminant species, *L. monocytogenes* also can be isolated from a number of non-ruminant species and non-ruminant agricultural environments. For example, *L. monocytogenes* has been isolated from the feces of wild birds [57], horses [59, 73], swine [58, 74], poultry [59] and other domestic animals [59] as well as from eviscerated farmed fish [75]. While *L. monocytogenes* in ruminants and on ruminant farms are more likely to contribute directly to human disease, i.e., through human consumption of raw milk [76], the presence of *L. monocytogenes* in other food animals

is more likely to contribute indirectly to food contamination and human disease, e.g., by facilitating introduction of this pathogen into food processing plants or onto vegetables through contaminated manure, e.g., references [39, 77]. One example of food products that appear to become contaminated both directly and indirectly is cold-smoked fish products. Production of cold-smoked fish products does not include a listeriocidal heat treatment. Matching *L. monocytogenes* subtypes have been found occasionally in both raw fish and cold smoked products produced from the contaminated raw materials, supporting a direct route of transmission [78], however, in most cases *L. monocytogenes* contamination of these products appears to occur from the processing plant environment rather than from the raw material [79, 80].

In addition to animal-associated agricultural environments, *L. monocytogenes* has also been isolated from a number of plant-based agricultural systems [38, 39, 81]. Reported prevalences of *L. monocytogenes* in raw vegetables have ranged from 1.1% to 85.7% with an average prevalence of 11.4% [81]. Raw vegetables have been implicated as sources in multiple human listeriosis outbreaks [82-84]. Vegetables may become increasingly important in human listeriosis transmission since current trends in food consumption patterns reflect increasing consumption of raw and ready-to-eat vegetables.

***Listeria monocytogenes* growth and survival in agricultural environments.** In addition to field studies that indicate survival and persistence of L. monocytogenes in agricultural environments over time periods up to 6 years [4], a number of laboratory studies have also provided evidence that Listeria spp. survive in animal feces [85] and agricultural soil [4], including manure amended soil [86], for prolonged time periods. For example, Listeria spp. have been shown to survive in bovine feces from 182 to 2190 days [4]; for several weeks in soil [86, 87]; and for >56 days [4] in sewage sludge cake sprayed onto fields.

Subtype analysis of *L. monocytogenes* found in agricultural environments. Serotyping data provided initial evidence that *L. monocytogenes* serotypes associated with human disease (i.e., serotypes 1/2a, 1/2b, and 4b) are present in agricultural environments [88, 89] ruminant fecal samples [64, 90], and raw milk collected on farms [35]. Molecular subtyping studies further support the presence of human disease-associated *L. monocytogenes* strains in agricultural environments and on farms. For example, in one study, 23% of *L. monocytogenes* isolates from human sporadic cases were found to have identical or similar PFGE patterns to *L. monocytogenes* isolates from farm environments [91]. Nightingale et al. [36] found that 25 of the 35 *Eco*RI ribotypes with greater than 5 occurrences among farm environments and ruminant fecal samples had also been isolated previously from human listeriosis cases [54]. This study also found that all three ribotypes linked to multiple human listeriosis outbreaks, which represent *L. monocytogenes* epidemic clones, were each found on multiple farms. Similarly, isolates with a PFGE type identical to the strain responsible for the 1985 listeriosis outbreak in Los Angeles [92] have also been isolated from a dairy farm by Borucki et al. [91]. Subtyping data thus clearly support that *L. monocytogenes* subtypes linked to sporadic human listeriosis cases and to human listeriosis outbreaks are commonly found in agricultural environments and on farms.

LISTERIA MONOCYTOGENES IN FOOD-ASSOCIATED ENVIRONMENTS AND FOODS

Since the vast majority of human listeriosis cases are foodborne, a body of knowledge on the presence of *L. monocytogenes* in food-associated environments and foods has been published. Since a comprehensive review of these studies is beyond the scope of this chapter, only a brief summary is provided here. Detailed information

on food-associated environments and outbreaks of listeriosis can be obtained from a number of review articles [2, 93] and book chapters (e.g., [76]). *L. monocytogenes*, including subtypes associated with human listeriosis cases and outbreaks, are not uncommon in natural, non-agricultural, and agricultural environments, including raw food commodities. Therefore, an understanding of *L. monocytogenes* ecology and survival in food-associated environments is critical for elucidating the transmission of human listeriosis and reducing human infections.

***Listeria monocytogenes* prevalence and loads in foods and food-associated environments.** A relatively small number of human listeriosis cases have been associated with food products contaminated during primary production in agricultural environments, such as raw vegetables, raw milk and raw milk dairy products, since heating regimes typically applied during food cooking or commercial processing effectively inactivate *L. monocytogenes*. Instead, most human listeriosis cases are caused by consumption of Ready-to-Eat (RTE) meats and dairy food products that contain *L. monocytogenes* as a consequence of contamination from environmental sources in processing plants and other food-associated environments (e.g., retail food businesses) [93, 94] and that are not thoroughly re-heated immediately prior to consumption. Foods at particular risk for transmission of this pathogen include RTE foods that are not aseptically packaged after processing and that require refrigerated storage. *L. monocytogenes* prevalences in these products have been reported at 5% and above [1, 95]. Since the infectious dose of *L. monocytogenes* appears to be high (10^6) [96] and initial contamination of RTE foods typically occurs at low levels, RTE foods with the highest likelihood of transmitting listeriosis are those that support the growth of this pathogen and that are stored long enough to allow bacterial numbers to increase to high levels. Foods that have been implicated as sources of human infections include deli meats [97], coleslaw [98] cheeses [99], hotdogs [100], and

smoked fish [101]. Recent *L. monocytogenes* risk assessments [102, 103] provide comprehensive information on food products most commonly linked to human listeriosis cases as well as *L. monocytogenes* contamination prevalences and levels in these foods. In the U.S., the foods that appear to be most commonly responsible for human listeriosis cases include RTE deli meats (average *L. monocytogenes* prevalence of 1.9%) followed by hot dogs (4.8%) [28]. *L. monocytogenes* is prevalent in smoked seafood (12.0%), fruit (11.8%), preserved fish (9.8%), raw seafood (7%) and pate/meat spreads (6.5%) [103]. In general, RTE foods that are handled extensively after heat treatment, such as deli salads that are prepared in retail environments, show the highest *L. monocytogenes* prevalences. Foods produced under poor hygienic conditions also can have very high *L. monocytogenes* prevalences. For example, Van Coiellie et al. [104] found that prepared minced meat collected in Belgian markets showed *L. monocytogenes* prevalences of 94.7%. In addition, specific food processing plants may manufacture products with extremely high prevalences of *L. monocytogenes*. For example, Fonnesbech Vogel et al., [105] reported that cold smoked fish from one processing plant in Denmark had a *L. monocytogenes* prevalence of 85% during one year of their study.

Due to the importance of the food processing environment as a source of post-processing contamination of RTE foods with *L. monocytogenes*, a variety of studies have evaluated the prevalence, transmission, and ecology of *L. monocytogenes* in food processing plants, including the presence of this pathogen on worker gloves and aprons, as well as in free standing water, aerosolized dust particles [106], walls [107], floors and drains [80, 107, 108], and on processing equipment [78-80, 109-113]. Reported *L. monocytogenes* prevalences in food processing environments have varied tremendously among studies. For example, some studies haven shown *L. monocytogenes* prevalences in food processing plant drains to be close to 100% [5]

while others have found very low prevalences in drains (<1%; [109]). *L. monocytogenes* prevalences in processing plant environments are greatly affected by the characteristics of a given processing plant (e.g., sanitary practices, age and design of the facility, processing run times [80, 111]) and also by the types of samples tested and the time of sample collection (e.g., sampling of processing equipment prior to production initiation should yield lower prevalences than sampling of drains at the end of a production day). Further, *L. monocytogenes* prevalences and contamination patterns in processing plant environments can vary considerably from week to week and even daily [112].

Since the mid 1990s, use of molecular subtyping techniques for characterizing *L. monocytogenes* has enhanced our understanding of the ecology and transmission of *L. monocytogenes* in food processing plants. Most importantly, subtyping studies [111, 114] have shown that one or more specific *L. monocytogenes* subtypes can persist in a given processing plant from a few months to up to 10 years [80, 111]. Persistent *L. monocytogenes* contamination has been identified in a variety of food processing environments, including in plants that produce milk [115], cheese [116], RTE meat [117], pork [118] and poultry [118, 119] products, RTE crawfish [120], and smoked seafoods [79]. Persistent isolates in plants often appear to be "plant-specific". For example, different plants in close geographical proximity can host distinct persistent strains [80], suggesting that specific strains can establish themselves as resident microflora in a given processing facility. Importantly, subtyping studies have also shown that subtypes persisting in a given processing environment are often also isolated from finished products, indicating that *L. monocytogenes* persistence in processing plants can be a major factor contributing to finished product contamination [111, 114], particularly if the persistent strain is also associated with food contact surfaces.

While *L. monocytogenes* transmission and persistence in food processing plants has been examined in numerous studies, the primary sources responsible for introduction of *L. monocytogenes* into the processing environment have not been clearly identified. Potential sources include contaminated raw materials, employee's shoes and attire as well as equipment introduced into the plant, including equipment tires [121]. Fecal shedding by healthy human carriers has also been proposed as a potential source for introduction of *L. monocytogenes* into food processing plants. Early studies in high-risk populations (e.g., household contacts of listeriosis patients [122]) suggested the potential for a high prevalence of fecal shedders among humans. However, more recent studies on broader human populations indicate that the prevalence of *L. monocytogenes* shedding is generally low (< 0.12%, [123] and <0.17% [124]) and usually of short duration [125], suggesting that human fecal shedding is likely to play a limited role as a source of *L. monocytogenes* introduction into food processing environments [126].

The majority of reports on *L. monocytogenes* ecology and transmission have focused on food processing environments, while less is known about *L. monocytogenes* prevalence and transmission in other food related environments, including the retail environment and consumer homes. A study by Gombas et al. [127] found that luncheon meats, deli salads, and seafood salads packaged at retail were 6.8, 2.6, and 5 times more likely to be contaminated with *L. monocytogenes* as compared to manufacturer-packaged equivalents, suggesting a considerable risk for retail *L. monocytogenes* contamination of at least some types of RTE foods. Further, a molecular subtyping study by Sauders et al. [128] showed that a number of different *L. monocytogenes* strains appear to persist in different retail environments, consistent with the well-established ability of this pathogen to establish persistent contamination in food processing environments. Not surprisingly, *L. monocytogenes* has also been

isolated from domestic household environments, including food preparation-associated surfaces such as kitchen sinks, dish-cloths, and washing-up brushes [129], indicating consumer households as potential contamination sources. The contributions of points after primary processing, including in retail and consumer or commercial kitchens, to *L. monocytogenes* contamination of RTE foods, and hence, to human disease incidence, remains to be elucidated. If time intervals are short between contamination and consumption of food products in food service operations and homes, it is less likely that *L. monocytogenes* will be able grow to levels associated with human disease, particularly if initial contamination levels are low.

***Listeria monocytogenes* growth and survival in foods and food-associated environments.** While *L. monocytogenes* can survive and grow under a variety of environmental conditions, including high salt concentrations and low pH, one of the most important characteristics contributing to human exposure is its ability to grow at refrigeration temperatures [130], which can enable it to multiply to high numbers in RTE foods that support its growth. Numerous studies have been conducted to characterize the ability of *L. monocytogenes* to grow in different foods held under different temperatures. *L. monocytogenes* has been found to grow, albeit slowly, even at refrigeration temperatures close to 0°C, with increasing growth rates as storage temperature increases. Many RTE foods that inherently permit growth of *L. monocytogenes* can allow growth of this pathogen to high numbers (e.g., up to 2.5 $x10^6$ CFU g^{-1} in corned beef and up to 1.8 x 10^7 CFU g^{-1} in ham [131]). A comprehensive review and summary of *L. monocytogenes* growth and survival characteristics in a variety of foods can be found in "Microorganisms in Foods 5" [132].

Subtype analysis of *L. monocytogenes* found in foods and food-associated environments. While serotyping studies indicated that a number of isolates from

RTE foods represented serotypes associated with human disease (i.e., serotypes 1/2a, 1/2b, and 4b) [76], a number of food isolates also represented serotypes rarely associated with human infections, in particular, serotype 1/2c [1]. The observations that serotypes 1/2b and 4b were often overrepresented among human clinical isolates as compared to their prevalence among food-associated isolates, while serotypes 1/2a, and particularly 1/2c, were generally overrepresented among food isolates as compared to their prevalence among human clinical isolates, provided initial evidence that *L. monocytogenes* strains and serotypes differ in their abilities to cause human disease. Molecular subtyping studies further supported these observations by showing that *L. monocytogenes* strains grouped into lineage I (which includes serotypes 1/2b and 4b) were more common among human clinical isolates as compared to food isolates, while strains grouped into lineage II (which includes serotypes 1/2a and 1/2c) were more common among food isolates than human clinical isolates [30, 33]. A study of almost 1,000 *L. monocytogenes* isolates from human clinical cases and foods showed that a number of specific ribotypes within the different *L. monocytogenes* lineages differed significantly in their prevalences among food and human isolates [25]. Three ribotypes that were overrepresented among isolates from human listeriosis patients represented subtypes previously associated with multiple human listeriosis cases (i.e., epidemic clones), and also showed significantly higher ability to spread from cell-to-cell in a tissue culture plaque assay, providing phenotypic data supporting enhanced mammalian virulence of these subtypes and epidemic clones [25]. Conversely, a number of the specific subtypes that were more common among food isolates in both the US and France also showed reduced invasion efficiencies for human intestinal epithelial Caco-2 cells. Interestingly, the reduced invasion phenotype was found to be caused by one of several possible mutations leading to premature stop codons in *inlA*, which encodes for internalin A, a listerial surface

molecule critical for invasion of human intestinal epithelial cells, thus providing a clear biological explanation for attenuated human virulence in these strains. Importantly, strains with premature stop codon mutations in *inlA* appear to represent about 30% of food isolates as established by independent studies in France [133] and the US [25, 26] supporting that a number of food isolates show reduced virulence. In addition, a smaller proportion of food isolates appears to show attenuated human virulence due to mutations in other virulence genes, including *hly* and *prfA* [134, 135]. The combination of molecular subtyping with phenotypic characterization has thus provided substantial evidence for virulence differences among *L. monocytogenes* subtypes. These experimental observations have also been supported by mathematical modeling data that indicate greater than 5 log differences in the likelihood of different *L. monocytogenes* subtypes to cause human disease [34].

Subtyping studies on *L. monocytogenes* isolated from food processing and retail environments showed that plant- or retail-specific *L. monocytogenes* subtypes can persist in these environments, and also showed that the subtypes found in these environments represent both subtypes rarely associated with human disease as well as those commonly associated with human disease, including subtypes that have caused multiple human listeriosis outbreaks (i.e., epidemic clones [128]). The observation that human disease-associated *L. monocytogenes* can persist in the environment without apparent loss of human virulence is further supported by the observation that a human listeriosis outbreak in 2000 was linked to contaminated RTE deli turkey produced in a processing plant in which the causative strain appears to have persisted for more than 10 years [136]. The strain responsible for the outbreak in 2000 also caused a single human listeriosis case in 1989 that was linked to consumption of contaminated hot dogs produced (and contaminated) in the same plant [136].

TRANSMISSION INTO THE MAMALIAN HOST

Transmission of *L. monocytogenes* into mammalian hosts and development of infection is dependent on various host factors, pathogen-related factors, environmental factors as well as interactions among these factors [3]. Critical factors for transmission of *L. monocytogenes* include (i) presence of sufficiently high numbers of the pathogen in food or feed, (ii) presence of *L. monocytogenes* strains of sufficient virulence in the food, as well as (iii) exposure of a sufficiently susceptible mammalian host. The interplay among these factors is also critical for development of listeriosis. For example, even a highly virulent *L. monocytogenes* strain present at high levels in a food is unlikely to cause disease if the exposed host is young and highly immunocompetent. At the other extreme, even a virulence attenuated *L. monocytogenes* strain (e.g., a strain with a premature stop codon in *inlA*) can cause human disease, even if present in foods at fairly low levels, if the exposed host is severely immunocompromised. Transmission of *L. monocytogenes* appears to differ between human and non-human mammalian hosts (specifically in different ruminants), therefore, key aspects of *L. monocytogenes* transmission into these host populations is discussed in the following sections.

Transmission of *L. monocytogenes* to non-human mammals. While *L. monocytogenes* can cause disease in various non-primate mammalian hosts, this section will focus predominantly on transmission in ruminants (e.g., cattle), since very little is known about natural transmission in other non-primate mammalian hosts. This selected focus is not intended to imply that other mammalian hosts do not have important roles in the overall ecology and transmission of *L. monocytogenes*.

Transmission of *L. monocytogenes* in silage-fed ruminant hosts and the ecology of *L. monocytogenes* in ruminant hosts and ruminant farm environments (specifically, those feeding silage) appears to be characterized by a high prevalence of

this pathogen in the environment [36], including high levels (up to 100%) of fecal shedding among cattle as well as potentially high *L. monocytogenes* loads in silage (up to $1x10^8$ CFU g^{-1} silage [20]). Preliminary analyses suggest intrahost amplification of ingested *L. monocytogenes* in cattle that show fecal shedding [36], even though fecal shedding generally appears to be short; further studies, including mathematical modeling of transmission, are needed to further confirm this hypothesis. The data available to date suggest that silage fed ruminants and the associated farm environment can maintain high *L. monocytogenes* densities, most likely due to a combination of multiple factors, including *L. monocytogenes* growth in poorly fermented silage as well as fecal shedding by animals. *L. monocytogenes* that are fecally shed are dispersed into the environment, e.g., onto plant material that may be used for subsequent silage production, thus effectively maintaining an infectious cycle. Interestingly, clinical disease in ruminant populations that are fed silage year-round appears to be uncommon and is generally limited to a single or few animals in a herd if disease occurs. In these ruminant populations, disease cases often appear to be linked to either consumption of silage contaminated with extremely high levels of *L. monocytogenes* or immunosuppression of cows or both [3, 61]. The lack of frequent signs of overt disease despite the presence of high pathogen numbers in the environment may possibly reflect herd immunity against *L. monocytogenes* due to frequent exposure to the pathogen through consumption of contaminated silage. In the silage fed large ruminant ecosystem, the host and pathogen may have established an equilibrium that allows high pathogen population densities with limited animal disease. The importance of constant or at least frequent *L. monocytogenes* exposure, and consequent immunity, is highlighted by the fact that small ruminant populations (e.g., sheep and goats) that are fed silage only seasonally (i.e., in the winter) show higher prevalences of listeriosis with more severe disease outcomes following

exposure to contaminated silage, possibly due to reduced immunity after extended time periods without silage feeding and thus without the *L. monocytogenes* exposure needed to maintain or build anti-listerial immunity.

Interestingly, preliminary data indicate that the majority of farm environment and ruminant-associated *L. monocytogenes* isolates examined to date are fully able to invade human intestinal epithelial cells (i.e., they do not carry *inlA* premature stop codon mutations that are responsible for virulence attenuation in a proportion of human food-associated *L. monocytogenes* strains [26]). It is thus tempting to speculate that silage fed ruminants and the associated agricultural environments represent an important, but unlikely sole, reservoir for human virulent *L. monocytogenes* strains. On-farm sources appear to be rarely linked directly to food contamination and human disease, as most RTE foods appear to be contaminated in the processing plant environment rather than from the farm environment, therefore, the importance of ruminant farms and agricultural environments as direct or indirect sources of human *L. monocytogenes* infections remain to be elucidated.

Transmission of *L. monocytogenes* into the human host. Transmission of *L. monocytogenes* into the human host is almost exclusively foodborne [6] and Ready-to-Eat foods (i.e., foods that do not undergo an additional cooking step before consumption) that allow growth of *L. monocytogenes* during storage are most commonly implicated as vehicles of human infections. As described in section 5.1, specific food categories most commonly associated with human listeriosis cases include ready-to-eat deli meats and hot dogs, deli salads, soft cheeses, raw milk and raw milk dairy products, paté, smoked seafoods, and vegetables [76, 103].

Unlike many other foodborne pathogens, such as *Salmonella* and enterohemorrhagic *E. coli*, human infections with *L. monocytogenes* usually require a high pathogen dose [96, 103]. In addition, human hosts that present clinical symptoms

after foodborne exposure to *L. monocytogenes* usually are fetuses or severely immunocompromised individuals. Since human foodborne exposure to *L. monocytogenes*, even at high doses, is not uncommon and usually does not result in human disease, it is tempting to speculate that the majority of the human population has some immunity against this pathogen. Considering the rarity of human infections as well as the apparent short duration and low prevalence of human fecal shedding, it appears that *L. monocytogenes* represents an opportunistic human pathogen and that human infections are likely to contribute little if anything to the ecological success or dispersal of *L. monocytogenes*.

OVERALL *L. MONOCYTOGENES* TRANSMISSION AND CONCLUSIONS

L. monocytogenes is a widely distributed, if not ubiquitous, bacterial pathogen. While the importance of feed- and foodborne transmission to ruminant and human hosts has been well defined, its routes of transmission among different ecosystems and compartments within food production systems appear complex and remain to be clearly elucidated. Despite the fact that human infections with *L. monocytogenes* appear rare, particularly given the frequent prevalence and occasional high load of *L. monocytogenes* in many different environments, including in human foods and animal feeds, it is tempting to propose an anthropocentric transmission pathway for *L. monocytogenes* from the general environment through animal populations to food processing environments and foods to humans. While subtyping studies have clearly shown that human disease-associated *L. monocytogenes* strains, including epidemic clones, can be found in many environments, including natural, urban and farm environments, directionality of transfer and transmission is difficult to establish. Consequently, future work remains to identify and characterize *L. monocytogenes* hosts, reservoirs, and transmission pathways, with consideration given to the

possibility that the true natural host(s) of *L. monocytogenes* could be currently unidentified mammalian or even non-mammalian species. The ecological success of *L. monocytogenes* as a globally distributed microorganism may lie in its ability to survive in a large number of hosts as well as in non-host associated environments, with the ability to establish high population densities in some host associated ecosystems.

A number of distinct *L. monocytogenes* phylogenetic lineages and clonal groups have been identified and classified based on differences in abilities to cause human disease. Key groups important to the overall picture of *L. monocytogenes* ecology and transmission include (i) virulence attenuated strains (such as those characterized by premature stop codons in *inlA* [26] or by mutations in other virulence genes [135]), (ii) epidemic clones, which appear to show increased human virulence as compared to other strains, and (iii) lineage III strains that appear to be associated with animal hosts [33] and that have limited ability to survive or multiply in non-host associated environments [25, 137]. Evolution of *L. monocytogenes* strains and lineages likely represents adaptation of specific strains to different niches, including many that may remain to be defined (e.g., in alternate host species). An improved understanding of the evolution of different *L. monocytogenes* ecotypes will thus provide an opportunity to better understand the ecology and transmission of *L. monocytogenes*, including its reservoirs and hosts.

REFERENCES

1. Farber JM, Peterkin PI: ***Listeria monocytogenes*, a food-borne pathogen.** *Microbiol Rev* 1991, **55:**476-511.
2. Kathariou S: ***Listeria monocytogenes* virulence and pathogenicity, a food safety perspective.** *Journal of food protection* 2002, **65:**1811-1829.
3. Roberts AJ, Wiedmann M: **Pathogen, host and environmental factors contributing to the pathogenesis of listeriosis.** *Cell Mol Life Sci* 2003, **60:**904-918.
4. Fenlon DR: ***Listeria monocytogenes* in the natural environment. In: Ryser ET and Marth EH (eds) *Listeria*, listeriosis, and food safety. 2nd edn. M Decker Inc, New York,.** 1999**:**pp 21-37.
5. Gravani R: **Incidence and control of *Listeria* in food-processing facilities. In: Ryser ET and Marth EH (eds) *Listeria*, listeriosis, and food safety. 2nd edn. M Decker Inc, New York.** 1999**:**pp 657-700.
6. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV: **Food-related illness and death in the United States.** *Emerg Infect Dis* 1999, **5:**607-625.
7. Kreft J, Vazquez-Boland JA, Ng E, Goebel W: **Virulence gene clusters and putative pathogenicity islands in Listeriae. In: Kaper JB, Hacker J (eds) Pathogenicity islands and other mobile virulence elements. 1st edn. ASM Press, Washington, D.C.** 1999**:**pp 219-232.
8. Wiedmann M: **Molecular subtyping methods for *Listeria monocytogenes*.** *J AOAC Int* 2002, **85:**524-531.
9. Wiedmann M: **Subtyping technologies for bacterial foodborne pathogens.** *Nutr Rev* 2002, **60:**201-208.

10. Windham K, Nightingale K, Wiedmann M: **Molecular evolution and diversity of foodborne pathogens. In: Shetty K, Pometto A, Paliyath G (eds) Food biotechnology, CRC Press, Boca Raton,.** 2005**:**pp 1259-1291.

11. Sauders BD, and Wiedmann, M: **Ecology of *Listeria* species and *L. monocytogenes* in the natural environment. In: Ryser ET and Marth EH (eds) *Listeria,* listeriosis, and food safety. 3rd edn. M Decker Inc, New York,.** submitted.

12. MacDonald F, Sutherland AD: **Important differences between the generation times of *Listeria monocytogenes* and *List. innocua* in two *Listeria* enrichment broths.** *J Dairy Res* 1994, **61:**433-436.

13. Donnelly CW: **Detection and isolation of *Listeria monocytogenes* from food samples: implications of sublethal injury.** *J AOAC Int* 2002, **85:**495-500.

14. Ryser ET, Arimi SM, Bunduki MM, Donnelly CW: **Recovery of different *Listeria* ribotypes from naturally contaminated, raw refrigerated meat and poultry products with two primary enrichment media.** *Appl Environ Microbiol* 1996, **62:**1781-1787.

15. Pritchard TJ, Flanders KJ, Donnelly CW: **Comparison of the incidence of *Listeria* on equipment versus environmental sites within dairy processing plants.** *Int J Food Microbiol* 1995, **26:**375-384.

16. Ryser ET, Arimi SM, Donnelly CW: **Effects of pH on distribution of *Listeria* ribotypes in corn, hay, and grass silage.** *Appl Environ Microbiol* 1997, **63:**3695-3697.

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

18. Yu LS, Fung DY: **Five-tube most-probable-number method using the Fung-Yu tube for enumeration of *Listeria monocytogenes* in restructured meat products during refrigerated storage.** *Int J Food Microbiol* 1993, **18:**97-106.

19. Wiedmann M, Arvik T, Bruce JL, Neubauer J, del Piero F, Smith MC, Hurley J, Mohammed HO, Batt CA: **Investigation of a listeriosis epizootic in sheep in New York state.** *Am J Vet Res* 1997, **58:**733-737.

20. Wiedmann M, Bruce JL, Knorr R, Bodis M, Cole EM, McDowell CI, McDonough PL, Batt CA: **Ribotype diversity of *Listeria monocytogenes* strains associated with outbreaks of listeriosis in ruminants.** *J Clin Microbiol* 1996, **34:**1086-1090.

21. Vela AI, Fernandez-Garayzabal JF, Vazquez JA, Latre MV, Blanco MM, Moreno MA, de La Fuente L, Marco J, Franco C, Cepeda A, et al: **Molecular typing by pulsed-field gel electrophoresis of Spanish animal and human Listeria monocytogenes isolates.** *Appl Environ Microbiol* 2001, **67:**5840-5843.

22. Lozniewski A, Humbert A, Corsaro D, Schwartzbrod J, Weber M, Le Faou A: **Comparison of sludge and clinical isolates of Listeria monocytogenes.** *Lett Appl Microbiol* 2001, **32:**336-339.

23. Cai S, Kabuki DY, Kuaye AY, Cargioli TG, Chung MS, Nielsen R, Wiedmann M: **Rational design of DNA sequence-based strategies for subtyping Listeria monocytogenes.** *J Clin Microbiol* 2002, **40:**3319-3325.

24. Salcedo C, Arreaza L, Alcala B, de la Fuente L, Vazquez JA: **Development of a multilocus sequence typing method for analysis of Listeria monocytogenes clones.** *J Clin Microbiol* 2003, **41:**757-762.

25. Gray MJ, Zadoks RN, Fortes ED, Dogan B, Cai S, Chen Y, Scott VN, Gombas DE, Boor KJ, Wiedmann M: ***Listeria monocytogenes*** **isolates from foods and humans form distinct but overlapping populations.** *Appl Environ Microbiol* 2004, **70:**5833-5841.

26. Nightingale KK, Windham K, Martin KE, Yeung M, Wiedmann M: **Select *Listeria monocytogenes* subtypes commonly found in foods carry distinct nonsense mutations in inlA, leading to expression of truncated and secreted internalin A, and are associated with a reduced invasion phenotype for human intestinal epithelial cells.** *Appl Environ Microbiol* 2005, **71:**8764-8772.

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

28. Piffaretti JC, Kressebuch H, Aeschbacher M, Bille J, Bannerman E, Musser JM, Selander RK, Rocourt J: **Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease.** *Proceedings of the National Academy of Sciences of the United States of America* 1989, **86:**3818-3822.

29. Nadon CA, Woodward DL, Young C, Rodgers FG, Wiedmann M: **Correlations between molecular subtyping and serotyping of *Listeria monocytogenes*.** *J Clin Microbiol* 2001, **39:**2704-2707.

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

31. Ward TJ, Gorski L, Borucki MK, Mandrell RE, Hutchins J, Pupedis K: **Intraspecific phylogeny and lineage group identification based on the prfA virulence gene cluster of *Listeria monocytogenes*.** *Journal of bacteriology* 2004, **186:**4994-5002.

32. Norton DM, Scarlett JM, Horton K, Sue D, Thimothe J, Boor KJ, Wiedmann M: **Characterization and pathogenic potential of *Listeria monocytogenes* isolates from the smoked fish industry.** *Appl Environ Microbiol* 2001, **67:**646-653.

33. Jeffers GT, Bruce JL, McDonough PL, Scarlett J, Boor KJ, Wiedmann M: **Comparative genetic characterization of *Listeria monocytogenes* isolates from human and animal listeriosis cases.** *Microbiology (Reading, England)* 2001, **147:**1095-1104.

34. Chen Y, Ross WH, Gray MJ, Wiedmann M, Whiting RC, Scott VN: **Attributing risk to *Listeria monocytogenes* subgroups: dose response in relation to genetic lineages.** *Journal of food protection* 2006, **69:**335-344.

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

36. Nightingale KK, Schukken YH, Nightingale CR, Fortes ED, Ho AJ, Her Z, Grohn YT, McDonough PL, Wiedmann M: **Ecology and transmission of *Listeria monocytogenes* infecting ruminants and in the farm environment.** *Appl Environ Microbiol* 2004, **70:**4458-4467.

37. Welshimer HJ, Donker-Voet J: ***Listeria monocytogenes* in nature.** *Appl Microbiol* 1971, **21:**516-519.

38. Weis J, Seeliger HP: **Incidence of *Listeria monocytogenes* in nature.** *Appl Microbiol* 1975, **30:**29-32.

39. Fenlon DR, Wilson J, Donachie W: **The incidence and level of *Listeria monocytogenes* contamination of food sources at primary production and initial processing.** *J Appl Bacteriol* 1996, **81:**641-650.

40. Sauders BD: **Molecular epidemiology, diversity, distribution, and ecology of *Listeria*.** *PhD Thesis* 2005.

41. Arvanitidou M, Papa A, Constantinidis TC, Danielides V, Katsouyannopoulos V: **The occurrence of *Listeria* spp. and *Salmonella* spp. in surface waters.** *Microbiol Res* 1997, **152:**395-397.

42. Frances N, Hornby H, Hunter PR: **The isolation of *Listeria* species from fresh-water sites in Cheshire and North Wales.** *Epidemiol Infect* 1991, **107:**235-238.

43. Colburn KG, Kaysner CA, Abeyta CJ, Wekell MM: ***Listeria* species in a California coast estuarine environment.** *Appl Environ Microbiol* 1990, **56:**2007-2011.

44. al-Ghazali MR, al-Azawi SK: **Storage effects of sewage sludge cake on the survival of *Listeria monocytogenes*.** *J Appl Bacteriol* 1988, **65:**209-213.

45. al-Ghazali MR, al-Azawi SK: **Effects of sewage treatment on the removal of *Listeria monocytogenes*.** *J Appl Bacteriol* 1988, **65:**203-208.

46. Garrec N, Picard-Bonnaud F, Pourcher AM: **Occurrence of *Listeria* sp. and *L. monocytogenes* in sewage sludge used for land application: effect of dewatering, liming and storage in tank on survival of *Listeria* species.** *FEMS Immunol Med Microbiol* 2003, **35:**275-283.

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

48. De Luca G ZF, Fateh-Moghadm P, Stampi S: **Occurrence of *Listeria monocytogenes* in sewage sludge.** *Zentralbl Hyg Umweltmed* 1998, **201:**269-277.

49. MacGowan AP, Bowker K, McLauchlin J, Bennett PM, Reeves DS: **The occurrence and seasonal changes in the isolation of *Listeria* spp. in shop bought food stuffs, human faeces, sewage and soil from urban sources.** *Int J Food Microbiol* 1994, **21:**325-334.

50. Geuenich HH, Muller HE, Schretten-Brunner A, Seeliger HP: **The occurrence of different Listeria species in municipal waste water.** *Zentralbl Bakteriol Mikrobiol Hyg [B]* 1985, **181:**563-565.

51. Welshimer HJ: **Survival of Listeria monocytogenes in soil.** *J Bacteriol* 1960, **80:**316-320.

52. Botzler RG, Cowan AB, Wetzler TF: **Survival of *Listeria monocytogenes* in soil and water.** *J Wildl Dis* 1974, **10:**204-212.

53. McLauchlin J: **Distribution of serovars of Listeria monocytogenes isolated from different categories of patients with listeriosis.** *Eur J Clin Microbiol Infect Dis* 1990, **9:**210-213.

54. Sauders BD, Durak MZ, Fortes E, Windham K, Schukken Y, Lembo AJ, Jr, Akey B, Nightingale KK, Wiedmann M: **Molecular characterization of *Listeria monocytogenes* from natural and urban environments.** *Journal of food protection* 2006, **69:**93-105.

55. Sauders BD, Fortes ED, Morse DL, Dumas N, Kiehlbauch JA, Schukken Y, Hibbs JR, Wiedmann M: **Molecular subtyping to detect human listeriosis clusters.** *Emerg Infect Dis* 2003, **9:**672-680.

56. Murray EGD, Webb RA, Swann MBR: **A disease of rabbits characterized by a large mononuclear leucocytosis, caused by a hitherto undescribed**

bacillus *Bacterium monocytogenes* (n.sp.). *J Pathol Bacteriol* 1926, **29:**407-439.

57. Fenlon DR: **Wild birds and silage as reservoirs of *Listeria* in the agricultural environment.** *J Appl Bacteriol* 1985, **59:**537-543.

58. Hayashidani H, Kanzaki N, Kaneko Y, Okatani A, Taniguchi T, Kaneko K, Ogawa M: **Occurrence of yersiniosis and listeriosis in wild boars in Japan.** *J Wildl Dis* 2002, **38:**202-205.

59. Weber A, Potel J, Schafer-Schmidt R, Prell A, Datzmann C: **Studies on the occurrence of *Listeria monocytogenes* in fecal samples of domestic and companion animals.** *Zentralbl Hyg Umweltmed* 1995, **198:**117-123.

60. Welshimer HJ: **Isolation of *Listeria monocytogenes* from vegetation.** *Journal of bacteriology* 1968, **95:**300-303.

61. Wesley IV: **Listeriosis in Animals. In: Ryser ET and Marth EH (eds) *Listeria,* listeriosis, and food safety. 2nd edn. M Decker Inc, New York,.** 1999**:**pp 39-73.

62. Sammarco ML, Ripabelli G, Fanelli I, Grasso GM: **Prevalence of *Listeria* spp. in dairy farm and evaluation of antibiotic-resistance of isolates.** *Ann Ig* 2005, **17:**175-183.

63. Unnerstad H, Romell A, Ericsson H, Danielsson-Tham ML, Tham W: **Listeria monocytogenes in faeces from clinically healthy dairy cows in Sweden.** *Acta Vet Scand* 2000, **41:**167-171.

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

65. Husu JR: **Epidemiological studies on the occurrence of Listeria monocytogenes in the feces of dairy cattle.** *Zentralbl Veterinarmed B* 1990, **37:**276-282.

66. Gronstol H: **Listeriosis in sheep. Listeria monocytogenes excretion and immunological state in healthy sheep.** *Acta Vet Scand* 1979, **20:**168-179.

67. Loken T, Aspoy E, Gronstol H: **Listeria monocytogenes excretion and humoral immunity in goats in a herd with outbreaks of listeriosis and in a healthy herd.** *Acta Vet Scand* 1982, **23:**392-399.

68. Gronstol H, Overas J: **Listeriosis in sheep. Eperythrozoon ovis infection used as a model to study predisposing factors.** *Acta Vet Scand* 1980, **21:**523-532.

69. Nightingale KK, Fortes ED, Ho AJ, Schukken YH, Grohn YT, Wiedmann M: **Evaluation of farm management practices as risk factors for clinical listeriosis and fecal shedding of *Listeria monocytogenes* in ruminants.** *J Am Vet Med Assoc* 2005, **227:**1808-1814.

70. Sanaa M, Poutrel B, Menard JL, Serieys F: **Risk factors associated with contamination of raw milk by *Listeria monocytogenes* in dairy farms.** *J Dairy Sci* 1993, **76:**2891-2898.

71. Fenlon DR: **Rapid quantitative assessment of the distribution of *Listeria* in silage implicated in a suspected outbreak of listeriosis in calves.** *Vet Rec* 1986, **118:**240-242.

72. Wiedmann M, Czajka J, Bsat N, Bodis M, Smith MC, Divers TJ, Batt CA: **Diagnosis and epidemiological association of *Listeria monocytogenes* strains in two outbreaks of listerial encephalitis in small ruminants.** *J Clin Microbiol* 1994, **32:**991-996.

73. Gudmundsdottir K, Svansson V, Aalbaek B, Gunnarsson E, Sigurdarson S: **Listeria monocytogenes in horses in Iceland.** *Vet Rec* 2004, **155:**456-459.

74. Yokoyama E, Saitoh T, Maruyama S, Katsube Y: **The marked increase of *Listeria monocytogenes* isolation from contents of swine cecum.** *Comp Immunol Microbiol Infect Dis* 2005, **28:**259-268.

75. Miettinen H, Wirtanen G: **Prevalence and location of *Listeria monocytogenes* in farmed rainbow trout.** *Int J Food Microbiol* 2005, **104:**135-143.

76. Ryser E: **Foodborne listeriosis. In: Ryser ET and Marth EH (eds) *Listeria,* listeriosis, and food safety. 2nd edn. M Decker Inc, New York,.** 1999**:**pp 299-358.

77. Rorvik LM, Aase B, Alvestad T, Caugant DA: **Molecular epidemiological survey of *Listeria monocytogenes* in broilers and poultry products.** *Journal of applied microbiology* 2003, **94:**633-640.

78. Markkula A, Autio T, Lunden J, Korkeala H: **Raw and processed fish show identical *Listeria monocytogenes* genotypes with pulsed-field gel electrophoresis.** *Journal of food protection* 2005, **68:**1228-1231.

79. Hoffman AD, Gall KL, Norton DM, Wiedmann M: ***Listeria monocytogenes* contamination patterns for the smoked fish processing environment and for raw fish.** *Journal of food protection* 2003, **66:**52-60.

80. Norton DM, McCamey MA, Gall KL, Scarlett JM, Boor KJ, Wiedmann M: **Molecular studies on the ecology of *Listeria monocytogenes* in the smoked fish processing industry.** *Appl Environ Microbiol* 2001, **67:**198-205.

81. Beuchat LR: ***Listeria monocytogenes* incidence on vegetables.** *Food Control* 1996, **7:**223-228.

82. Ho JL, Shands KN, Friedland G, Eckind P, Fraser DW: **An outbreak of type 4b *Listeria monocytogenes* infection involving patients from eight Boston hospitals.** *Arch Intern Med* 1986, **146:**520-524.

83. Allerberger F, Guggenbichler JP: **Listeriosis in Austria--report of an outbreak in 1986.** *Acta Microbiol Hung* 1989, **36:**149-152.

84. Schlech III WF, Lavigne PM, Bortolussi RA, Allen AC, Haldane EV, Wort AJ, Hightower AW, Johnson SE, King SH, Nicholls ES, Broome CV: **Epidemic listeriosis--evidence for transmission by food.** *N Engl J Med* 1983, **308:**203-206.

85. Dijkstra RG: **Investigations on the survival times of Listeria bacteria in suspensions of brain tissue, silage and faeces and in milk.** *Zentralbl Bakteriol [Orig]* 1971, **216:**92-95.

86. Jiang X, Islam M, Morgan J, Doyle MP: **Fate of *Listeria monocytogenes* in bovine manure-amended soil.** *Journal of food protection* 2004, **67:**1676-1681.

87. Nicholson FA, Groves SJ, Chambers BJ: **Pathogen survival during livestock manure storage and following land application.** *Bioresour Technol* 2005, **96:**135-143.

88. Dijkstra RG: **Incidence of Listeria monocytogenes in the intestinal contents of broilers on different farms.** *Tijdschr Diergeneeskd* 1978, **103:**229-231.

89. Borucki MK, Gay CC, Reynolds J, McElwain KL, Kim SH, Call DR, Knowles DP: **Genetic diversity of *Listeria monocytogenes* strains from a high-prevalence dairy farm.** *Appl Environ Microbiol* 2005, **71:**5893-5899.

90. Ralovich B, Audurier A, Hajtos I, Berkessy E, Pitron-Szemeredi M: **Comparison of *Listeria* serotypes and phage types isolated from sheep, other animals and humans.** *Acta Microbiol Hung* 1986, **33:**9-17.

91. Borucki MK, Reynolds J, Gay CC, McElwain KL, Kim SH, Knowles DP, Hu J: **Dairy farm reservoir of *Listeria monocytogenes* sporadic and epidemic strains.** *Journal of food protection* 2004, **67:**2496-2499.

92. Linnan M, Mascola L, Lou X, Goulet V, May S, Salminen C, Hird D, Yonkura M, Hayes PS, Weaver RE, et al: **Epidemic listeriosis associated with Mexican-style cheese.** *N Engl J Med* 1988, **319:**823-828.

93. Tompkin RB: **Control of *Listeria monocytogenes* in the food-processing environment.** *Journal of food protection* 2002, **65:**709-725.

94. Reij MW, Den Aantrekker ED: **Recontamination as a source of pathogens in processed foods.** *Int J Food Microbiol* 2004, **91:**1-11.

95. Farber JM, Peterkin PI: **Incidence and Behavior of *Listeria monocytogenes* in Meat Products In: Ryser ET and Marth EH (eds) *Listeria,* listeriosis, and food safety. 2nd edn. M Decker Inc, New York,.** 1999**:**pp 505-564.

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

97. Gottlieb SL, Newbern EC, Griffin PM, Graves LM, Hoekstra RM, Baker NL, Hunter SB, Holt KG, Ramsey F, Head M, et al: **Multistate outbreak of listeriosis linked to turkey deli meat and subsequent changes in US regulatory policy.** *Clin Infect Dis* 2006, **42:**29-36.

98. Ryser E: **Listeriosis in animals. In: Ryser ET and Marth EH (eds) *Listeria,* listeriosis, and food safety. 2nd edn. M Decker Inc, New York,.** 1999**:**pp 229-358.

99. Goulet V, Jacquet, C., Vaillant, V., Rebiere, I., Mouret, E., Lorente, C., Maillot, E., Stainer, F., Rocourt, J.: **Listeriosis from consumption of raw-milk cheese.** *Lancet* 1995, **345:**1581-1582.

100. CDC: **Update: multistate outbreak of listeriosis -- United States.** *MMWR* 1999, **47:**1117-1118.

101. Ericsson H, Eklow A, Danielsson-Tham M, Loncarevic S, Mentzing L, Persson I, Unnerstad H, Tham W: **An outbreak of listeriosis suspected to have been caused by rainbow trout.** *J Clin Microbiol* 1997, **35:**2904-2907.

102. FAO/WHO: **Microbiological risk assessment series 5: risk assessment of *Listeria monocytogenes* in ready-to-eat foods. Available at: http://www.fao.org/documents.** 2004.

103. FDA/USDA: **Quanitative assessment of relative risk to public health from foodborne *Listeria monocytogenes* among selected categories of ready-to-eat foods.** *Available at: http://wwwfoodsafetygov/~dms/LMr2-tochtml* 2003.

104. Van Coillie E, Werbrouck H, Heyndrickx M, Herman L, Rijpens N: **Prevalence and typing of *Listeria monocytogenes* in ready-to-eat food products on the Belgian market.** *Journal of food protection* 2004, **67:**2480-2487.

105. Fonnesbech Vogel B, Huss HH, Ojeniyi B, Ahrens P, Gram L: **Elucidation of *Listeria monocytogenes* contamination routes in cold-smoked salmon processing plants detected by DNA-based typing methods.** *Appl Environ Microbiol* 2001, **67:**2586-2595.

106. De Roin MA, Foong SC, Dixon PM, Dickson JS: **Survival and recovery of *Listeria monocytogenes* on ready-to-eat meats inoculated with a desiccated and nutritionally depleted dustlike vector.** *Journal of food protection* 2003, **66:**962-969.

107. Chasseignaux E, Toquin MT, Ragimbeau C, Salvat G, Colin P, Ermel G: **Molecular epidemiology of *Listeria monocytogenes* isolates collected from the environment, raw meat and raw products in two poultry- and pork-processing plants.** *Journal of applied microbiology* 2001, **91:**888-899.

108. Rorvik LM, Skjerve E, Knudsen BR, Yndestad M: **Risk factors for contamination of smoked salmon with *Listeria monocytogenes* during processing.** *Int J Food Microbiol* 1997, **37:**215-219.

109. Autio T, Hielm S, Miettinen M, Sjoberg AM, Aarnisalo K, Bjorkroth J, Mattila-Sandholm T, Korkeala H: **Sources of *Listeria monocytogenes* contamination in a cold-smoked rainbow trout processing plant detected by pulsed-field gel electrophoresis typing.** *Appl Environ Microbiol* 1999, **65:**150-155.

110. Rorvik LM, Caugant DA, Yndestad M: **Contamination pattern of *Listeria monocytogenes* and other Listeria spp. in a salmon slaughterhouse and smoked salmon processing plant.** *Int J Food Microbiol* 1995, **25:**19-27.

111. Thimothe J, Nightingale KK, Gall K, Scott VN, Wiedmann M: **Tracking of *Listeria monocytogenes* in smoked fish processing plants.** *Journal of food protection* 2004, **67:**328-341.

112. Hu Y, Gall, K, Ho, A, Ivanek, R, Grohn, YT, Wiedmann, M: **Daily variability of *Listeria* contamination patterns in a cold-smoked salmon processing operations.** *Journal of food protection* submitted.

113. Eklund M, Poysky F, Paranjpye R, Lashbrook L, Peterson M, Pelroy G: **Incidence and sources of *Listeria monocytogenes* in cold smoking fishery products and processing plants.** *Journal of food protection* 1995, **58:**502-508.

114. Lappi VR, Thimothe J, Nightingale KK, Gall K, Scott VN, Wiedmann M: **Longitudinal studies on *Listeria* in smoked fish plants: impact of intervention strategies on contamination patterns.** *Journal of food protection* 2004, **67:**2500-2514.

115. Kells J, Gilmour A: **Incidence of *Listeria monocytogenes* in two milk processing environments, and assessment of *Listeria monocytogenes* blood agar for isolation.** *Int J Food Microbiol* 2004, **91:**167-174.

116. Kabuki DY, Kuaye AY, Wiedmann M, Boor KJ: **Molecular subtyping and tracking of *Listeria monocytogenes* in latin-style fresh-cheese processing plants.** *J Dairy Sci* 2004, **87:**2803-2812.

117. Samelis J, Metaxopoulous J: **Incidence and principal sources of *Listeria* spp. and *L. monocytogenes* contamination in processed meats and a meat processing plant.** *Food Microbiology* 1999, **16:**465-477.

118. Chasseignaux E, Gerault P, Toquin MT, Salvat G, Colin P, Ermel G: **Ecology of *Listeria monocytogenes* in the environment of raw poultry meat and raw pork meat processing plants.** *FEMS Microbiol Lett* 2002, **210:**271-275.

119. Lawrence LM, Gilmour A: **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* 1995, **61:**2139-2144.

120. Lappi VR, Thimothe J, Walker J, Bell J, Gall K, Moody MW, Wiedmann M: **Impact of intervention strategies on *Listeria* contamination patterns in crawfish processing plants: a longitudinal study.** *Journal of food protection* 2004, **67:**1163-1169.

121. Rocourt J, Cossart P: ***Listeria monocytogenes*. In: Doyle MP, Beuchat LR, and Montville TJ (eds) Food microbiology, fundamentals and frontiers. 1st edn. ASM Press, Washington, D.C.** 1997**:**pp 337-352.

122. Schuchat A, Deaver K, Hayes PS, Graves L, Mascola L, Wenger JD: **Gastrointestinal carriage of *Listeria monocytogenes* in household contacts of patients with listeriosis.** *J Infect Dis* 1993, **167:**1261-1262.

123. Sauders BD, Pettit D, Currie B, Suits P, Evans A, Stellrecht K, Dryja DM, Slate D, Wiedmann M: **Low prevalence of *Listeria monocytogenes* in human stool.** *Journal of food protection* 2005, **68:**178-181.

124. Schlech III WF, Schlech IV WF, Haldane H, Mailman TL, Warhuus M, Crouse N, Haldane DJM: **Does sporadic *Listeria* gastroenteritis exist? A 2-year population-based survey in Nova Scotia, Canada.** *Clin Infect Dis* 2005, **41:**778-784.

125. Grif K, Patscheider G, Dierich MP, Allerberger F: **Incidence of fecal carriage of *Listeria monocytogenes* in three healthy volunteers: a one-year prospective stool survey.** *Eur J Clin Microbiol Infect Dis* 2003, **22:**16-20.

126. Ivanek R, Grohn, YT, and Wiedmann, M: **Modeling *Listeria monocytogenes* in multiple habitats and host populations: review of available data.** *Journal of food protection* submitted.

127. Gombas DE, Chen Y, Clavero RS, Scott VN: **Survey of *Listeria monocytogenes* in ready-to-eat foods.** *Journal of food protection* 2003, **66:**559-569.

128. Sauders BD, Mangione K, Vincent C, Schermerhorn J, Farchione CM, Dumas NB, Bopp D, Kornstein L, Fortes ED, Windham K, Wiedmann M: **Distribution of *Listeria monocytogenes* molecular subtypes among human and food isolates from New York State shows persistence of human**

disease--associated *Listeria monocytogenes* strains in retail environments. *Journal of food protection* 2004, **67:**1417-1428.

129. Beumer RR, te Giffel MC, Spoorenberg E, Rombouts FM: ***Listeria* species in domestic environments.** *Epidemiol Infect* 1996, **117:**437-442.

130. Gray ML, Killinger AH: ***Listeria monocytogenes* and listeric infections.** *Bacteriol Rev* 1966, **30:**309-382.

131. Sim J, Hood D, Finnie L, Wilson M, Graham C, Brett M, Hudson JA: **Series of incidents of *Listeria monocytogenes* non-invasive febrile gastroenteritis involving ready-to-eat meats.** *Lett Appl Microbiol* 2002, **35:**409-413.

132. ICMSF: **Listeria monocytogenes. In: Roberts TA, Baird-Parker, AC, and Tompkin, RB (eds) Microorganisms in foods 5. 1st edn. Blackie Academic & Professional, London.** 1996**:**pp 141-182.

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

134. Roche SM, Gracieux P, Milohanic E, Albert I, Virlogeux-Payant I, Temoin S, Grepinet O, Kerouanton A, Jacquet C, Cossart P, Velge P: **Investigation of specific substitutions in virulence genes characterizing phenotypic groups of low-virulence field strains of *Listeria monocytogenes*.** *Appl Environ Microbiol* 2005, **71:**6039-6048.

135. Roberts A, Chan Y, Wiedmann M: **Definition of genetically distinct attenuation mechanisms in naturally virulence-attenuated *Listeria monocytogenes* by comparative cell culture and molecular characterization.** *Appl Environ Microbiol* 2005, **71:**3900-3910.

136. Olsen SJ, Patrick M, Hunter SB, Reddy V, Kornstein L, MacKenzie WR, Lane K, Bidol S, Stoltman GA, Frye DM, et al: **Multistate outbreak of *Listeria***

***monocytogenes* infection linked to delicatessen turkey meat.** *Clin Infect Dis* 2005, **40:**962-967.

137. Roberts A, Nightingale K, Jeffers G, Fortes E, Kongo JM, Wiedmann M: **Genetic and phenotypic characterization of *Listeria monocytogenes* lineage III.** *Microbiology (Reading, England)* 2006, **152:**685-693.

CHAPTER TWO

σ^B Contributes to Stress Response and Virulence in Select *Listeria monocytogenes* Strains Representing Lineages I, II, IIIA, and IIIB

ABSTRACT

Phylogenetic analysis of *L. monocytogenes* strains has identified three distinct lineages. While lineages I and II are both common among human clinical and food isolates, lineage I strains are overrepresented among human clinical isolates, and lineage II strains are overrepresented among food and environmental isolates. Lineage III, which includes subgroups IIIA and IIIB, is rare and predominantly associated with animal disease. σ^B, encoded by *sigB*, is a sigma factor previously demonstrated to critically contribute to stress response and virulence in lineage II strains. We used transcriptomic and phenotypic analyses to characterize the role of σ^B in *L. monocytogenes* strains representing lineages I, II, IIIA, and IIIB. Whole-genome expression microarrays were used to compare stationary phase wildtype and *sigB* null mutant transcriptomes. Mutant and wildtype strains were tested for acid and oxidative stress survival, ability to invade Caco-2 human intestinal epithelial cells, and for differences in virulence in the guinea pig gastrointestinal model of infection. A total of 63 genes were positively regulated by sigma B in all four strains evaluated. In addition, a number of genes were found to be positively regulated by σ^B only in specific strains, including genes apparently positively regulated by σ^B in the lineage I strain but not in the lineage II representative. While σ^B contributed significantly to acid and oxidative stress survival and Caco-2 cell invasion in lineage I, II, and IIIB strains, σ^B contributions to stress survival and invasion in the lineage IIIA strain were

not significant under the conditions tested. Further, σ^B contributed to virulence in the guinea pig animal model for listeriosis in each lineage representative. Our results indicate that the role of σ^B in stress response and virulence differs among *L. monocytogenes* strains, which may contribute to the differences in the distribution of *L. monocytogenes* lineages among different sources (e.g., humans and foods).

INTRODUCTION

Listeria monocytogenes is a Gram-positive, rod-shaped, saprophytic organism found ubiquitously in nature (i.e., soil, water, plants, manure), and it is therefore not possible to completely exclude this pathogen from the food chain. However, *L. monocytogenes* is the etiological agent of listeriosis, a life-threatening invasive disease in humans and animals. While listeriosis rarely occurs in healthy individuals, the elderly, the immunocompromised, and pregnant women and their fetuses are particularly at risk. An overwhelming majority of listeriosis cases and *L. monocytogenes* infections are foodborne (99%) and an estimated 20% of cases result in 500 human deaths each year in the U.S [1].

L. monocytogenes' ability to grow in a wide range of temperatures (0 to 45°C) [2], pH (4.4 to 9.4) [2], and under other stressful environmental conditions, makes it particularly difficult to control in food processing environments, allowing it to subsequently contaminate foods. It is well established that the general stress regulator, σ^B, contributes to *L. monocytogenes'* ability to overcome a number of environmental stresses, including acid, oxidative, and energy stresses as shown by reduced survival of an isogenic Δ*sigB* mutant under the same conditions [3-5]. Homologues of σ^B have been identified in many Gram-positive bacteria such as *Staphylococcus aureus* [6], *B. anthracis* [7], and *B. licheniformis* [8]. σ^S (RpoS) is the stationary phase stress response alternative sigma factor in Gram-negative including *Escherichia coli* [9], and

Salmonella Typhimurium [10]; its roles vary among species, but include survival of environmental stress [11] and virulence gene regulation. In *L. monocytogenes,* σ^B contributes to transcription of a number of virulence genes (i.e. *prfA*, *bsh*, *inlA*, and *inlB*) [12-17] . Consequently, a Δ*sigB* null mutant has reduced invasiveness in human intestinal epithelial cells [18] and reduced virulence in intragastrically inoculated guinea pigs [19] . Furthermore, σ^B directly regulates more than 55 genes [12] and a more recent study indicates that it may regulate, directly and indirectly, as many as 200 genes [16].

Molecular subtyping methods, including ribotyping [20, 21], pulsed-field gel electrophoresis (PFGE) [22, 23], and more recently, multilocus sequence-based typing (MLST) [24, 25], have provided recent advances in our understanding of *L. monocytogenes* transmission and have been used in many studies on *L. monocytogenes* ecology. Characterization of *L. monocytogenes* isolates from a variety of different hosts and environments by multiple subtyping methods, including initial multilocus enzyme electrophoresis (MLEE) work by Pifaretti et al. [26], has shown that strains comprising the species *L. monocytogenes* represent at least three distinct genetic lineages. While different nomenclatures have been used to designate these *L. monocytogenes* lineages [27], the main lineages described in different studies are a consistent grouping of specific *L. monocytogenes* serotypes [26, 28]. *L. monocytogenes* strains group into two major divisions, designated lineages I and II, and a third, distinct division that can be further sub-divided into lineages IIIA and IIIB, as shown by application of a number of molecular subtyping strategies [20, 26], including PFGE [29], and virulence gene sequencing [20, 30]. Based on the lineage designations used by most groups [20, 30, 31], lineage I predominantly includes serotypes 1/2b, 3b, 3c, and 4b strains and lineage II primarily includes serotypes 1/2a, 1/2c, and 3a [28]. Interestingly, previous reports have shown that lineage I strains are

significantly overrepresented among human clinical listeriosis cases as compared to their prevalence among animal listeriosis cases and contaminated foods [28, 30, 32]. On the other hand, lineage II strains show a significantly higher prevalence among food isolates than among human listeriosis cases [30]. In addition, lineage I isolates appear to have significantly greater pathogenic potential, as determined by their ability to spread to neighboring host cells in a cell culture plaque assay, when compared to lineage II isolates [30, 32]. Lineage III predominantly includes serotypes 4a and 4c, as well as some serotype 4b strains that are distinct from those grouped into lineage I [28]. Strains classified in lineage III appear to be associated with isolation from animals and are occasionally isolated from human listeriosis cases with clinical disease, but are rarely isolated from foods [30, 33]. Increasing evidence exists that *L. monocytogenes* strains represent multiple lineages that appear to differ in their abilities to be transmitted to humans, as also supported by recent subtype-specific mathematical modeling data, which indicate that the likelihood of human disease caused by *L. monocytogenes* classified into different lineages can differ by more than 2 logs [34].

While the contributions of σ^B to stress response and virulence is well characterized in lineage II strains (e.g. 10403S [12-16], EGDe [35]), there is limited study of the contributions of σ^B to lineages I, IIIA, and IIIB. Moorhead et al. concluded that the relative importance of σ^B in stress response is not the same in all strains of *L. monocytogenes*, as assessed by comparing the survival of two Δ*sigB* mutants from different serotypes under various stress conditions [36]. Specifically, serotype 1/2a was more dependent on σ^B under stress conditions compared to serotype 4a. Furthermore, σ^B contribution to virulence in the guinea pig model has been conducted with laboratory strain 10403S and isogenic Δ*sigB* which are lineage II strains, therefore the relative contribution of σ^B with respect to other lineages in

gastrointestinal infections has not yet been explored. A recent comparative transcriptomic study found that *sigB* (lmo0895) was overexpressed in lineage II strains when compared to lineage I strains as were a number of genes under the control of σ^B [37] identified previously [12, 16] including *opuCA* and *lmo1421* which encode known and putative compatible solute transporter proteins. We postulated that the inherent differences between the lineages' ability to cause disease may be partially attributable to differences in contributions from σ^B. To expand our understanding of σ^B and its role in stress response and virulence, we have evaluated its role in four strains selected to represent lineages I, II, IIIA, and IIIB, utilizing a combination of transcriptomic and phenotypic analyses. We used a *L. monocytogenes* multi-genome microarray to initially characterize differences in the σ^B regulons of *L. monocytogenes* lineage representatives. The isogenic parent and Δ*sigB* strain pairs were assayed for acid and oxidative stress survival, invasiveness in a Caco-2 cell model, and virulence in the guinea pig gastrointestinal model of infection to evaluate differences in σ^B contributions to stress response and virulence amongst diverse strains.

MATERIALS AND METHODS

Bacterial strains and storage. Four strains were selected to represent the inherent diversity of *L. monocytogenes* species. FSL J1-194 was selected as a lineage I strain from a sporadic case; it represents serotype 1/2b which is commonly associated with human disease [27], and is Ribotype DUP-1042B, which may be a representative of a major epidemic clone with enhanced virulence characteristics [30]. 10403S is a common laboratory type strain [38]. It was selected as a lineage II representative and as an internal control in order that the results of this study could be compared to previous studies (e.g. [4, 19, 39]). FSL J2-071 and FSL J1-208 represent lineages IIIA and IIIB, respectively. Both strains are animal clinical cases and are serotype 4c and

4a, respectively, which are serotypes common to lineage III [28]. FSL A1-254, a 10403S isogenic Δ*sigB* mutant, was used in previous studies (e.g. [3-5]). Isogenic *sigB* null mutants (Δ*sigB*) were created in each lineage representative using allelic exchange mutagenesis, described previously [3, 40]; all strains used in this study are summarized in Table 2.1. Primers used for mutant generation are described in Table A2 [S2.1]. Stock cultures were stored at -80°C in BHI containing 15% glycerol. Cultures were streaked onto brain heart infusion agar (BHI; Difco, Detroit, MI) and incubated at 37°C for 24 h to obtain isolated colonies for inoculation of overnight cultures. Detailed growth conditions are described for each experiment below.

Microarray, cDNA labeling and microarray hybridization.

Growth conditions and RNA isolation. Bacteria were grown in 5 ml of BHI broth at 37ºC with shaking (230 rpm) for 15 h. A 1% inoculum was transferred to a fresh 5 ml tube of pre-warmed BHI and grown to OD_{600} = 0.4, then a 1% inoculum of the OD_{600}=0.4 culture was transferred to a 300 ml nephelo flask containing 50 ml of pre-warmed BHI. Following the second passage, cells were collected at stationary phase (defined as growth to an OD_{600} = 1.0, followed by incubation for an additional 3 h). Prior to centrifugation, RNAProtect bacterial reagent (Qiagen, Valencia, CA) was added to the cultures according to manufacturer's instructions to stabilize the mRNA; pellets were stored at -80°C prior to RNA isolation. RNA was isolated as previously described [16]. Briefly, bacterial cells were lysed enzymatically using lysozyme, and mechanically, using 6 sonication bursts at 18W on ice for 30 seconds. Total RNA was isolated and purified using the RNeasy Midi kit (Qiagen) according to the manufacturer's protocol. RNA was eluted from the column using RNase-free water. Total RNA was incubated with RNasin (Promega, Madison, WI) to inhibit RNases, and RQ1 DNase (Promega) to remove DNA contamination. Subsequently, two phenol:chloroform treatments were used to remove the enzymes followed by one

Table 2.1 Strains used in this study

Lineage	Stain	Serotype	Ribotype	Origin	Reference
I	FSL J1-194	1/2b	DUP-1042B	Human clinical isolate	30
I	FSL C6-001, Δ*sigB*	1/2b	DUP-1042B	FSL J1-194	This study
II	10403S	1/2a	DUP-1030A	Laboratory type strain	38
II	FSL A1-254, Δ*sigB*	1/2a	DUP-1030A	10403S	3
IIIA	FSL J2-071	4c	DUP-1061A	Bovine clinical isolate	33
IIIA	FSL O1-006, Δ*sigB*	4c	DUP-1061A	FSL J2-071	This study
IIIB	FSL J1-208	4a	DUP-10142	Caprine clinical isolate	33
IIIB	FSL O1-005, Δ*sigB*	4a	DUP-10142	FSL J1-208	This study

chloroform treatment to remove excess phenol. UV spectrophotometry (Nanodrop, Wilmington, DE) was used to quantify and assess purity of the RNA. RNA integrity was assessed by agarose gel electrophoresis. Purified RNA samples were stored in RNase-free water at -80°C prior to reverse transcription. RNA was precipitated for long-term storage.

The Pathogen Function Genomics Resource Center / Craig Venter Institute (PFGRC/JCVI) *L. monocytogenes microarray.* The PFGRC/JCVI *L. monocytogenes* microarray *version 2* was used to identify mRNA expression differences between wildtype strains and their respective Δ*sigB* mutant. Differences in expression patterns across lineages were compared.

RNA aminoallyl labeling and cDNA synthesis for microarrays. The PFGRC/JCVI microbial RNA aminoallyl labeling for microarray standard operating procedure M007 (publically available at http://pfgrc.jcvi.org/) was used to reverse transcribe and label the total RNA.

Analysis of labeling reaction. UV spectrophotometry (Nanodrop Thermo Fisher, Wilmington, DE) was used to quantify cDNA and total picomoles of Cy3 or Cy5 (Amersham Biosciences, Piscataway, NJ) incorporation. The Cy3- and Cy5-labeled cDNA probes were combined and dried to completion prior to hybridization.

Hybridization of labeled cDNA probes. PFGRC/JCVI standard operating procedure M008 (publically available at http://pfgrc.jcvi.org/) was used to hybridize the Cy-dye labeled cDNA probes to the microarray. Briefly, microarray slides were blocked in a prehybridization buffer supplemented with bovine serum albumen (BSA) (Sigma-Aldrich), washed with deionized water then isopropyl alcohol. Dried Cy3- and Cy5-labeled probes were hydrated with hybridization buffer containing 0.6 mg/ml sheared salmon sperm DNA (Invitrogen), denatured twice at 95ºC for 5 minutes and briefly centrifuged. Probes were hybridized to the array at 42ºC for 16 hours.

Following hybridization, slides were washed per the PFGRC/JCVI M008 protocol and centrifuged dry.

Microarray image processing, replicates and statistical analysis. Image processing and analysis were performed as previously described (17, 18, 19). Data pre-processing and statistical analyses were performed using LIMMA package, available from BioConductor software project for R programming environment [41, 42]. Background correction was performed using "normexp" method to produce more robust ratios for low-intensity spots and print-tip normalization was used to correct for spatial effects and dye-intensity bias [43]. The empirical Bayes approach was used to assess differential expression; empirical Bayes reduced standard errors of the estimated log-fold changes resulting in more stable inference and improved power for experiments with small numbers of arrays [44]. Three replicates were used for each transcriptome comparison. For each probe, fold changes, moderated t-statistics, and p-values (adjusted for multiple testing by controlling for the false discovery rate) were calculated. An χ^2 test for trend was used to determine if there was an overrepresentation of σ^B-dependent genes among The J. Craig Venter Institute (JCVI) role categories. Subsequent χ^2 tests were used to determine which role categories were overrepresented by σ^B-dependent genes.

Hybridization index and identification of differentially expressed genes. The PFGRC/JCVI *L. monocytogenes* microarray had multiple probes designed for some loci if a consensus sequence sufficient for a unique 70mer could not be identified among the four strains represented. In instances where multiple probes were present for a single locus, the probe that best matched the strains used in this study was selected as the representative probe for that strain. BLASTN was used to calculate percent identity between each probe and each strain; blast results were parsed to generate a table containing the percent identity between each strain and the probes.

The percent identity between each strain and each probe on the microarray was *L. monocytogenes* EGD-e locus names were used to describe genes in FSL J1-194 (lineage I), FSL J2-071 (lineage IIIA), and FSL J1-208 (lineage IIIB) as these genomes have not yet been annotated and were also used in lieu of 10403S locus names to facilitate comparisons between all strains in this study. Genes unique to F2365 (i.e. no EGD-e locus name) maintained an F2365 locus name. A list of positively differentially expressed genes was generated; genes with adjusted p-values <0.05 and fold changes $\geq$ 1.5 between the parent and respective isogenic Δ*sigB* mutant strain were considered positively differentially expressed, and thus σ^B-dependent. Raw and normalized microarray data is available at the NCBI Gene Expression Omnibus (GEO).

TaqMan qRT-PCR. TaqMan qRT-PCR was used to confirm lineage-specific σ^B-dependent genes. Total RNA from cells grown as described above for microarray analyses was extracted using the Ambion MicrobExpress kit (Ambion). RNA quality and integrity was assessed by the Agilent 2100 Bioanalzyer (Agilent, Santa Clara, CA); total RNA concentration was checked by spectrophotometry (Nanodrop). RNA was reverse transcribed to cDNA using the High Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Applied Biosystems) using 1 μg of total RNA and random hexamers. qRT-PCR was performed on an ABI Prism 7000 Sequence Detection Systme (Applied Biosystems). To check for residual genomic DNA, the same reaction was performed for each sample without reverse transcriptase (RT); *rpoB* primers and probe were used to calculate C_t values for no RT reactions. The efficiencies of each primer set were calculated from ten-fold serial dilution standard curves for each wildtype strain and the relative RNA transcript levels were calculated by the Pfaffl method [45]. We used *rpoB* transcript levels (C_t) as the reference gene to calculate the relative expression ratio. In previous studies by our group [13, 46], *rpoB*

transcript levels have been used to calculate standard curves for absolute transcript quantification as it is a housekeeping gene which shows minimal variation in transcript levels under varying conditions. Primers (IDT DNA, Coralville, IA) and probes (Applied Biosystems) used in this study were designed with Primer Express (Applied Biosystems) based on the consensus sequence of the gene among the four wildtype strains. Primers and probes are described in Table A2 [S2.2]. In the event complete genomic DNA sequence was not available for a gene and at least 200 nt upstream of the coding region for promoter sequence comparison, primers were designed based on the consensus sequences of other available wildtype strains in this study.

Acid and oxidative stress survival. Strains were grown in 5 ml of BHI broth at 37ºC with shaking (230 rpm) for 12 h. A 1% inoculum was transferred to a fresh 5 ml tube of pre-heated BHI and grown to OD_{600} = 0.4. A 1% inoculum was transferred to a fresh tube containing 5 ml pre-warmed BHI. After growth to an OD_{600} of approximately 0.4, a 1% inoculum (adjusted based on actual OD_{600} reading) was transferred to a 300 ml nephelo flask (Bellco, Vineland, NJ) containing 50 ml of pre-warmed BHI. Following the second passage, strains were grown to stationary phase (defined as 10h post-inoculation). Two 5 ml aliquots of stationary phase cells were transferred to 16 mm autoclaved test tubes. For acid stress, one aliquot was used to determine the amount (µl) of 12N HCl (VBR, Westchester, PA) needed to reduce the pH of the culture to 2.5 as determined by direct measurement with a pH meter (Beckman, Coulter Inc., Fullterton, CA). The second aliquot served as the culture used for experimentation. Upon addition of 12N HCl to yield pH 2.5, the cultures were gently vortexed and immediately returned to 37ºC with shaking. Non-stressed cultures were enumerated by serial dilution and standard plate count. Aliquots of the stressed cultures were removed at 10, 30 and 60 min post-acidification. Bacterial

numbers were quantified by serial dilution and standard plate count. For oxidative stress, cumene hydroperoxide (CHP) was added to 5 ml of 10 h cultures in 16 mm test tube at a final concentration of 13.0 mM CHP in DMSO for 15 minutes at 37°C with shaking. An equal volume of DMSO was added to non-stressed cultures; survival was assessed by serial dilution and standard plate count. At least three independent replicates were performed for each *L. monocytogenes* strain tested under each condition.

Caco-2 cell invasion assay. Caco-2 invasion assays were performed as previously described [47]. Confluent Caco-2 monolayers were inoculated with 10 µl of stationary phase culture (approx. 4.8 x 10^7 cells/well) grown as described for acid and oxidative stress survival. Intracellular *L. monocytogenes* cells were enumerated by spiral-plating (Autoplate 4000, Spiral Biotech, Norwood, MA) of lysed Caco-2 cell suspensions and ten-fold dilutions in PBS on BHI agar plates. Invasion efficiency was reported as the log ratio of *L. monocytogenes* cells recovered / initial inoculum. Three independent invasion assays were performed for each *L. monocytogenes* strain tested.

Statistical analyses of stress and invasion assays. Statistical analyses were performed with the Statistical Analysis Software (SAS) 9.0 (SAS Institute, Inc., Cary, NC). Regression analysis was used to calculate the death rate which was expressed as average log CFU death per hour for each strain. Repeated measures analysis of variance was used to test "time", "strain", and the "time*strain" interaction effects. While the variable "time" was expected to be significant for all strains as cells die over time under these conditions, a significant "time*strain" interaction indicated that death rates were significantly different. Oxidative stress survival was reported as the log CFU death of each strain. For Caco-2 invasion assays, differences in the log ratio of CFU recovered after invasion to CFU of the initial inoculum between each parent strain and its isogenic Δ*sigB* mutant were assessed. Two-sided t-tests were used to test

for differences in oxidative death and invasion efficiencies between wildtype and Δ*sigB* strain representatives. Significance was set at $p<0.05$ for all statistical analyses.

Guinea pig model of listeriosis. The guinea pig animal model of listeriosis was used to assess the contribution of σ^B to virulence in *L. monocytogenes* lineages as previously described by Garner et al. [19].

L. monocytogenes growth conditions and cell collection. Bacteria were grown as described for RNA isolation with the exception that cells were grown to early stationary phase (defined as growth to an $OD_{600} = 0.8$ followed by incubation for an additional 1 h) to be consistent with previous guinea pig infection model work conducted in this laboratory. For infection, aliquots of early stationary phase bacterial cultures were concentrated by centrifugation, resuspended in 1 ml of phosphate-buffered saline (PBS) pH 7.4 containing 15% glycerol. Cells were then frozen and stored in multiple aliquots at –80°C for use in three replicates. An aliquot was thawed and enumerated prior to infection by serial dilution and spiral plating to determine cell viability. On the day of infection, cells were thawed just prior to infection; the concentration was adjusted to 1.0×10^{10} CFU ml^{-1}. Actual inoculum was confirmed by serial dilution and plating immediately after infection.

Animal care and housing condition. Animal protocols (# 2002-0060) were approved by the Institutional Animal Care and Use Committee prior to initiation of the experiments. Male Hartley guinea pigs (Elm Hill, Chelmsford, MA) weighing 348 ± 43g at about 3 weeks of age were housed individually, allowing for collection of each animal's fecal material. Animals were provided with feed and water ad libitum. Cages were changed daily, and animal health and weight were monitored and recorded daily. Animals were acclimated for 5 days prior to infection. Experiments were preformed in triplicate.

Intragastric infection of guinea pigs. Intragastric infections of guinea pigs was performed as described previously [19]. Briefly, feed was withheld 12 h prior to infection. Animals were anesthetized with isoflurane administered via inhalation, using oxygen as a carrier gas [48]. *L. monocytogenes* was inoculated intragastrically by gavage. The stomach pH was neutralized with 1 ml of PBS containing 125 mg calcium carbonate (pH 7.4) prior to inoculation with *L. monocytogenes* (1.0×10^{10} CFU ml^{-1}). An additional 1 ml of PBS was used to flush the catheter.

Guinea pig euthanasia, organ harvest, and L. monocytogenes enumeration. Cages and water containers were changed daily. Each guinea pig was weighed daily and prior to euthanasia by CO_2 72 h after infection. The brain, liver, spleen, mesenteric lymph node, and small intestine were harvested and evaluated individually for *L. monocytogenes* as previously described [19]. Recovered organs were held on ice until processing. All organs were weighed and visually inspected for lesions. A 20 cm portion of the small intestine distal to the cecum was harvested. After the contents were removed, the small intestine segments were rinsed two times in 20 ml PBS, incubated at room temperature for 90 min in 20ml of DMEM containing 100 mg L^{-1} gentamicin to kill extracellular bacteria, followed by three rinses in PBS prior to homogenization. The liver was homogenized with 60 ml of sterile PBS in a small autoclaved blender unit for 30 s; the brain, spleen, mesenteric lymph nodes, and small intestine were homogenized with 30 ml of PBS for 30s. Following homogenization, homogenate was directly spread plated and spiral plated onto BHI media in duplicate. Dilutions were exponentially spiral plated onto BHI agar. All samples were enriched to detect presence/absence of *L. monocytogenes* in all organs: 10 ml homogenate was added to 90 ml *Listeria* enrichment broth (LEB) (Difco, Sparks, MD), incubated at 30°C, then plated (50µl) on Oxford medium (ThermoFisher, Waltham, MA) after 24 hours and 48 hours. Detection limits were established for (i) standard plate count and

(ii) enrichment detection methods. The detection limit for standard plate counts was set as the mean between 1 CFU/g organ plated and the calculated enrichment detection limit. The detection limit for enriched sample was set as the mean of 0 and 1 CFU per average g enriched for each organ (i.e. 0.5 CFU/g of average organ weigh). The absence of *L. monocytogenes* after enrichment indicated a CFU initially below the detection limit; all *L. monocytogenes* negative enriched samples were reported at the enrichment detection limit.

Enumeration of L. monocytogenes from feces. Post-infection, feces were collected from all guinea pigs on a daily basis and processed as described by Garner et al. [19]. 0.5 g of feces were homogenized in 4.5 ml of PBS. Homogenized samples were serially diluted in PBS; 10^{-1} dilution was spread plated on Oxford medium, subsequent dilutions were spiral plated on Oxford Medium and grown 24 h at 30ºC. *L. monocytogenes* colonies were enumerated and confirmed on *L. monocytogenes* plating media (LMPM) (Biosynth, Naperville, IL).

Statistical analysis. Statistical analyses were performed with the Statistical Analysis Software (SAS) 9.0 (SAS Institute, Inc., Cary, NC). The recovery levels (in log CFU/g) of *L. monocytogenes* from organs (i.e., liver, spleen, mesenteric lymph nodes, and small intestine) were used as the main measure of virulence. One-sided t-tests were used to compare *L. monocytogenes* numbers (log CFU/g) recovered from each organ, weight at 72h post infection, and *L. monocytogenes* shed in feces (log CFU/g) from each wildtype and Δ*sigB* isogenic pairs. Significance was set at $p<0.05$.

RESULTS

The core σ^B regulon consists of at least 63 genes. A total of 63 genes were positively differentially expressed (fold change ≥1.5; adj. p-value <0.05) in all lineage representatives (Table 2.2) indicating higher transcript levels in the wildtype strains

Table 2.2 σ^B dependent genes significant in all lineage representatives.

lmo[a]	Common Name[b]	Lineage I fold change[c]	Lineage I adj p-value[d]	Lineage II fold change[c]	Lineage II adj p-value[d]	Lineage IIIA fold change[c]	Lineage IIIA adj p-value[d]	Lineage IIIB fold change[c]	Lineage IIIB adj p-value[d]
lmo0133	conserved hypothetical protein	4.5	0.0003	5.1	<0.0001	9	<0.0001	14.5	<0.0001
lmo0134	acetyltransferase, GNAT family	6.9	0.0001	5.6	<0.0001	14.3	<0.0001	7.7	<0.0001
lmo0169	similar to a glucose uptake protein	5.5	0.0002	3.4	<0.0001	5.1	<0.0001	6.9	<0.0001
lmo0170	conserved hypothetical protein	3.6	0.0006	3	0.0008	5.6	<0.0001	7.4	<0.0001
lmo0210	similar to L-lactate dehydrogenase	4.5	<0.0001	2.6	<0.0001	3.8	0.0001	4.6	0.0058
lmo0211	similar to B. subtilis general stress protein	1.7	0.0335	1.5	0.0001	2.1	0.0021	2.4	<0.0001
lmo0405	phosphate transporter family protein	2.7	0.0410	1.7	0.0106	2.1	0.0053	2.1	0.0036
lmo0433	Internalin A	3.6	0.0004	3.1	<0.0001	4.8	<0.0001	7	<0.0001
lmo0515	conserved hypothetical protein	3.3	0.0116	3.4	0.0002	3.5	<0.0001	5.4	<0.0001
lmo0539	similar to tagatose-1,6-diphosphate aldolase	14.4	<0.0001	7.5	<0.0001	19.5	<0.0001	24.2	<0.0001
lmo0555	similar to di-tripeptide transporter	7.2	0.0002	4.1	<0.0001	7.1	<0.0001	9.3	<0.0001
lmo0593	similar to transport proteins (formate?)	6	0.0032	5.7	<0.0001	9	0.0003	18.2	<0.0001
lmo0596	similar to unknown proteins	14.2	0.0001	22.7	<0.0001	33.8	<0.0001	36	<0.0001
lmo0602	weakly similar to transcription regulator	2.6	0.0023	3.7	<0.0001	2.8	0.0003	2.5	0.0121
lmo0610	similar to internalin proteins, putative peptidoglycan bound protein (LPXTG motif)	1.9	0.0380	3.7	<0.0001	5.5	<0.0001	8.2	<0.0001
lmo0642	putative membrane protein	3.4	0.0028	2	0.0002	2	0.0029	3.1	0.0005

Table 2.2 (Continued)

lmo[a]	Common Name[b]	Lineage I fold change[c]	Lineage I adj p-value[d]	Lineage II fold change[c]	Lineage II adj p-value[d]	Lineage IIIA fold change[c]	Lineage IIIA adj p-value[d]	Lineage IIIB fold change[c]	Lineage IIIB adj p-value[d]
lmo0655	similar to phosphoprotein phosphatases	4.4	<0.0001	2.9	<0.0001	2.5	0.0014	3.1	0.0007
lmo0722	similar to pyruvate oxidase	4.6	0.0002	5.4	<0.0001	8.8	<0.0001	3.8	0.0070
lmo0781	similar to mannose-specific phosphotransferase system (PTS) component IID	10	<0.0001	15.6	<0.0001	18.2	<0.0001	25.4	<0.0001
lmo0782	similar to mannose-specific phosphotransferase system (PTS) component IIC	12.9	<0.0001	13.5	<0.0001	20.3	<0.0001	22	<0.0001
lmo0783	similar to mannose-specific phosphotransferase system (PTS) component IIB	6.4	<0.0001	12	<0.0001	14.8	<0.0001	18	<0.0001
lmo0784	PTS system, IIAB component, authentic frameshift	2	0.0341	5.7	<0.0001	5.5	<0.0001	6.4	<0.0001
lmo0794	similar to B. subtilis YwnB protein	6.1	0.0007	12.8	<0.0001	10.3	<0.0001	25.4	<0.0001
lmo0796	conserved hypothetical protein	1.9	0.0029	4	<0.0001	8.3	<0.0001	12.3	<0.0001
lmo0880	lysM domain protein	7.4	<0.0001	6.7	<0.0001	16.4	<0.0001	6.8	0.0006
lmo0896	Indirect negative regulation of sigma B dependant gene expression (serine phosphatase)	1.5	0.0244	2.3	<0.0001	1.7	0.0197	2.7	0.0017
lmo0911	unknown	2.1	0.0152	2.1	<0.0001	9.2	<0.0001	1.8	0.0200
lmo0913	succinate-semialdehyde dehydrogenase	6.6	<0.0001	13.4	<0.0001	16	<0.0001	22.3	<0.0001
lmo0937	unknown	6.4	<0.0001	10.4	<0.0001	16.2	<0.0001	18.9	<0.0001
lmo0953	unknown	3.2	0.0156	6.5	<0.0001	12	<0.0001	16.3	<0.0001

Table 2.2 (Continued)

lmo[a]	Common Name[b]	Lineage I fold change[c]	Lineage I adj p-value[d]	Lineage II fold change[c]	Lineage II adj p-value[d]	Lineage IIIA fold change[c]	Lineage IIIA adj p-value[d]	Lineage IIIB fold change[c]	Lineage IIIB adj p-value[d]
lmo0956	similar to N-acetylglucosamine-6P-phosphate deacetylase (EC 3.5.1.25)	2.6	0.0073	2.1	<0.0001	4.4	<0.0001	4.3	<0.0001
lmo0957	glucosamine-6-phosphate isomerase	2.3	0.0006	1.6	0.0027	3.1	0.0027	2.4	0.0441
lmo0994	unknown	24.4	<0.0001	14.1	<0.0001	54.2	<0.0001	79.2	<0.0001
lmo1140	unknown	4	0.0084	3.5	<0.0001	4.9	0.0001	4.9	<0.0001
lmo1241	conserved hypothetical protein	2.7	0.0012	2.2	0.0002	4.2	<0.0001	5	<0.0001
lmo1295	similar to host factor-1 protein	1.8	0.0120	3.4	<0.0001	5.4	<0.0001	8	<0.0001
lmo1375	peptidase, M20/M25/M40 family	2.4	0.0168	3	<0.0001	2.4	0.0025	2.1	0.0381
lmo1425	similar to betaine/carnitine/choline ABC transporter (membrane p)	1.9	0.0062	3.4	<0.0001	3.5	<0.0001	4.7	<0.0001
lmo1428	similar to glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)	2.2	0.0008	2.9	<0.0001	2.7	0.0001	3.5	0.0004
lmo1433	pyridine nucleotide-disulfide oxidoreductase family protein	4.7	0.0006	4.2	<0.0001	3	0.0378	3.7	0.0309
lmo1602	similar to unknown proteins	4.2	<0.0001	4.6	<0.0001	1.7	0.0417	2.7	<0.0001
lmo1605	UDP-N-acetylmuramate--alanine ligase	7.4	0.0001	2.2	<0.0001	2.4	0.0047	2.8	<0.0001
lmo1606	FtsK/SpoIIIE family protein	9.1	<0.0001	5.6	<0.0001	3.5	<0.0001	4.1	<0.0001

Table 2.2 (Continued)

lmo[a]	Common Name[b]	Lineage I fold change[c]	Lineage I adj p-value[d]	Lineage II fold change[c]	Lineage II adj p-value[d]	Lineage IIIA fold change[c]	Lineage IIIA adj p-value[d]	Lineage IIIB fold change[c]	Lineage IIIB adj p-value[d]
lmo1694	similar to CDP-abequose synthase	8	<0.0001	8	<0.0001	26.2	<0.0001	1.8	0.0162
lmo1799	putative peptidoglycan bound protein (LPXTG motif)	2.8	0.0017	1.7	0.0002	4.8	<0.0001	4.7	<0.0001
lmo2085	cell wall surface anchor family protein	11	<0.0001	12.2	<0.0001	14.1	0.0007	16.8	0.0067
lmo2130	similar to unknown protein	2.1	0.0112	2.6	<0.0001	1.9	0.0104	2.3	0.0006
lmo2191	similar to unknown proteins	2.2	0.0066	3	<0.0001	2.3	<0.0001	2.3	<0.0001
lmo2269	unknown	4.3	0.0096	5.7	<0.0001	5.6	<0.0001	6.2	<0.0001
lmo2391	conserved hypothetical protein similar to B. subtilis YhfK protein	6.2	<0.0001	9.1	<0.0001	21.6	<0.0001	28.5	<0.0001
lmo2434	glutamate decarboxylase gamma	3.2	0.0193	2.7	<0.0001	4	0.0097	3.8	<0.0001
lmo2454	unknown	3.6	0.0003	4.6	<0.0001	6.6	<0.0001	8.2	<0.0001
lmo2463	similar to transport protein	3	0.0426	3.9	<0.0001	6.2	<0.0001	4.3	0.0266
lmo2485	PspC domain protein, truncated	1.9	0.0113	4.4	<0.0001	1.8	0.0006	2	0.0021
lmo2570	putative membrane protein	5.5	<0.0001	4.5	<0.0001	7.3	<0.0001	8.9	<0.0001
lmo2571	similar to nicotinamidase	4.4	<0.0001	5.8	<0.0001	7.3	<0.0001	8.3	<0.0001
lmo2572	similar to Chain A, Dihydrofolate Reductase	4.4	0.0002	1.6	0.0018	7.5	<0.0001	3.5	<0.0001
lmo2573	alcohol dehydrogenase, zinc-dependent	3.4	0.0002	4.6	<0.0001	5.6	<0.0001	7.5	<0.0001
lmo2673	conserved hypothetical protein	6.9	0.0009	9.7	<0.0001	26	<0.0001	37.2	<0.0001
lmo2674	similar to ribose 5-phosphate epimerase	9.8	<0.0001	3.8	<0.0001	10.1	<0.0001	13.7	<0.0001

Table 2.2 (Continued)

lmo[a]	Common Name[b]	Lineage I fold change[c]	Lineage I adj p-value[d]	Lineage II fold change[c]	Lineage II adj p-value[d]	Lineage IIIA fold change[c]	Lineage IIIA adj p-value[d]	Lineage IIIB fold change[c]	Lineage IIIB adj p-value[d]
lmo2724	similar to unknown proteins	2.8	0.0030	2.9	<0.0001	4.7	<0.0001	4.9	<0.0001
lmo2748	similar to B. subtilis stress protein YdaG	11.6	<0.0001	10	<0.0001	15.3	<0.0001	5.1	0.0063
lmof2365_0703	conserved hypothetical protein	2.2	0.0277	3.3	<0.0001	2.2	0.0038	2.4	0.0186

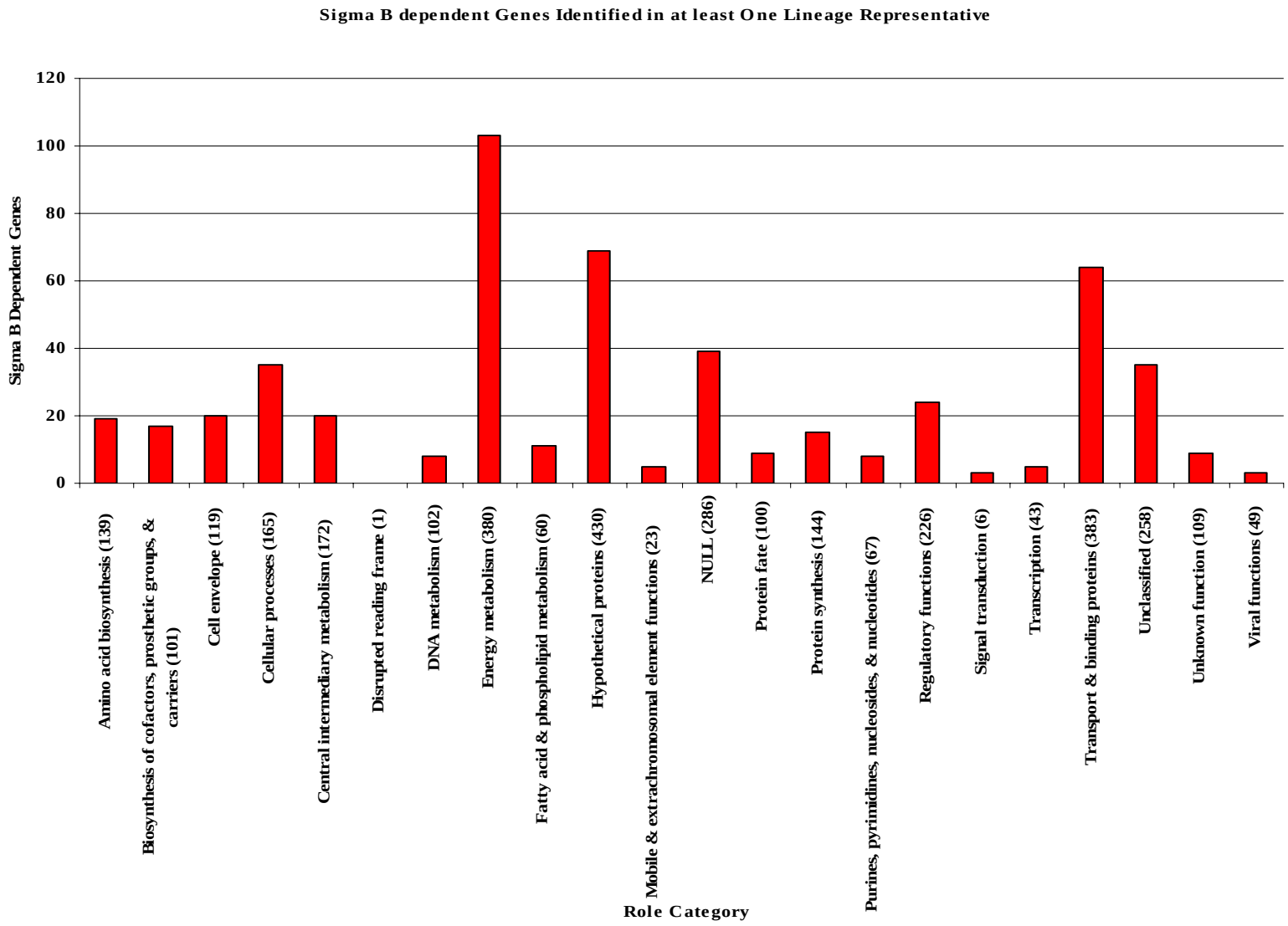
[a] Probe name based on *L. monocytogenes* EGD-e gene

[b] Common name based on EGD-e annotation

[c] Fold changes represent transcript levels in the parent strain compared to the Δ*sigB* strain; significant genes ≥1.5-fold and adjusted p-value <0.05

[d] Adjusted p-value <0.05 considered significant

Figure 2.1 Sigma B dependent genes are overrepresented in 5 *L. monocytogenes* EGDe role categories. 425 genes were identified as positively differentially expressed in at least one strain. The number of sigma B dependent genes in each role category was based on the *L. monocytogenes* EGDe genome. Significant overrepresentation of sigma B dependent genes in a given role category was determined by χ^2 test. Significant categories are indicated by * where $p>0.05$. Numbers in parentheses represent number of genes in each role category in the *L. monocytogenes* EGDe genome.

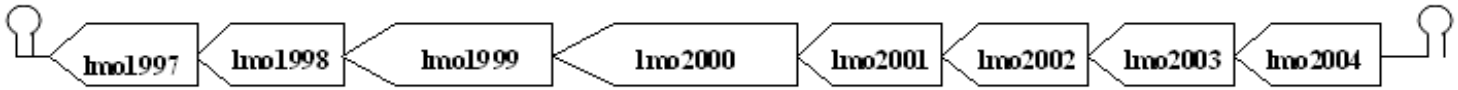

Sigma B dependent Genes Identified in at least One Lineage Representative
Sigma B Dependent Genes
120
100
80
60
40
20
0
Amino acid biosynthesis (139)
Biosynthesis of cofactors, prosthetic groups, & carriers (101)
Cell envelope (119)
Cellular processes (165)
Central intermediary metabolism (172)
Disrupted reading frame (1)
DNA metabolism (102)
Energy metabolism (380)
Fatty acid & phospholipid metabolism (60)
Hypothetical proteins (430)
Mobile & extrachromosomal element functions (23)
NULL (286)
Protein fate (100)
Protein synthesis (144)
Purines, pyrimidines, nucleosides, & nucleotides (67)
Regulatory functions (226)
Signal transduction (6)
Transcription (43)
Transport & binding proteins (383)
Unclassified (258)
Unknown function (109)
Viral functions (49)
Role Category

lmo2593). The GntR family of regulators has been characterized as global regulators of primary metabolism in a number of bacteria [52-54] and MerR-like regulators have been shown to play a role in optimizing σ^{70}-dependent promoters with atypical distances between the -35 and -10 elements [55]. These findings support that *L. monocytogenes* σ^B appears to be involved in a number of transcriptional regulatory networks. Finally, 109 of 425 genes (25.5%) identified as σ^B-dependent in stationary phase cells have an unknown function; efforts to characterize these gene products will likely provide further insight into the role of σ^B in stress response and virulence.

σ^B-dependent genes were identified as exclusively lineage I or lineage II, suggesting differences in σ^B regulons among lineage representatives. In an effort to elucidate potential mechanisms responsible for the differences in *L. monocytogenes* strains' ability to cause disease, we compared the σ^B regulons of lineage I and II representatives. While some genes described hereafter had higher transcripts levels in lineage IIIA and/or lineage IIIB wildtype strains compared to their Δ*sigB* mutant (Figure 2.2, Figure 2.3), we focused our comparison between lineages I and II, as the majority of human listeriosis cases result from infections caused by strains in these lineages [30]. In order for genes to be considered exclusively σ^B-dependent in one lineage and not the other, the corresponding genes not differentially expressed in a lineage had to have ≥ 95% hybridization index; this was necessary to reduce false positives. Of the 170 genes in lineage I and 252 genes in lineage II which had higher transcript levels in the wildtype strain compared to the Δ*sigB* mutant, a total of 106 genes were significant in both lineage representatives. Of the genes determined to have higher transcript levels in the wildtype strain compared to its isogenic Δ*sigB*, 55 were uniquely σ^B-dependent genes in lineage I when compared to the lineage II σ^B regulon under the HI criteria (Table 2.3).

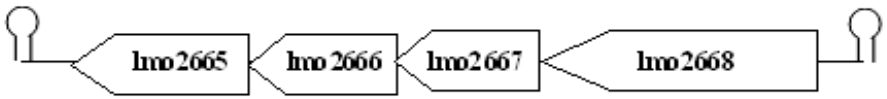
Figure 2.2 σ^B-dependent operons in lineage I but not lineage II. The figures depict three σ^B-dependent putative operons that had higher transcript levels in the lineage I wildtype strain compared to the Δ*sigB* mutant which was not observed in the lineage II representative. A) lmo2004-lmo1997. Genes in this operon were flanked by Rho-independent terminator found by TransTermHP. Numbers below each gene indicate the fold changes representing transcript levels in the parent strain compared to the Δ*sigB* strain for each lineage representative; numbers in bold are significant (fold change ≥1.5; adj. p-value <0.05). Values in parentheses are the hybridization index (%) for each strain; some hybridization indices could not be calculated as BLAST results found no match as these genomes are not yet complete. B) Fold change, adj. p-value and hybridization index for the σ^B-dependent putative operon comprised of lmo2668-lmo2665. C) Fold change, adj. p-value and hybridization index for the σ^B-dependent putative operon comprised of lmo1539-lmo1538.

A



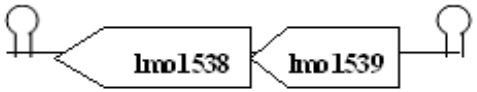
Lin I	3.5 (97.1)	**5.3** (95.7)	4.3 (100)	**13.5** (100)	**10.4** (.)	2.4 (100)	2.3 (95.7)	**3.8** (.)
Lin II	0.8 (100)	0.7 (100)	0.6 (100)	0.9 (98.6)	0.7 (100)	0.8 (95.7)	0.9 (95.7)	0.9 (98.6)
Lin IIIA	**3.5** (94.2)	**4.6** (.)	3.9 (100)	**5.0** (100)	**5.0** (100)	**7.4** (98.6)	**5.1** (95.7)	**4.3** (100)
Lin IIIB	**3.3** (.)	**5.2** (.)	**4.8** (100)	**3.9** (98.6)	**13.1** (.)	**3.4** (95.7)	**4.4** (94.3)	**3.9** (.)

B



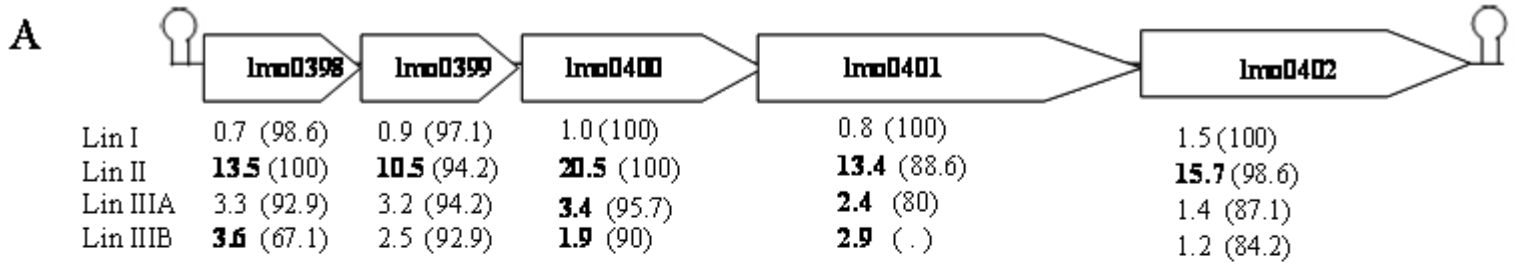
Lin I	**2.9** (95.7)	**2.9** (94.3)	**2.3** (94.3)	**2.2** (100)
Lin II	1 (100)	0.7 (100)	0.8 (98.6)	0.7 (97.1)
Lin IIIA	1.1 (95.7)	1.2 (94.3)	1.1 (97.1)	5.0 (97.1)
Lin IIIB	**1.8** (97.1)	**1.9** (94.3)	1.7 (.)	**2.0** (97.1)

C



Lin I	**2.9** (100)	**3.9** (100)
Lin II	0.7 (100)	0.7 (100)
Lin IIIA	**2.4** (97.1)	**3.8** (100)
Lin IIIB	**4.1** (92.9)	**5.6** (97.1)

Figure 2.3 σ^B-dependent operons in lineage II but not lineage I. The figures depict two σ^B-dependent putative operons that had higher transcript levels in the lineage II wildtype strain compared to the Δ*sigB* mutant which was not observed in the lineage I representative. A) lmo0398-lmo0402. Genes in this operon were flanked by Rho-independent terminator found by TransTermHP. Numbers below each gene indicate the fold changes representing transcript levels in the parent strain compared to the Δ*sigB* strain for each lineage representative; numbers in bold are significant (fold change ≥1.5; adj. p-value <0.05). Values in parentheses are the hybridization index (%) for each strain; some hybridization indices could not be calculated as BLAST results found no match as these genomes are not yet complete. B) Fold change, adj. p-value and hybridization index for the σ^B-dependent putative operon comprised of lmo2164-lmo2158.


A
lmo0398
lmo0399
lmo0400
lmo0401
lmo0402
Lin I 0.7 (98.6) 0.9 (97.1) 1.0 (100) 0.8 (100) 1.5 (100)
Lin II 13.5 (100) 10.5 (94.2) 20.5 (100) 13.4 (88.6) 15.7 (98.6)
Lin IIIA 3.3 (92.9) 3.2 (94.2) 3.4 (95.7) 2.4 (80) 1.4 (87.1)
Lin IIIB 3.6 (67.1) 2.5 (92.9) 1.9 (90) 2.9 (.) 1.2 (84.2)

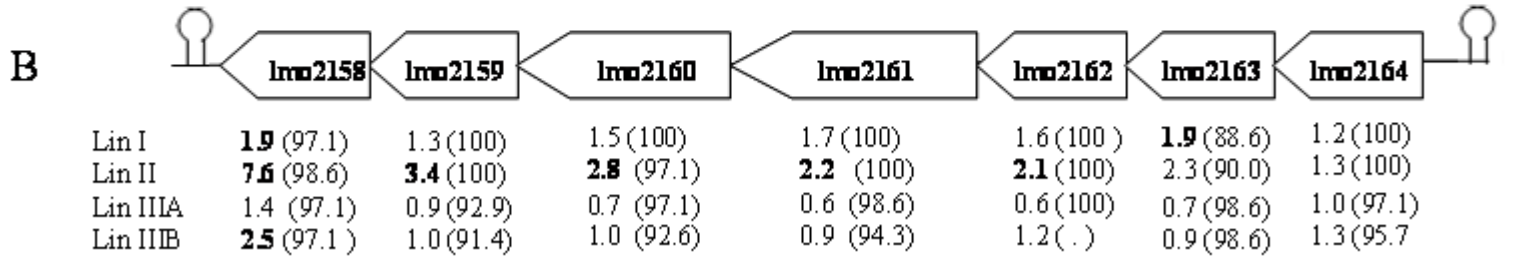

B
lmo2158
lmo2159
lmo2160
lmo2161
lmo2162
lmo2163
lmo2164
Lin I 1.9 (97.1) 1.3 (100) 1.5 (100) 1.7 (100) 1.6 (100) 1.9 (88.6) 1.2 (100)
Lin II 7.6 (98.6) 3.4 (100) 2.8 (97.1) 2.2 (100) 2.1 (100) 2.3 (90.0) 1.3 (100)
Lin IIIA 1.4 (97.1) 0.9 (92.9) 0.7 (97.1) 0.6 (98.6) 0.6 (100) 0.7 (98.6) 1.0 (97.1)
Lin IIIB 2.5 (97.1) 1.0 (91.4) 1.0 (92.6) 0.9 (94.3) 1.2 (.) 0.9 (98.6) 1.3 (95.7

Table 2.3 σ^B-dependent genes significant in lineage I and not in lineage II

lmo[a]	Common Name[bc]	Lineage I fold change[d]	Lineage II fold change[d]	Lineage I adj p-value[e]	Lineage II adj p-value[e]
lmo0130	Ser/Thr protein phosphatase family protein	2.1	1.3	0.0046	0.0764
lmo0188	dimethyladenosine transferase (16S rRNA dimethylase)	1.7	1.1	0.0348	0.3556
lmo0217	similar to B. subtilis DivIC protein	1.8	1.2	0.0167	0.0230
lmo0239	hypothetical protein	1.9	1.1	0.0244	0.4018
lmo0315	similar to thiamin biosynthesis protein	2.5	1.0	0.0395	0.9741
lmo0640	similar to oxidoreductase	1.8	1.2	0.0081	0.0748
lmo0958	transcriptional regulator, GntR family	1.7	1.4	0.0082	0.0100
lmo0959	llm protein	1.8	1.1	0.0398	0.3736
lmo1076	N-acetylmuramoyl-L-alanine amidase, family 4	1.6	1.2	0.0264	0.2002
lmo1237	similar to glutamate racemase	1.7	1.0	0.0324	0.8658
lmo1255	PTS system, trehalose-specific, IIBC component	2.7	1.3	0.0227	0.0495
lmo1293	similar to glycerol 3 phosphate dehydrogenase	1.8	0.5	0.0200	0.0026
lmo1348	similar to aminomethyltransferase	1.7	0.9	0.0447	0.1346
lmo1357	acetyl-CoA carboxylase, biotin carboxylase	1.9	1.3	0.0290	0.0092
lmo1389	similar to sugar ABC transporter, ATP-binding protein	2.2	1.2	0.0020	0.0868
lmo1390	similar to ABC transporter (permease proteins)	2.2	1.2	0.0041	0.0399
lmo1391	putative ABC transporter, permease protein	1.8	1.2	0.0187	0.2869
lmo1538	similar to glycerol kinase	2.9	0.7	0.0001	0.0195

Table 2.3 (Continued)

lmo[a]	Common Name[bc]	Lineage I fold change[d]	Lineage II fold change[d]	Lineage I adj p-value[e]	Lineage II adj p-value[e]
lmo1539	similar to glycerol uptake facilitator	3.9	0.7	0.0000	0.0746
lmo1542	ribosomal protein L21	1.5	1.4	0.0354	0.0355
lmo1570	highly similar to pyruvate kinases	1.8	1.1	0.0227	0.7807
lmo1658	30S ribosomal protein S2	2.2	1.3	0.0029	0.0530
lmo1849	similar to metal cations ABC transporter, ATP-binding proteins	2.3	1.0	0.0224	0.8362
lmo1956	similar to transcriptional regulator (Fur family)	2.0	1.0	0.0148	0.9788
lmo1998	*similar to opine catabolism protein*	*5.3*	*0.7*	*0.0467*	*0.0471*
lmo2000	*PTS system, mannose/fructose/sorbose family, IID component*	*13.5*	*0.9*	*0.0019*	*0.7744*
lmo2001	*PTS system, IIC component*	*10.4*	*0.7*	*0.0089*	*0.0102*
lmo2004	*transcriptional regulator, GntR family*	*3.8*	*0.9*	*0.0379*	*0.4411*
lmo2020	similar to cell-division initiation protein (septum placement)	2.5	1.4	0.0096	0.0014
lmo2038	UDP-N-acetylmuramoylalanyl-D-glutamate--2,6-diaminopimelate ligase	2.9	1.2	0.0124	0.0404
lmo2058	similar to heme O oxygenase	2.1	1.2	0.0462	0.1411
lmo2101	pyridoxine biosynthesis protein	2.4	1.1	0.0011	0.5643
lmo2102	unknown	1.9	1.2	0.0066	0.0845
lmo2118	similar to phosphoglucomutase	1.5	0.9	0.0367	0.6917
lmo2167	metallo-beta-lactamase family protein	2.0	1.2	0.0201	0.1227
lmo2208	hydrolase, haloacid dehalogenase-like family	2.0	1.0	0.0153	0.9761
lmo2216	similar to histidine triad (HIT) protein	1.7	1.3	0.0444	0.0018
lmo2217	similar to unknown protein	1.5	1.4	0.0329	0.0010
lmo2223	conserved hypothetical protein	2.0	1.4	0.0136	0.0056

Table 2.3 (Continued)

lmo[a]	Common Name[bc]	Lineage I fold change[d]	Lineage II fold change[d]	Lineage I adj p-value[e]	Lineage II adj p-value[e]
lmo2232	CBS domain protein	2.5	1.2	0.0123	0.1146
lmo2240	similar to ABC transporter (ATP-binding protein)	1.7	1.1	0.0137	0.4621
lmo2397	similar to NifU protein	1.8	1.2	0.0244	0.3112
lmo2415	similar to ABC transporter, ATP-binding protein	2.0	1.0	0.0254	0.8350
lmo2507	highly similar to the cell-division ATP-binding protein FtsE	2.7	1.3	0.0290	0.0181
lmo2547	highly similar to homoserine dehydrogenase	1.5	0.9	0.0341	0.6523
lmo2633	ribosomal protein S10	1.6	1.0	0.0187	0.8983
lmo2660	similar to transketolase	1.7	0.9	0.0305	0.1802
lmo2664	similar to sorbitol dehydrogenase	2.2	1.1	0.0027	0.7538
lmo2665	*similar to PTS system galactitol-specific enzyme IIC component*	*2.9*	*1.0*	*0.0064*	*0.9200*
lmo2666	*similar to PTS system galactitol-specific enzyme IIB component*	*2.9*	*0.7*	*0.0004*	*0.0005*
lmo2667	*similar to PTS system galactitol-specific enzyme IIA component*	*2.3*	*0.8*	*0.0015*	*0.0316*
lmo2668	*similar to transcriptional antiterminator (BglG family)*	*2.2*	*0.7*	*0.0014*	*0.0058*
lmo2758	similar to inosine-monophosphate dehydrogenase	1.7	1.1	0.0239	0.5999
lmo2791	Partition protein, ParA homolog	1.9	1.0	0.0260	0.7694
lmoh7858_0080.5	hypothetical protein	3.2	1.0	0.0002	0.9174

Table 2.3 (Continued)

[a] Probe name based on *L. monocytogenes* EGD-e gene
[b] Common name based on EGD-e annotation
[c] Italics indicate operons
[d] Fold changes represent transcript levels in the parent strain compared to the ΔsigB strain; significant genes ≥1.5-fold and adjusted p-value <0.05
[e] Adjusted p-value <0.05 considered significant

Of particular interest are lmo1997-lmo2004, lmo2665-lmo2668, and lmo1538-lmo1539, as these putative operons are comprised of a number of genes that had higher transcript levels in the lineage I wildtype representative compared to its Δ*sigB* mutant (Figure 2.2). We found that 4 of 8 genes in the lmo1997-lmo2004 putative operon had higher transcript levels in the lineage I representative compared to its isogenic Δ*sigB* mutant, while no genes in the operon were significant in the lineage II strain (Table 3, Figure 2.2A). Lmo2003 and lmo2004 are GntR family transcriptional regulators and are upstream of 6 genes comprising components of a PTS system. In particular, lmo2001, lmo2000, and lmo1998 had 10.4, 13.5, and 5.3 fold higher transcript levels in the wildtype strain (Figure 2.2A). Similarly, lmo2665-lmo2667 encode components of a PTS system and are preceded by lmo2668 (Figures 2.2B), which is similar to a BglG family transcriptional antiterminator, suggesting that these particular PTS systems are universally important in stationary phase, but there may be differences in PTS systems utilized by different strains. Lmo1538 and lmo1539 (Figure 2.2C) encode a glycerol kinase and glycerol uptake facilitator, respectively and a recent study demonstrated that *L. monocytogenes* cells growing in the presence of glycerol results in up-regulation of all PrfA-controlled genes of the LIPI-1 cluster as well as of *inlAB*, *inlC*, and *hpt* [56]. Both lmo1538 and lmo1539 had significantly higher transcript levels in the lineage I wildtype compared to the isogenic Δ*sigB*. Transcript level differences were not significant in the lineage II representative despite a HI=100% in this study. This was consistent with observations by Raengpradub et al. [16] under the same conditions and in EGD-e (also a lineage II strain) in a recent study [35].

We found 108 genes that had higher transcript levels in the lineage II wildtype representative compared to its isogenic Δ*sigB* but that were not significantly different in lineage I (Table 2.4). In particular, the lmo0398-lmo402 putative operon, which

Table 2.4 σ^B-dependent genes identified in lineage II but not lineage I

lmo[a]	Common Name[bc]	Lineage I fold change[d]	Lineage II fold change[d]	Lineage I adj p-value[e]	Lineage II adj p-value[e]
lmo0019	conserved hypothetical protein	1.4	2.9	0.1649	0.0000
lmo0043	similar to arginine deiminase	1.9	3.7	0.3567	0.0000
lmo0135	oligopeptide ABC transporter, oligopeptide-binding protein	0.8	2.0	0.2953	0.0000
lmo0136	oligopeptide ABC transporter, permease protein	0.6	1.5	0.1425	0.0000
lmo0264	internalin C2	1.6	5.3	0.3567	0.0000
lmo0265	peptidase, M20/M25/M40 family	1.5	8.2	0.7760	0.0000
lmo0292	similar to heat-shock protein htrA serine protease	1.4	1.6	0.1117	0.0017
lmo0321	similar to unknown proteins	3.5	5.3	0.1010	0.0000
lmo0342	similar to transketolase	1.5	1.6	0.5712	0.0185
lmo0343	similar to transaldolase	1.1	2.0	0.9782	0.0011
lmo0344	similar to dehydrogenase/reductase	1.3	1.8	0.8013	0.0083
lmo0345	similar to sugar-phosphate isomerase	1.3	1.5	0.8159	0.0002
lmo0346	similar to triosephosphate isomerase	1.3	1.9	0.8578	0.0062
lmo0348	dihydroxyacetone kinase	1.2	1.8	0.9272	0.0098
lmo0398	*similar to phosphotransferase system enzyme IIA*	*0.7*	*13.5*	*0.1649*	*0.0000*
lmo0399	*PTS system, IIABC component, degenerate*	*0.9*	*10.5*	*0.8049*	*0.0000*
lmo0400	*PTS system, IIABC component, degenerate*	*1.0*	*20.5*	*0.9922*	*0.0000*
lmo0401	*glycosyl hydrolase, family 38*	*0.8*	*13.4*	*0.3601*	*0.0000*
lmo0402	*similar to transcriptional antiterminator (BglG family)*	*1.5*	*15.7*	*0.1404*	*0.0000*
lmo0439	conserved hypothetical protein	3.8	5.6	0.1291	0.0000
lmo0449	unknown	1.7	1.5	0.4564	0.0238
lmo0584	conserved hypothetical membrane protein	1.2	1.7	0.3709	0.0001
lmo0589	unknown	1.7	1.7	0.1977	0.0079
lmo0590	similar to a fusion of two types of conserved hypothetical protein, conserved hypothetical	1.1	1.7	0.9228	0.0003
lmo0591	membrane protein, putative	1.2	1.5	0.4479	0.0166
lmo0626	similar to unknown protein	1.4	2.0	0.8736	0.0005
lmo0628	unknown	2.9	3.3	0.2091	0.0000
lmo0647	unknown	1.4	2.0	0.1241	0.0003
lmo0648	magnesium transporter, CorA family	1.7	1.7	0.2015	0.0014
lmo0649	transcriptional regulator, GntR family	2.7	1.7	0.1298	0.0004

Table 2.4 (Continued)

lmo[a]	Common Name[bc]	Lineage I fold change[d]	Lineage II fold change[d]	Lineage I adj p-value[e]	Lineage II adj p-value[e]
lmo0650	conserved membrane protein	1.8	1.9	0.0763	0.0000
lmo0759	glyoxalase family protein	0.8	1.6	0.5762	0.0049
lmo0760	unknown	0.9	1.6	0.5681	0.0016
lmo0811	carbonic anhydrase	1.7	1.7	0.1107	0.0002
lmo0818	cation transport ATPase, E1-E2 family	0.7	1.7	0.4525	0.0023
lmo0819	unknown	1.5	1.7	0.0833	0.0010
lmo0928	similar to 3-methyladenine DNA glycosylase	2.0	1.5	0.1498	0.0150
lmo0929	sortase family protein	1.8	1.6	0.0663	0.0032
lmo0995	membrane protein, putative	1.1	3.4	0.9752	0.0001
lmo1037	*B. subtilis* YoaT protein homolog lmo1037 [imported]	1.0	1.6	0.9873	0.0064
lmo1064	hypothetical protein	1.4	1.7	0.7066	0.0002
lmo1072	highly similar to pyruvate carboxylase	1.2	1.5	0.5247	0.0062
lmo1121	unknown	1.9	2.4	0.0539	0.0000
lmo1226	similar to transporter, (to B. subtilis YdgH protein)	1.2	1.5	0.5341	0.0023
lmo1242	B. subtilis YdeI protein homolog lin1206	1.1	1.6	0.8038	0.0004
lmo1243	uncharacterized conserved protein, phnB family CAC3689	0.9	1.5	0.8989	0.0117
lmo1360	highly similar to methylenetetrahydrofolate dehydrogenase and methenyltetrahydrofolate cyclohydrolase	1.6	1.6	0.1421	0.0017
lmo1388	CD4 T cell-stimulating antigen, lipoprotein	1.6	1.7	0.1146	0.0066
lmo1421	similar to glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)	1.7	2.2	0.1477	0.0003
lmo1426	similar to glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein)	1.8	3.1	0.0541	0.0000
lmo1427	glycine betaine/L-proline ABC transporter, permease protein	1.3	3.0	0.4386	0.0000
lmo1527	similar to protein-export membrane protein SecDF	1.8	1.5	0.1750	0.0006
lmo1534	L-lactate dehydrogenase	1.5	1.6	0.1649	0.0005
lmo1571	6-phosphofructokinase	1.4	1.5	0.3590	0.0001
lmo1580	universal stress protein family	1.3	2.0	0.2022	0.0000
lmo1622	conserved hypothetical protein	1.2	1.6	0.5714	0.0011

Table 2.4 (Continued)

lmo[a]	Common Name[bc]	Lineage I fold change[d]	Lineage II fold change[d]	Lineage I adj p-value[e]	Lineage II adj p-value[e]
lmo1635	conserved hypothetical protein	2.1	1.7	0.1056	0.0037
lmo1636	similar to similar to ABC transporter (ATP-binding protein)	2.0	1.9	0.0936	0.0001
lmo1637	putative ABC transporter, permease protein	1.4	1.6	0.2716	0.0002
lmo1666	conserved hypothetical protein	1.4	1.6	0.7464	0.0002
lmo1681	similar to cobalamin-independent methionine synthase	1.2	1.5	0.7274	0.0061
lmo1696	putative membrane protein	1.2	1.5	0.4402	0.0073
lmo1698	acetyltransferase, GNAT family	1.1	3.5	0.8156	0.0001
lmo1702	glyoxalase family protein	1.0	1.6	0.9402	0.0369
lmo1713	cell shape-determining protein	1.5	1.5	0.4402	0.0244
lmo1749	similar to shikimate kinase	1.5	1.9	0.5157	0.0003
lmo1790	metallo-beta-lactamase family protein	1.5	1.7	0.2399	0.0001
lmo1806	highly similar to acyl carrier proteins	1.2	1.9	0.7644	0.0000
lmo1883	chitinase	2.6	3.2	0.3466	0.0000
lmo1929	*similar to nucleoside diphosphate kinase*	*1.1*	*1.8*	*0.9873*	*0.0001*
lmo1930	*heptaprenyl diphosphate syntase component II [imported]*	*2.2*	*1.6*	*0.1056*	*0.0003*
lmo1931	*2-heptaprenyl-1,4-naphthoquinone methyltransferase*	*1.9*	*1.8*	*0.0763*	*0.0002*
lmo1932	*heptaprenyl diphosphate synthase component I, putative*	*1.8*	*1.5*	*0.1616*	*0.0272*
lmo1933	*similar to GTP cyclohydrolase I*	*2.2*	*1.9*	*0.5761*	*0.0002*
lmo2031	conserved hypothetical protein TIGR00044	1.2	1.6	0.4204	0.0111
lmo2033	highly similar to cell-division protein FtsA	1.4	1.6	0.2091	0.0015
lmo2159	*oxidoreductase, Gfo/Idh/MocA family*	*1.3*	*3.4*	*0.3295*	*0.0000*
lmo2160	*AP endonuclease family 2 C terminus family*	*1.5*	*2.8*	*0.0988*	*0.0000*
lmo2161	*ThuA protein*	*1.7*	*2.2*	*0.1642*	*0.0000*
lmo2162	*conserved hypothetical protein*	*1.6*	*2.1*	*0.1513*	*0.0001*
lmo2169	unknown	1.4	1.7	0.2512	0.0000
lmo2196	similar to pheromone ABC transporter (binding protein)	0.9	1.6	0.9436	0.0006
lmo2230	similar to arsenate reductase	10.5	18.7	0.0947	0.0000
lmo2231	similar to unknown proteins	0.9	3.1	0.9720	0.0009
lmo2368	MutT/nudix family protein	1.6	1.5	0.0577	0.0054
lmo2386	similar to B. subtilis YuiD protein	1.1	1.9	0.9436	0.0000

Table 2.4 (Continued)

lmo[a]	Common Name[bc]	Lineage I fold change[d]	Lineage II fold change[d]	Lineage I adj p-value[e]	Lineage II adj p-value[e]
lmo2387	conserved hypothetical protein	1.4	4.0	0.3857	0.0000
lmo2389	similar to NADH dehydrogenase	1.9	1.8	0.1056	0.0000
lmo2399	similar to conserved hypothetical proteins	1.5	1.8	0.2577	0.0004
lmo2437	conserved hypothetical protein	1.0	2.0	0.9866	0.0000
lmo2465	conserved hypothetical protein	1.5	1.8	0.3884	0.0000
lmo2511	similar to conserved hypothetical proteins like to B. subtilis YvyD protein	1.2	1.9	0.5123	0.0001
lmo2520	N-acylamino acid racemase	0.8	1.8	0.6061	0.0000
lmo2522	similar to hypothetical cell wall binding protein from B. subtilis	1.0	1.6	0.9825	0.0028
lmo2534	ATP synthase F0, C subunit	1.4	1.5	0.1426	0.0006
lmo2536	highly similar to ATP synthase subunit i	1.4	1.6	0.2900	0.0009
lmo2568	unknown	1.4	1.6	0.6145	0.0236
lmo2611	adenylate kinase	1.0	1.5	0.9866	0.0143
lmo2638	pyridine nucleotide-disulfide oxidoreductase family protein	1.0	1.5	0.9914	0.0025
lmo2670	hypothetical protein	.	2.3	NA	0.0000
lmo2695	dihydroxyacetone kinase, Dak1 subunit, putative	1.6	1.6	0.1042	0.0003
lmo2696	dihydroxyacetone kinase family protein	1.5	1.8	0.1100	0.0001
lmo2720	acetyl-coenzyme A synthetase	1.0	1.5	0.9951	0.0022
lmo2739	transcriptional regulator, Sir2 family	1.7	1.5	0.0538	0.0038
lmo2741	major facilitator family transporter	1.7	1.5	0.0698	0.0010
lmo2742	SH3 domain protein	0.7	1.5	0.1447	0.0015
lmo2832	glycerate kinase 2	1.2	1.5	0.8437	0.0083
lmof2365_1394	hypothetical protein	1.9	1.5	0.1498	0.0068

[a] Probe name based on *L. monocytogenes* EGD-e locus tag

[b] Common name based on EGDe annotation

[c] Italics indicate operons

[d] Fold changes represent transcript levels in the parent strain compared to the Δ*sigB* strain; significant genes ≥1.5-fold and adjusted p-value <0.05

encodes components of a PTS system, a glycosyl hydrolase, and a BglG family transcriptional terminator, had at least 10.5 fold higher transcript levels in the wildtype strain (Figure 2.3A). No genes in the operon had a significant fold change in the lineage I representative despite HI >95% for each gene. Finally, the putative operon lmo2164-lmo2158 consists of 6 genes (Figure 2.3B) which includes lmo2159 and lmo2160 which encode an oxidoreductase and an endonuclease, respectively, which may play role in stationary phase survival [4]

Confirmation of select differentially expressed genes by TaqMan qRT-PCR. We used TaqMan qRT-PCR to confirm differences in transcript levels between wildtype and Δ*sigB* mutants in select genes which microarray analyses suggested were not universally σ^B-dependent among all lineage representatives. Specifically, we used TaqMan qRT-PCR to compare transcript levels between the parent strain and its isogenic Δ*sigB* mutant using probe and primer pairs based on the consensus sequence for all strains; TaqMan qRT-PCR accounted for potential differences in probe and primer binding efficiencies via the Pfaffl method [45]. As detailed in the Materials and Methods section, *rpoB* transcript levels were used as the reference gene for the relative expression ratio calculation. We found that *opuCA* was σ^B-dependent in all lineage representatives by microarray analysis (all fold changes ≥1.5; adj p-values <0.05) (Table 2.5; the *opuCABCD* operon is σ^B-dependent and is preceded by a σ^B promoter and *opuCA* has been used as an indictor of σ^B activity in other transcriptional studies [12, 57]. TaqMan qRT-PCR expression ratios (wildtype transcript levels to Δ*sigB* transcript levels) were consistent with σ^B-dependent transcription identified by microarray analyses, but were low for the lineage IIIA strain relative to the fold change determined by microarray. This may be attributable to the fact that *opuCA* is not solely dependent on □^{B} as it is transcribed from multiple promoters [57-59]. We also found that *inlA,* which encodes the well-characterized virulence factor Internalin

Table 2.5 Confirmation of select differentially expressed genes by TaqMan qRT-PCR and comparison of promoter regions determined by RNA-Seq

inlA (lmo0433)[a]

Lineage	Promotor Contruct (5'→3')[b]	Micorarray Fold Change[c]	Microarray adj. p-value[d]	Hybridization Index[e]	qRT-PCR Expression Ratio[f]
Lineage I	ATGTGTTATTTTGAACATAAAGGGTAGAGGATA	3.6	0.0004	95.7%	5.1
Lineage II	ATGTGTTATTTTGAACATAAAGGGTAGAGGATA	3.1	0.0000	98.5%	5.8
Lineage IIIA	ATGTGTTATTTTGAACATAAAGGGTAGAGGATA	4.8	0.0000	94.3%	9.6
Lineage IIIB	ATGTGTTATTTTGAACATAAAGGGTAGAGGATA	7	0.0000	94.3%	5.7*

opuCA (lmo1428)

Lineage	Promotor Contruct (5'→3')	Micorarray Fold Change[c]	Microarray adj. p-value[d]	Hybridization Index[e]	qRT-PCR Expression Ratio[f]
Lineage I	AAGTTTAAATCTATACTAGTTAGGGAAATTAGTT	2.2	0.0008	100.0%	1.6
Lineage II	AAGTTTAAATCTATACTAGTTAGGGAAATTAGTT	2.9	<0.0001	100.0%	5.9
Lineage IIIA	AAGTTTAAATCTATACTAGTTAGGGAAATTAGTT	2.7	0.0001	95.7%	1.4
Lineage IIIB	AAGTTTAAATCTATACTAGTTAGGGAAATTAGTT	3.5	0.0004	.	4.2*

Table 2.5 (Continued)

lmo0398 Lineage	Putative Promotor Contruct (5'→3')	Micorarray Fold Change[c]	Microarray adj. p-value[d]	Hybridization Index[e]	qRT-PCR Expression Ratio[f]
Lineage I	CGGTTTCATTAGAATGTAATTGTAAGCAAGGCATT	0.7	0.1649	98.5%	2.4
Lineage II	CGGTTTCATTAGAATGTAATTGTAAGCAAGGCATT	13.5	0.0000	100.0%	21.0
Lineage IIIA	CGGTTTCATTAGAATGTAATTGTAAGCAAGGCATT	3.3	0.0761	92.9%	269.3
Lineage IIIB	CGGTTTCATTAGAATGTAATTGTAAGCAAGGCATT	3.6	0.0102	77.27 (51/66)	0.9*

lmo1539 Lineage	Promotor Contruct (5'→3')	Micorarray Fold Change[c]	Microarray adj. p-value[d]	Hybridization Index[e]	qRT-PCR Expression Ratio[f]
Lineage I	AGGTTATAACTCTCGCGAATTGGGGTAAAAGTA	3.9	0.0000	100.0%	5.8
Lineage II	TGGTTATAACTCTCGCGAATTGGGGTAAAAGTA	0.7	0.0746	100.0%	1.6
Lineage IIIA	TGGTTATAACTCTCGCGAAATGGGGTAAAAGTA	3.8	0.0000	100.0%	1.1
Lineage IIIB	AGGTTATAACTCTCGCGAATTGGGGTAAAAGTA	5.6	0.0000	97.1%	5.8*

Table 2.5 (Continued)

lmo2668 Lineage	Putative Promotor Contruct (5'→3')	Micorarray Fold Change[c]	Microarray adj. p-value[d]	Hybridization Index[e]	qRT-PCR Expression Ratio[f]
Lineage I	AGATTTATAATTAAAACGAACAGGAGGGAACGAG	2.2	0.0014	100.0%	4.5
Lineage II	AGATTTATAATTAAAACGAACAGGAGGGAACGAT	0.7	0.0058	97.1%	0.4
Lineage IIIA	AGATTTATAATTAAAACGAACAGGAGGGAACGAG	1.2	0.2125	97.1%	1.8
Lineage IIIB	AGATTTATAATTAAAACGAACAGGAGGGAACGAG	2	0.0071	97.1%	3.1*

[a]*L. monocytogenes* EGD-e gene name

[b]Promoter sequenced of σ^B promoter determined previously by RACE-PCR, visual inspection, or RNA-Seq transcription startsite mapping

[c]Fold changes represent transcript levels in the parent strain compared to the Δ*sigB* strain; significant genes ≥1.5-fold and adjusted p-value <0.05

[d]Adjusted p-value <0.05 considered significant

[e]Hybridization index is the percent match between a strain and the probe used for differential expression analyses

[f]Average expression ratio of transcript levels in the parent strain compared to the Δ*sigB* strain

*These data represent two RNA extraction replicates

A, was σ^B-dependent in all lineage representatives by microarray analysis (Table 5). TaqMan qRT-PCR expression ratios of *inlA* (wildtype transcript levels to Δ*sigB* transcript levels) in all lineages were consistent with transcript level differences determined by microarray analyses. Microarray analyses indicated that there were differences in σ^B-dependent transcript levels of lmo0398, which is similar to a phosphotransferase system enzyme IIA component, among lineage representatives evidenced by large differences in fold changes (Table 5). TaqMan qRT-PCR expression ratios of lmo0398 were consistent with microarray analyses in that wide variation in transcript level differences exist among lineages. These differences may be attributable to (i) considerable variation in qRT-PCR expression ratios among replicates, and (ii) HI < 100% which underestimates transcript copy numbers, particularly in the lineage IIIA strain which had a HI=92.9% and an average qRT-PCR expression ratio of 269.3. The putative σ^B promoter region upstream of lmo0398 (determined by RNA-Seq transcriptional start site mapping [49] and visual inspection) is 100% conserved among lineage representatives while the annotated CDS is highly polymorphic (Table 5).

We also investigated the apparent lineage-specific σ^B-dependence of lmo1539 which is similar to a glycerol uptake facilitator and has been previously identified as σ^B-dependent [16, 37]. Specifically, microarray analyses indicated that lmo1539 was σ^B-dependent in lineage I, IIIA, and IIIB, but not in lineage II (Table 5); all HI >95%. TaqMan qRT-PCR supported observations that lmo1539 is σ^B-dependent in lineages I and IIIB. Differences in σ^B-dependence of lmo1539 among lineages may be partially attributed to differences in the promoter region identified by HMM and confirmed with RNA-Seq transcription start site mapping [49] as lineages I and IIIB representatives had conserved promoter sequences. Among the four lineage representatives, σ^B promoter sequences differed by 3 single nucleotide

polymorphisms, while the coding sequence of lmo1539 was highly conserved among all strains (Table 5). Similarly, a difference in σ^B-dependence among lineages was found by microarray analyses for lmo2668 which is similar to a BglG family transcriptional antiterminator. qRT-PCR expression ratios were consistent with microarray analyses supporting differences in apparent σ^B-dependence among lineage representatives. Differences may be partially due to a SNP in the promoter region identified by RNA-Seq transcription start site mapping and visual inspection [49]. This suggests that that diversification of σ^B promoter sequences among lineages may modulate some genes in the σ^B regulon and hence stress response systems among *L. monocytogenes* strains.

σ^B contributions to survival of stationary phase cells at pH 2.5 differ among strains. Wildtype and mutant cells were exposed to pH 2.5 at 37ºC with shaking for one hour; survival was assessed at three time points (10, 30, and 60 min). There was a significant difference in death rate between the lineage I ($p= 0.0004$), lineage II ($p<0.0001$), and lineage IIIB ($p=0.0047$) strains and their isogenic Δ*sigB* mutant as indicated by a statistically significant interaction between the effect of "time*strain" (Table 2.6). There was no significant difference in death rate between the lineage IIIA wildtype representative strain and its isogenic Δ*sigB* mutant ($p=0.2920$). Therefore, σ^B contributes to survival of the lineage I, II, and IIIB representatives at pH 2.5 but offers little apparent contribution to the survival of the IIIA representative under the same conditions. Further, the largest difference in death rate (~5 log CFU/h) was observed between the lineage II wildtype and its Δ*sigB* mutant suggesting that σ^B played the largest apparent role in acid stress survival in lineage II.

σ^B contributions to survival of stationary phase cells under oxidative stress differ among strains. After exposure to 13mM CHP in DMSO for 15 min, there were

Table 6. σ^B contributions to acid and oxidative stress survival among lineage represetnatives

	Acid Stress, pH 2.5 for 1 h			Oxidative stress, 13mM CHP for 15 min		
Strain	Wildtype[a]	Δ*sigB*[b]	p-value[c]	Wildtype[d]	Δ*sigB*[e]	p-value[f]
Lineage I	3.95 ± 0.93	6.62 ± 0.66	0.0004	1.27 ± 0.04	2.10 ± 0.23	0.0225
Lineage II	1.75 ± 0.49	6.76 ± 0.79	<0.0001	1.10 ± 0.08	2.57 ± 0.12	<0.0001
Lineage IIIA	4.02 ± 2.23	5.74 ± 1.89	0.2920	1.14 ± 0.22	1.74 ± 0.40	0.0827
Lineage IIIB	3.65 ± 0.77	6.37 ± 0.73	0.0047	1.38 ± 0.30	2.09 ± 0.22	0.0291

[a]Average death rate (log CFU/h) of wildtype parent strain

[b]Average death rate (log CFU/h) of isogenic Δ*sigB* strain

[c]P-value of time*strain interaction; p-value <0.05 indicates significant difference in the average death rate between wildtype and isogenic Δ*sigB* strain

[d]Average death (log CFU) of wildtype parent strain

[e]Average death (log CFU) of isogenic Δ*sigB* strain

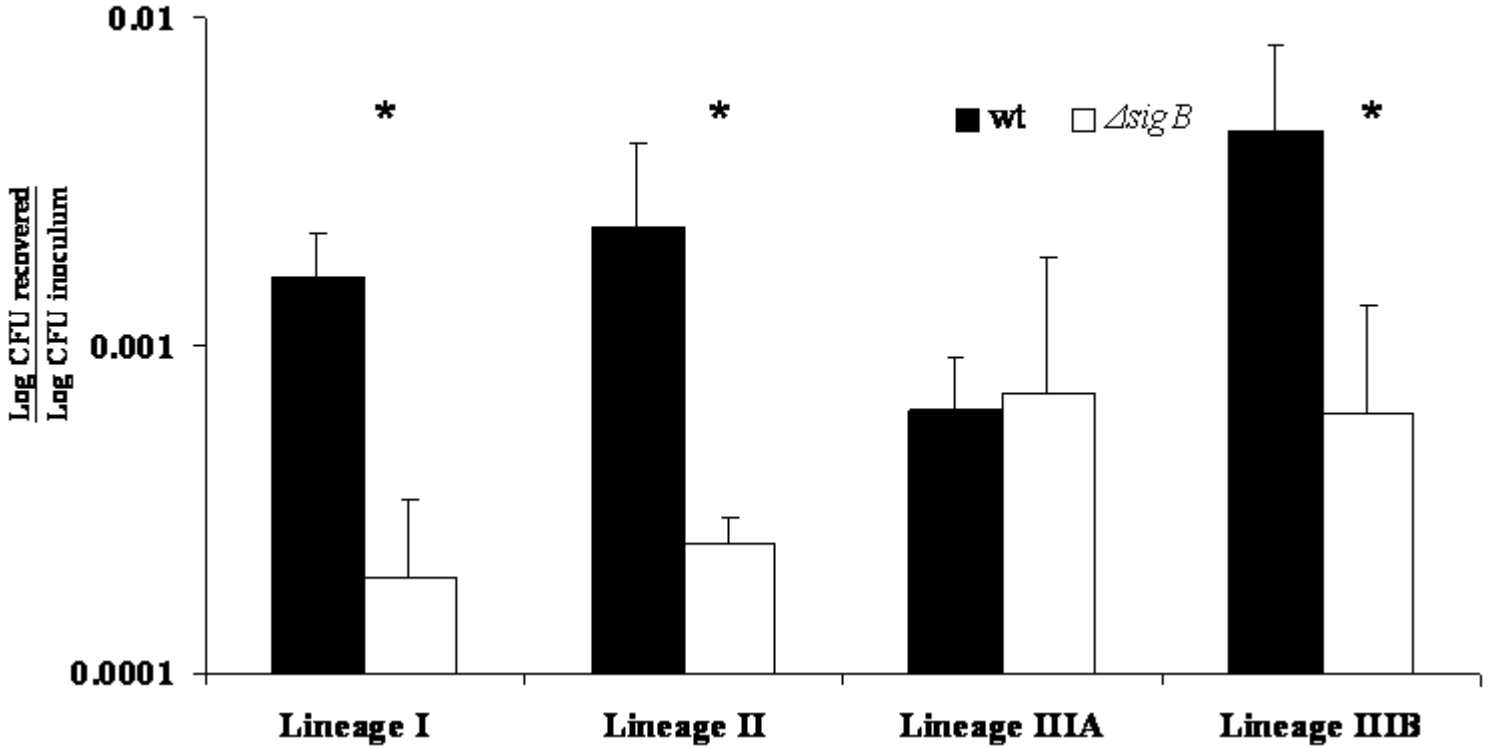
[f]P-value of one-sided t-test; p-value <0.05 indicates significant difference in the average death between wildtype and isogenic Δ*sigB* strain

significant differences in the number of cells killed (log CFU) between lineage I (p=0.0225), lineage II (p=<0.0001) and lineage IIIB (p=0.0291) wildtype strains and their Δ*sigB* mutants (Table 6). While σ^B did not significantly contribute to the survival of the lineage IIIA representative under these conditions (p= 0.0827), it played a significant role in oxidative stress survival in lineage I, II, and IIIB. Finally, σ^B played the largest apparent role in oxidative stress survival in the lineage II representative evidenced by the greatest difference been wildtype and Δ*sigB* mutant average death (~1.4 log CFU).

σ^B contributions to invasion of Caco-2 cells by stationary phase *L. monocytogenes* cells differ among lineage representatives. To assess the contribution of σ^B to invasion of Caco-2 epithelial cells by stationary phase *L. monocytogenes* cells, Caco-2 cells were exposed to each strain for 30 minutes followed by 3 washes with PBS and treatment with media containing gentamicin to kill extracellular bacteria (Figure 2.4). Significant differences were found between lineage I (p=0.0004), lineage II (p=0.0319), and lineage IIIB (p=0.0192) wildtype strains when compared to their isogenic Δ*sigB* mutants. No significant difference was found in invasion efficiency between lineage IIIA representative and its Δ*sigB* mutant (p=0.1991). Therefore, σ^B significantly contributed to the invasion of Caco-2 cells in lineages I, II, and IIIB, representative wildtype strains in stationary phase. σ^B played a less evident role in the lineage IIIA representative strain's ability to invade under the same conditions.

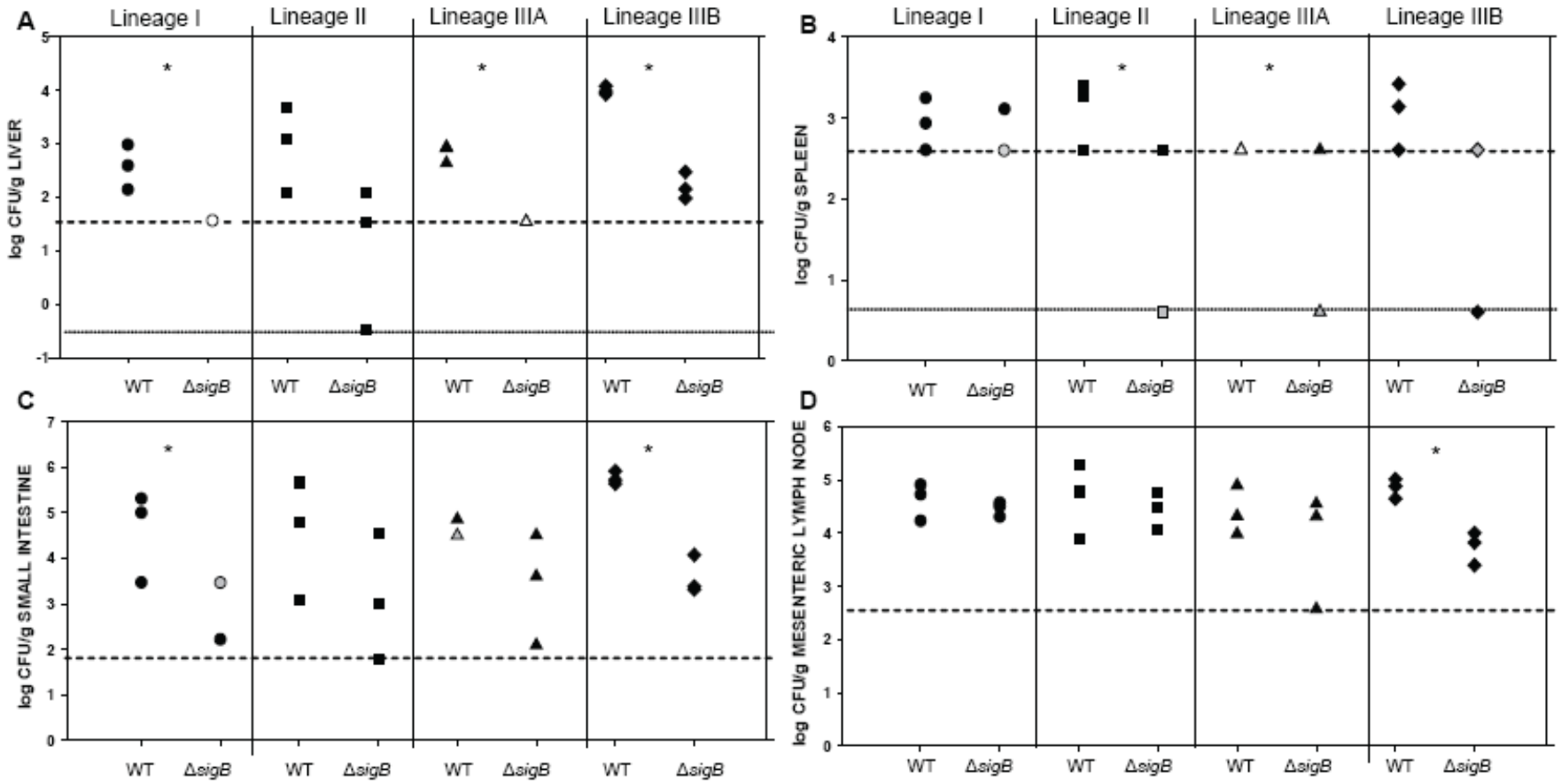
Guinea pig intragastric infection model. A previous study using the guinea pig intragastric infection model system of listeriosis demonstrated that σ^B significantly contributed to virulence [19] as evidenced by fewer bacterial numbers recovered from animals infected with the Δ*sigB* mutant relative to the wildtype strain. We used the guinea pig intragastric infection model to determine if σ^B contributed to virulence in strains representing *L. monocytogenes* lineages. Bacterial numbers (in log CFU/g)

Figure 2.4 Invasion efficiency (log ratio of CFU recovered to log CFU initial inoculum) of stationary phase cells. Data shown are an average of at least three independent experiments and error bars indicate standard deviation. Each lineage wildtype representative is paired with its isogenic Δ*sigB* mutant. Each pair was assessed by two-sample t-test; * = $p < 0.05$.


0.01
0.001
0.0001
Log CFU recovered
Log CFU inoculum
*
*
*
wt
Δsig B
Lineage I
Lineage II
Lineage IIIA
Lineage IIIB

were determined in five organs (brain, liver, mesenteric lymph node, spleen, and small intestine) harvested from each animal intragastrically infected with the *L. monocytogenes* mutant or wild-type strains at 72 h post-infection (Figure 2.5;) three animals were infected with each strain. One-sided t-tests were used to determine significant differences between bacterial numbers from organs in animals infected with *L. monocytogenes* wildtype or Δ*sigB* strain. In intragastrically infected animals, the Δ*sigB* mutants were present in lower numbers than those recovered from animals infected with the wildtype strain in some organs (Figure 2.5). Specifically, there were significantly lower bacterial numbers recovered from the liver ($p=0.0459$) and small intestine ($p=0.0254$) of animals infected with the Δ*sigB* mutants representing lineages I compared to the wildtype strain. Among the lineage II wildtype and Δ*sigB* mutant pair, there were significantly lower bacterial numbers recovered from the spleen ($p=0.0314$) of animals infected with the Δ*sigB* mutant. There were significantly lower bacterial numbers recovered from the liver ($p=0.0028$) and spleen ($p=0.0028$) of animals infected with the Δ*sigB* mutants representing lineages IIIA compared to the wildtype strain and there were lower bacterial numbers recovered from the liver ($p=0.0006$), small intestine ($p=0.0031$), and mesenteric lymph node ($p=0.0002$) of animals infected with the Δ*sigB* mutant representing lineage IIIB. For some organs, bacterial numbers recovered from animals intragastrically infected with the Δ*sigB* strain were reduced, but not significantly different from the wildtype (e.g. lineage II Δ*sigB* mutant was reduced in the liver compared to the wildtype strain). *L. monocytogenes* was not recovered from the brain of any animal regardless of the strain with which they were infected under these conditions (data not shown). While differences in this study are reduced compared to the disparity in log CFU/g between *L. monocytogenes* wildtype and Δ*sigB* mutant described previously [19], these differences are likely attributable to i) fewer replicates used for each comparison in an

Figure 2.5 Log CFU/g *L. monocytogenes* recovered from organs. Scatter plot of *L. monocytogenes* recovered from the organs of guinea pigs at 72 h post-intragastric infection. Strains (wildtype and Δ*sigB* for each lineage representative) are indicated on the *x* axes. Bacterial numbers, in log CFU/g, from the liver (A), spleen (B), small intestine (C) and mesenteric lymph nodes (D) are shown on the *y* axes in independent panels. Data were obtained from three guinea pigs intragastrically infected with each strain. Closed symbols indicate a single data point, shaded symbols indicate two data points, and open symbols indicate three data points. The detection limits, which differ among organs due to different organ weights, are indicated by horizontal broken and solid lines in each panel. The broken horizontal line indicates the plating detection limit; the solid horizontal lines in panels A and B indicate the enrichment detection limit. Data reported at the plating detection limit were positive for *L. monocytogenes* after enrichment, but had bacterial counts below that detectable by standard plate count. Data reported at the enrichment detection limit had no recovery of *L. monocytogenes* after enrichment. Asterisks indicate significantly (one-sided t-test; p-value<0.05) higher bacterial numbers recovered from organs from animals infected with a wildtype strain compared to its isogenic Δ*sigB* mutant.


A
Lineage I
Lineage II
Lineage IIIA
Lineage IIIB
log CFU/g LIVER
B
log CFU/g SPLEEN
C
log CFU/g SMALL INTESTINE
D
log CFU/g MESENTERIC LYMPH NODE
WT
ΔsigB
*

effort to reduce animal usage, which subsequently results in a decrease in the power to detect differences, and ii) the inherent variation resulting from a complex biological system. Holistically, σ^B contributed to virulence in all strains representing *L. monocytogenes* lineages.

Using a one-sided t-test, we tested if there was a difference in the log CFU/g *L. monocytogenes* shed in feces from animals infected with a wildtype and its isogenic Δ*sigB* mutant strains at 72 h post-infection. We found that there were significantly reduced number of Δ*sigB* mutant cells recovered from feces (compared to their respective wildtype strain) in the lineage I (p=0.0163), II (p=0.0163), and IIIA (p=0.0472) representatives (Table 2.7). There were fewer Δ*sigB* mutant cells recovered from feces in the lineage IIIB representative compared to the wildtype strain, but it was not statistically significant (p=0.0532) (Table 7). This suggests that the Δ*sigB* mutants have a decreased ability to i) attach to intestinal epithelial cells, and/or ii) survive in the intestinal tract. A previous study demonstrated that the Δ*sigB* mutant was recovered from feces at significantly lower levels compared to the wildtype strain 72 h post-infection [19].

Finally, we tested whether there was a significant difference between the weights of animals at 72 h post-infection which had been infected with either i) wildtype, or ii) Δ*sigB* mutant strains, using a one-sided t-test. Weight at 72 h post infection was calculated as the percentage of the weight of the animal at time of infection, which was set to 100%. While there was not a significant difference in weights of animal infected with the wildtype or Δ*sigB* mutant from lineage I (p=0.2129) or lineage II (p=0.1230), animals infected with wildtype strains representing lineages IIIA (p=0.0465) and lineage IIIB (p=0.0305) weighed significantly less than animals infected with the respective isogenic Δ*sigB* mutant (Table 7).

Table 7. Guinea pig weight and fecal shedding of *L. monocytogenes*72 h post-infection

	Weight (%)			Fecal Shedding, log CFU/g		
Strain	Wildtype[a]	Δ*sigB*[b]	p-value[c]	Wildtype[d]	Δ*sigB*[e]	p-value[f]
Lineage I	101.1 ± 8.9	106.5 ± 5.6	0.2129	5.3 ± 0.9	3.1 ± 0.7	0.0163
Lineage II	99.0 ± 8.4	106.6 ± 4.9	0.123	4.8 ± 1.5	0.0 ± 0.0	0.0163
Lineage IIIA	99.4 ± 3.7	105.7 ± 3.3	0.0465	6.1 ± 1.4	2.6 ± 2.4	0.0472
Lineage IIIB	91.6 ± 2.6	105.9 ± 9.2	0.0305	6.1 ± 0.1	3.3 ± 2.9	0.0532

[a]Average percentage of weight of guinea pig infected with wildtype strains at 72 h post-infection (euthanasia) relative to 0 h (infection) set at 100%

[b]Average percentage of weight of guinea pig infected with Δ*sigB* strain at 72 h post-infection relative to 0 h set at 100%

[c]P-value of t-test of weights of guinea pigs infected with wildtype or Δ*sigB* strains; p-value <0.05 indicates significant difference in animal weight 72 h post-infection

[d]Average fecal shedding (log CFU/g) of wildtype parent strain 72 h post-infection

[e]Average fecal shedding (log CFU/g) of Δ*sigB* strain 72 h post-infection

[f]P-value of t-test; p-value <0.05 indicates significant difference in the average log CFU/g *L. monocytogenes* shed in feces between wildtype and isogenic Δ*sigB*

DISCUSSION

In this study, we used transcriptomic and phenotypic approaches to define and characterize the contribution of σ^B to stress response and virulence in four strains representing *L. monocytogenes* strain diversity including lineages I, II, IIA and IIIB. The data generated using these approaches showed that (i) σ^B-dependent genes in *L. monocytogenes* include a pan-regulon of approximately 400 genes that are σ^B -dependent in at least one strain and a core regulon of at least 60 genes that are σ^B-dependent in all strains characterized here, (ii) contributions of σ^B to acid and oxidative stress resistance differ among strains; and (iii) while σ^B only contributes to *in vitro* intestinal epithelial cell invasion in some strains, it contributes to guinea pig virulence in all *L. monocytogenes* strains tested, further supporting strain specific contributions of σ^B to gene regulation in *L. monocytogenes* virulence.

σ^B-dependent genes in *L. monocytogenes* include a pan-regulon of approximately 400 genes and a core regulon of at least 60 genes. Whole genome microarray approaches previously identified more than 200 genes, which include both positively and negatively regulated genes, in the σ^B regulon in *L. monocytogenes* [15, 16, 35]. However, these studies focused on well-characterized laboratory type strains, which primarily included lineage II strains. Thus, there is a disparity in our understanding of the role of σ^B in lineage I strains, which are responsible for the majority of human listeriosis cases, and lineage III strains, which appear to be associated with animal listeriosis [30, 33]. In an effort to capture the inherent transcript sequence differences among strains, we used a multi-strain construct microarray to investigate positively differentially expressed σ^B-dependent genes in *L. monocytogenes* strains representing lineages I, II, IIIA, and IIIB. We found that, among four sets of paired strains representing *L. monocytogenes* lineages, approximately 400 genes comprise the σ^B

pan-regulon. Over 60 of these genes were σ^B-dependent among all four strains and thus constitute the core σ^B regulon. Our data indicate that the core σ^B regulon identified here is consistent with σ^B regulons identified in previous studies [15, 16, 35]. Of the 63 genes identified as □^{B}-dependent in all strains, 56 were previously identified by (i) Raengpradub et al. [16], who identified □^{B}-dependent genes using *L. monocytogenes* strains and growth conditions identical to those in this study, and 59 were previously identified by (ii) Ollinger et al. [15], who identified σ^B-dependent genes by comparing transcripts from *L. monocytogenes* 10403S with a PrfA* (G155S) allele [60], which constitutively expresses PrfA-regulated virulence genes [60-62], with those from an isogenic Δ*sigB* mutant grown to stationary phase under the same conditions. Further, we compared our results with those from a microarray study using another *L. monocytogenes* strain (EGD-e) and its isogenic Δ*sigB* mutant, grown under similar conditions (i.e., growth to early stationary phase [35]); where 45 of the 63 core σ^B-dependent genes were identified. While it is likely that the core σ^B regulon is underestimated in this study due to (i) low hybridization indexes (HI) for some genes for some strains, and (ii) use of a single growth condition, using a single microarray platform, we identified genes that are σ^B-dependent in a diverse set of strains representing the major genetic lineages in *L. monocytogenes*. A large proportion of genes identified in the current study to be σ^B-dependent among all strains play a role in metabolism or have not yet been characterized, warranting further investigation in to the function of these genes that appear to be universally σ^B-dependent.

Using our approach, we found that, among four strains representing *L. monocytogenes* lineages, the σ^B pan-regulon consists of over 400 genes. σ^B regulons have been characterized in a number of Gram positive organisms including *Staphylococcus aureus,* which includes over 120 σ^B-dependent genes including

virulence genes [63], *Bacillus subtilis* which has more than 150 σ^B-dependent genes [64], and *Bacillus cereus* which includes just over 20 σ^B-dependent genes under the conditions tested [65]. These studies found that genes in the σ^B regulons of Gram positive bacteria constitute a wide variety of genes, including genes involved in energy metabolism, regulatory functions, and pathogenesis [16, 35, 63-65]. In this study, *L. monocytogenes* σ^B-dependent genes were overrepresented in among the following role categories: Cellular Processes, DNA Metabolism, Energy Metabolism, Regulatory Function and Unknown Function *L. monocytogenes* JCVI role categories clearly indicate that σ^B is critically involved in a vast array of mechanisms necessary for cell function and survival. This is consistent with a recent study [35] of the σ^B regulon in *L. monocytogenes* EGD-e which was found to include genes involved in wide range of metabolic functions, general stress proteins, and unknown functions. The majority of genes categorized in the Regulatory Functions role category were involved with transcriptional regulation and included regulators in the GntR family and MerR-like regulators which have been shown to be global regulators of primary metabolism [52-54] and play a role in optimizing σ^{70}-dependent promoters in Gram negative bacteria [55], respectively. *L. monocytogenes* transcriptional regulatory networks were recently reviewed by Chaturongakul et al. [66], who summarized that σ^B may serve as a master regulator involved in a number of transcriptional regulatory networks necessary to fine-tune stress response and virulence and thus form the *L. monocytogenes* σ^B "modulon". Similarly, σ^B-mediated general stress response in *B. subtilis* is a critical component of the gene expression network and has been shown to connect multiple regulons thus serving as a master regulator [64]. Our findings contribute to an emerging body of evidence supporting that σ^B plays an important role in the global regulation of stress-response and virulence across *L. monocytogenes* lineages.

While σ^B clearly serves as a major transcription regulator in *L. monocytogenes* [15, 16, 35], we found evidence that the σ^B regulon varies among strains representing *L. monocytogenes* lineages. Although differences the σ^B regulon exist among all strains in this study, we focused on the differences between lineages I and II as the majority of human listeriosis cases result from infections caused by strains in these lineages [30]. Considerable differences in the σ^B regulons between lineage I and II representative strains were observed; where genes involved in energy metabolism and transport in addition to genes encoding hypothetical or unknown proteins differentiated the lineage I and II sigma B regulons. While many genes with significant fold changes unique to lineage I or II were marginally different from fold changes observed in the corresponding lineage representative, differences in PTS systems, fine-tuned differences in a number of σ^B-dependent transcript levels, and functions of known and hypothetical proteins may account for differences in the ability to survive stress and subsequently cause disease. Recent work by Severino et al. found that transcriptional profiles differed between wildtype strains representing *L. monocytogenes* lineages I and II [37]. Using a macroarray, differences in transcript levels in genes involved in metabolism, virulence-associated genes, and σ^B-related genes were found to be overexpressed in lineage II compared to lineage I (using lineage designations described by Wiedmann et al.[20]). Among the σ^B-related genes found to be overexpressed in lineage II by macroarray analyses, we found five genes which were σ^B-dependent in lineage II but not in lineage I.

In this study, we found evidence that differences in σ^B-dependence of a gene among lineage representatives may be attributable to differences in σ^B promoter sequences (i.e. lmo1539, lmo2668; Table 2.5). Specifically, diversification of σ^B promoter sequences among lineages may modulate the σ^B regulon and hence stress

response systems among *L. monocytogenes* strains. Interesting, van Schaik et al. found that, despite vast numbers of genes in the σ^B regulon in Gram positive organisms, only three genes (*rsbV*, *rsbW*, and *sigB*) were σ^B-dependent among *L. monocytogenes, S. aureus, B. cereus*, and *B. subtilis* [65]. This suggests that σ^B regulons may have evolved to perform niche-specific functions and may partially account for the differences in σ^B regulons in *L. monocytogenes* indentified in this study. More specifically, variation in transcriptional patterns, as a result of diversification, may provide insight into differences in *L. monocytogenes* strain-specific abilities to respond to environmental stress, interactions with host cells, and potential to cause disease.

Phenotypic characterization reveals contributions of σ^B to stress survival differ among lineage representatives. In our study, microarray analyses indicated that there were differences in the positively regulated σ^B regulons among strains. To determine if these transcriptional differences correlated with phenotypic differences, we tested the ability of each paired set of wildtype and Δ*sigB* mutant to resist acid and oxidative stress. We found that σ^B played a significant role in acid and oxidative stress resistance in lineages I, II, and IIIB. This is consistent with previous studies that demonstrated σ^B contributes to environmental stress survival (e.g. acid, oxidative, and energy stresses) as shown by reduced survival of an isogenic lineage II Δ*sigB* mutant under the same conditions [3-5]. Further, it has been shown that σ^B played a significant role in acid and salt stress resistance in *Listeria innocua* [16], as well as in *S. aureus* [67], *B. cereus* [68], *B. subtilis* [69, 70] when exposed to a myriad of stress conditions. This is the first report demonstrating that σ^B contributes to stress survival in lineage I, which is responsible for the majority of human sporadic and epidemic

listeriosis cases. Importantly, these results support that σ^B may play an important role in the survival and transmission of those *L. monocytogenes* strains that represent the most significant human health risk (i.e., lineage I strains) along the food continuum.

However, we found that σ^B played a limited role in stress resistance in *L. monocytogenes* lineage IIIA representative as there were no significant difference between wildtype and Δ*sigB* mutant's ability to resist acid and oxidative stress. Therefore despite significant evidence for σ^B-dependent differential gene expression and virulence, the *L. monocytogenes* lineage IIIA strain (serotype 4c) showed limited σ^B-dependent phenotype under the conditions tested. The apparent σ^B-independent phenotype of the lineage IIIA representative observed under these conditions is consistent with a previous study [36] which identified differences in σ^B contributions to environmental stress between a serotype 1/2a and 4c strains. Specifically, σ^B played a negligible role in acid, oxidative, and heat stress resistance in the serotype 4c strain as there were no significant differences between the wildtype strain and its isogenic Δ*sigB* mutant survival [36]. Phenotypic diversification of response regulation has also been observed in other bacteria including *Escherichia coli* [71] and *Salmonella* Typhimurium [72].

While σ^B only contributes to *in vitro* intestinal epithelial cell invasion in some strains, it contributes to guinea pig virulence in all *L. monocytogenes* strains tested, further supporting strain specific contributions of σ^B to gene regulation in *L. monocytogenes*. In this study, we used the Caco-2 intestinal epithelial cell model to determine if σ^B contributions to invasion efficiency differed between *L. monocytogenes* lineages representatives. Previous studies [18, 19] demonstrated that in *L. monocytogenes* 10403S (a lineage II strain), σ^B was important for invasion of

Caco-2 cells evidenced by reduced recovery of Δ*sigB* mutant cells compared to the parent strain. Here we demonstrated that wildtype strains from lineage I, II, and IIIB had significantly higher invasion efficiency than their isogenic Δ*sigB* mutants which indicated that σ^B played a role in invasion of Caco-2 cells. Kim et al. demonstrated that σ^B significantly contributed to *L. monocytogenes* invasion of human enterocytes and hepatocytes, predominantly through InlA- and InlB-mediated pathways [14], as both *inlA* and *inlB* are at least partially transcribed from σ^B promoters [14, 18]. Further, a recent study by Ollinger et al. provided evidence that σ^B regulates overall PrfA activity "switching" from transcriptional activation at the P2(prfA) promoter to posttranscriptional downregulation of PrfA regulon expression [15] which facilitates tightly coordinated repression and induction of virulence gene expression. Interestingly, we found that there was no contribution of σ^B in the *L. monocytogenes* lineage IIIA representative to invasion efficiency of Caco-2 cells despite the fact that (i) *L. monocytogenes* lineage IIIA transcribes a full-length *inlA* transcript (no *inlA* premature stop codon), and (ii) microarray analyses identified over 200 genes with higher transcript levels in the wildtype strain compared to the Δ*sigB* mutant indicating σ^B-dependence of these genes including well-characterized σ^B-dependent genes (e.g. *bsh, opuCA, inlA, gadB* [12]).

We used the guinea pig intragastric model for listeriosis as an *in vivo* system to quantify and characterize σ^B contributions to virulence in *L. monocytogenes* strains representing lineages I, II, IIIA, and IIIB; the guinea pig model allowed for InlA-mediated adhesion and invasion [50]. We found that all *L. monocytogenes* Δ*sigB* strains had reduced virulence in the guinea pig intragastric infection model evidenced by reduced recovery of the Δ*sigB* strains relative to their isogenic parent in at least one organ. This, in conjunction with reduced invasion efficiencies of human Caco-2 intestinal epithelial cells for most Δ*sigB* strains, (relative to their isogenic parent)

demonstrated the importance of σ^B for *L. monocytogenes* virulence in intragastric infection. These results are consistent with a previous study [19] which investigated the role of σ^B in virulence using *L. monocytogenes* 10403S and an isogenic Δ*sigB* mutant grown under then same conditions in a guinea pig intrgastric infection model. This study clearly demonstrated reduced virulence of the Δ*sigB* strain [19]. Consistent with Garner et al., we saw the largest differences (CFU/g recovered) between wildtype and Δ*sigB* mutants in the liver and small intestine [19], higher weights (relative to time of infection) in animals infected with Δ*sigB* mutants, and reduced fecal shedding of Δ*sigB* mutants compared to wildtype in most lineage representatives. Therefore, while σ^B only contributes to *in vitro* intestinal epithelial cell invasion and acid and stress survival in some strains, it contributes to guinea pig virulence in all *L. monocytogenes* strains tested, further supporting strain specific contributions of σ^B to gene regulation in *L. monocytogenes* virulence. Phenotypic diversification provides species with a capacity to survive environmental adversity and is a key player in niche adaptation [73] and therefore may partially contribute to differences in *L. monocytogenes* strains' abilities to cause disease.

ACKNOWLEDGEMENTS

We thank S. Chaturongakul, M.E. Palmer, E.D. Fortes, and Y. Soyer for their assistance with the guinea pig gastrointestinal model for infection.

Financial Disclosure: This work was funded by NIH-NIAID (R01 AI052151 to K.J.B.).

REFERENCES

1. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV: **Food-related illness and death in the United States.** *Emerg Infect Dis* 1999, **5:**607-625.
2. FAO/WHO: **Microbiological risk assessment series 5: risk assessment of *Listeria monocytogenes* in ready-to-eat foods. Available at:** http://www.fao.org/documents. 2004.
3. Wiedmann M, Arvik TJ, Hurley RJ, Boor KJ: **General stress transcription factor s^B and its role in acid tolerance and virulence of *Listeria monocytogenes*.** *J Bacteriol* 1998, **180:**3650-3656.
4. Ferreira A, O'Byrne CP, Boor KJ: **Role of s^B in heat, ethanol, acid, and oxidative stress resistance and during carbon starvation in *Listeria monocytogenes*.** *Appl Environ Microbiol* 2001, **67:**4454-4457.
5. Chaturongakul S, Boor KJ: **RsbT and RsbV contribute to s^B-dependent survival under environmental, energy, and intracellular stress conditions in *Listeria monocytogenes*.** *Appl Environ Microbiol* 2004, **70:**5349-5356.
6. Wu S, de Lencastre H, Tomasz A: **Sigma-B, a putative operon encoding alternate sigma factor of *Staphylococcus aureus* RNA polymerase: molecular cloning and DNA sequencing.** *J Bacteriol* 1996, **178:**6036-6042.
7. Fouet A, Namy O, Lambert G: **Characterization of the operon encoding the alternative sigma(B) factor from *Bacillus anthracis* and its role in virulence.** *J Bacteriol* 2000, **182:**5036-5045.
8. Brody MS, Price CW: ***Bacillus licheniformis* sigB operon encoding the general stress transcription factor sigma B.** *Gene* 1998, **212:**111-118.
9. Loewen PC, Hengge-Aronis R: **The role of the sigma factor sigma S (KatF) in bacterial global regulation.** *Ann Rev Microbiol* 1994, **48:**53-80.

10. Fang FC, Chen CY, Guiney DG, Xu Y: **Identification of sigma S-regulated genes in *Salmonella typhimurium*: complementary regulatory interactions between sigma S and cyclic AMP receptor protein.** *J Bacteriol* 1996, **178:**5112-5120.

11. Hengge-Aronis R: **Survival of hunger and stress: The role of rpoS in early stationary phase gene regulation in *E. coli*.** *Cell* 1993, **72:**165-168.

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

13. Kazmierczak MJ, Wiedmann M, Boor KJ: **Contributions of *Listeria monocytogenes* σ^B and PrfA to expression of virulence and stress response genes during extra- and intracellular growth.** *Microbiology (Read)* 2006, **152:**1827-1838.

14. Kim H, Marquis H, Boor KJ: **σ^B contributes to *Listeria monocytogenes* invasion by controlling expression of *inlA* and *inlB*.** *Microbiology (Read)* 2005, **151:**3215-3222.

15. Ollinger J, Bowen B, Wiedmann M, Boor KJ, Bergholtz TM: ***Listeria monocytogenes* σ^B modulates PrfA-mediated virulence factor expression.** *Infect Immun* 2009, **77:**2113-2124.

16. Raengpradub S, Wiedmann M, Boor KJ: **Comparative analysis of the s^B-dependent stress responses in *Listeria monocytogenes* and *Listeria innocua* strains exposed to selected stress conditions.** *Appl Environ Microbiol* 2008, **74:**158-171.

17. Rauch M, Luo Q, Muller-Altrock S, Goebel W: **SigB-Dependent In Vitro Transcription of prfA and Some Newly Identified Genes of *Listeria monocytogenes* Whose Expression Is Affected by PrfA In Vivo.** *J Bacteriol* 2005, **187:**800-804.

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

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

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

21. Wiedmann M, Bruce JL, Knorr R, Bodis M, Cole EM, McDowell CI, McDonough PL, Batt CA: **Ribotype diversity of *Listeria monocytogenes* strains associated with outbreaks of listeriosis in ruminants.** *J Clin Microbiol* 1996, **34:**1086-1090.

22. Lozniewski A, Humbert A, Corsaro D, Schwartzbrod J, Weber M, Le Faou A: **Comparison of sludge and clinical isolates of *Listeria monocytogenes*.** *Lett Appl Microbiol* 2001, **32:**336-339.

23. Vela AI, Fernandez-Garayzabal JF, Vazquez JA, Latre MV, Blanco MM, Moreno MA, de La Fuente L, Marco J, Franco C, Cepeda A, et al: **Molecular typing by pulsed-field gel electrophoresis of Spanish animal and human *Listeria monocytogenes* isolates.** *Appl Environ Microbiol* 2001, **67:**5840-5843.

24. Salcedo C, Arreaza L, Alcala B, de la Fuente L, Vazquez JA: **Development of a multilocus sequence typing method for analysis of *Listeria monocytogenes* clones.** *J Clin Microbiol* 2003, **41:**757-762.

25. Cai S, Kabuki DY, Kuaye AY, Cargioli TG, Chung MS, Nielsen R, Wiedmann M: **Rational design of DNA sequence-based strategies for subtyping *Listeria monocytogenes*.** *J Clin Microbiol* 2002, **40:**3319-3325.

26. Piffaretti JC, Kressebuch H, Aeschbacher M, Bille J, Bannerman E, Musser JM, Selander RK, Rocourt J: **Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease.** *P NAS U S A* 1989, **86:**3818-3822.

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

28. Nadon CA, Woodward DL, Young C, Rodgers FG, Wiedmann M: **Correlations between molecular subtyping and serotyping of *Listeria monocytogenes*.** *J Clin Microbiol* 2001, **39:**2704-2707.

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

30. Gray MJ, Zadoks RN, Fortes ED, Dogan B, Cai S, Chen Y, Scott VN, Gombas DE, Boor KJ, Wiedmann M: ***Listeria monocytogenes* isolates from foods and humans form distinct but overlapping populations.** *Appl Environ Microbiol* 2004, **70:**5833-5841.

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

32. Norton DM, Scarlett JM, Horton K, Sue D, Thimothe J, Boor KJ, Wiedmann M: **Characterization and pathogenic potential of *Listeria monocytogenes* isolates from the smoked fish industry.** *Appl Environ Microbiol* 2001, **67:**646-653.

33. Jeffers GT, Bruce JL, McDonough PL, Scarlett J, Boor KJ, Wiedmann M: **Comparative genetic characterization of *Listeria monocytogenes* isolates from human and animal listeriosis cases.** *Microbiology (Reading, England)* 2001, **147:**1095-1104.

34. Chen Y, Ross WH, Gray MJ, Wiedmann M, Whiting RC, Scott VN: **Attributing risk to *Listeria monocytogenes* subgroups: dose response in relation to genetic lineages.** *J Food Prot* 2006, **69:**335-344.

35. Hain T, Hossain H, Chatterjee SS, Machata S, Volk U, Wagner S, Brors B, Haas S, Kuenne CT, Billion A, et al: **Temporal transcriptomic analysis of the *Listeria monocytogenes* EGD-e s^B regulon.** *BMC Microbiol* 2008, **8:**20.

36. Moorhead SM, Dykes GA: **The role of the *sigB* gene in the general stress response of *Listeria monocytogenes* varies between a strain of serotype 1/2a and a strain of serotype 4c.** *Curr Microbiol* 2003, **46:**461-466.

37. Severino P, Dussurget O, Vencio RZ, Dumas E, Garrido P, Padilla G, Piveteau P, Lemaitre JP, Kunst F, Glaser P, Buchrieser C: **Comparative transcriptome analysis of *Listeria monocytogenes* strains of the two major lineages reveals differences in virulence, cell wall, and stress response.** *Appl Environ Microbiol* 2007, **73:**6078-6088.

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

39. Chaturongakul S, Boor KJ: **s^B Activation under environmental and energy stress conditions in *Listeria monocytogenes*.** *Appl Environ Microbiol* 2006, **72:**5197-5203.

40. Horton RM, Cai ZL, Ho SN, Pease LR: **Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction.** *BioTechniques* 1990, **8:**528-535.

41. Smyth GK: **Limma: linear models for microarray data. In: Bioinformatics and Computational Biology Solutions using R and Bioconductor, R.** V Carey, S Dudoit, R Irizarry, W Huber (eds), Springer, New York,397-420 2005.

42. Gentleman R, V. Carey, D. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Yang, and J. Zhang: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5:R80**.

43. Smyth GK, and T. Speed.: **Normalization of cDNA microarray data.** *Methods* 2003, **31:**265-273.

44. Smyth GK: **Linear models and empirical bayes methods for assessing differential expression in microarray experiments.** *Stat Appl Genet Mol Biol* 2004, **3:**Article3.

45. Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR.** *Nucleic Acids Res* 2001, **29:**e45.

46. Sue D, Fink D, Wiedmann M, Boor KJ: **s^B-dependent gene induction and expression in *Listeria monocytogenes* during osmotic and acid stress conditions simulating the intestinal environment.** *Microbiology (Read)* 2004, **150:**3843-3855.

47. Nightingale KK, Windham K, Martin KE, Yeung M, Wiedmann M: **Select *Listeria monocytogenes* subtypes commonly found in foods carry distinct nonsense mutations in inlA, leading to expression of truncated and secreted internalin A, and are associated with a reduced invasion phenotype for human intestinal epithelial cells.** *Appl Environ Microbiol* 2005, **71:**8764-8772.

48. Bakardjiev AI, Stacy BA, Fisher SJ, Portnoy DA: **Listeriosis in the Pregnant Guinea Pig: a Model of Vertical Transmission.** *Infect Immun* 2004, **72:**489-497.

49. Oliver HF, Orsi RH, Ponnala L, Keich U, Wang W, Sun Q, Cartinhour SW, Filiatrault MJ, Wiedmann M, Boor KJ: **Peering inside a killer's tool kit:**

mapping *L. monocytogenes* coding and noncoding RNAs. *BMC Genome* Submitted.

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

51. Lecuit M, Ohayon H, Braun L, Mengaud J, Cossart P: **Internalin of *Listeria monocytogenes* with an intact leucine-rich repeat region is sufficient to promote internalization.** *Infect Immun* 1997, **65:**5309-5319.

52. Chai Y, Kolter R, Losick R: **A widely conserved gene cluster required for lactate utilization in *Bacillus subtilis* and its involvement in biofilm formation.** *J Bacteriol* 2009, **191:**2423-2430.

53. Hillerich B, Westpheling J: **A new GntR family transcriptional regulator in *Streptomyces coelicolor* is required for morphogenesis and antibiotic production and controls transcription of an ABC transporter in response to carbon source.** *J Bacteriol* 2006, **188:**7477-7487.

54. Ogasawara H, Ishida Y, Yamada K, Yamamoto K, Ishihama A: **PdhR (pyruvatedehydrogenase complex regulator) controls the respiratory electron transport system in *Escherichia coli*.** *J Bacteriol* 2007, **189:**5534-5541.

55. Nigel L Brown JVSSPKJLH: **The MerR family of transcriptional regulators.** *FEMS Microbiol Rev* 2003, **27:**145-163.

56. Joseph B, Mertins S, Stoll R, Schar J, Umesha KR, Luo Q, Muller-Altrock S, Goebel W: **Glycerol metabolism and PrfA activity in *Listeria monocytogenes*.** *J Bacteriol* 2008, **190:**5412-5430.

57. Chan YC, Boor KJ, Wiedmann M: **s^B-dependent and s^B-independent mechanisms contribute to transcription of *Listeria monocytogenes* cold stress genes during cold shock and cold growth.** *Appl Environ Microbiol* 2007, **73:**6019-6029.

58. Cetin MS, Zhang C, Hutkins RW, Benson AK: **Regulation of transcription of compatible solute transporters by the general stress sigma factor, s^B, in *Listeria monocytogenes*.** *J Bacteriol* 2004, **186:**794-802.

59. Fraser KR, Sue D, Wiedmann M, Boor K, O'Byrne CP: **Role of s^B in regulating the compatible solute uptake systems of *Listeria monocytogenes*: osmotic induction of *opuC* is s^B dependent.** *Appl Environ Microbiol* 2003, **69:**2015-2022.

60. Shetron-Rama LM, Mueller K, Bravo JM, Bouwer HG, Way SS, Freitag NE: **Isolation of *Listeria monocytogenes* mutants with high-level in vitro expression of host cytosol-induced gene products.** *Mol Microbiol* 2003, **48:**1537-1551.

61. McGann P, Raengpradub S, Ivanek R, Wiedmann M, Boor KJ: **Differential regulation of *Listeria monocytogenes* internalin and internalin-like genes by s^B and PrfA as revealed by subgenomic microarray analyses.** *Food Path Dis* 2008, **5:**417-435.

62. Mueller KJ, Freitag NE: **Pleiotropic enhancement of bacterial pathogenesis resulting from the constitutive activation of the *Listeria monocytogenes* regulatory factor PrfA.** *Infect Immun* 2005, **73:**1917-1926.

63. Pané-Farré J, Jonas B, Förstner K, Engelmann S, Hecker M: **The σ^B regulon in *Staphylococcus aureus* and its regulation.** *Intl J Med Microbiol* 2006, **296:**237-258.

64. Hecker M, Reder A, Fuchs S, Pagels M, Engelmann S: **Physiological proteomics and stress/starvation responses in *Bacillus subtilis* and *Staphylococcus aureus*.** *Res Microbiol,* **In Press, Accepted Manuscript**.

65. van Schaik W, van der Voort M, Molenaar D, Moezelaar R, de Vos WM, Abee T: **Identification of the s^B regulon of *Bacillus cereus* and conservation of s^B-**

egulated genes in low-GC-content Gram-positive bacteria. *J Bacteriol* 2007, **189:**4384-4390.

66. Chaturongakul S, Raengpradub S, Wiedmann M, Boor KJ: **Modulation of stress and virulence in *Listeria monocytogenes*.** *Trends Microbiol* 2008, **16:**388-396.

67. Cebrián G, Sagarzazu N, Aertsen A, Pagán R, Condón S, Mañas P: **Role of the alternative sigma factor σ^B on *Staphylococcus aureus* resistance to stresses of relevance to food preservation.** *J App Microbiol* 2009, (in press).

68. van Schaik W, Tempelaars MH, Wouters JA, de Vos WM, Abee T: **The alternative sigma factor s^B of *Bacillus cereus*: response to stress and role in heat adaptation.** *J Bacteriol* 2004, **186:**316-325.

69. Holtmann G, Brigulla M, Steil L, Schutz A, Barnekow K, Volker U, Bremer E: **RsbV-independent induction of the SigB-dependent general stress regulon of *Bacillus subtilis* during growth at high temperature.** *J Bacteriol* 2004, **186:**6150-6158.

70. Volker U, Maul B, Hecker M: **Expression of the sigma B-dependent general stress regulon confers multiple stress resistance in *Bacillus subtilis*.** *J Bacteriol* 1999, **181:**3942-3948.

71. Notley-McRobb L, King T, Ferenci T: ***rpoS* mutations and loss of general stress resistance in *Escherichia coli* populations as a consequence of conflict between competing stress responses.** *J Bacteriol* 2002, **184:**806-811.

72. Jorgensen F, Leach S, Wilde SJ, Davies A, Stewart GSAB, Humphrey T: **Invasiveness in chickens, stress resistance and RpoS status of wild-type *Salmonella enterica* subsp. *enterica* serovar Typhimurium definitive type 104 and serovar Enteritidis phage type 4 strains.** *Microbiology (Reading, England)* 2000, **146:**3227-3235.

73. Fraser D, Kærn M: **A chance at survival: gene expression noise and phenotypic diversification strategies.** *Mol Microbiol* 2009, **71:**1333-1340.

CHAPTER THREE

Peering Inside a Killer's Tool Kit: Mapping *L. monocytogenes* Coding and Noncoding RNAs

ABSTRACT

Comprehensive, quantitative measurements of the transcriptional responses of bacterial pathogens under a variety of environmental conditions will identify specific genes and gene expression patterns important for bacterial survival, transmission and pathogenesis. The stationary phase stress response transcriptome of the human bacterial pathogen *Listeria monocytogenes* was defined using RNA sequencing (RNA-Seq) with the Illumina Genome Analyzer. Specifically, bacterial transcriptomes were compared between stationary phase cells of *L. monocytogenes* 10403S and an isogenic Δ*sigB* mutant, which does not express the alternative sigma factor σ^B, a major regulator of genes contributing to stress response. Overall, 83% of all genes were transcribed under these conditions. A total of 96 genes had significantly higher transcript levels in 10403S than in Δ*sigB*, indicating σ^B-dependent transcription of these genes. RNA-Seq analyses suggested 65 noncoding RNA molecules (ncRNAs) are transcribed in stationary phase *L. monocytogenes*, including (i) 15 previously unrecognized putative ncRNAs; one of which was σ^B-dependent, (ii) 38 ncRNAs resembling ncRNAs described in other bacteria, but not previously experimentally validated in *L. monocytogenes*, and (iii) 12 ncRNAs previously reported in *L. monocytogenes*. A dynamically trained Hidden Markov Model, in combination with RNA-Seq data, identified 65 putative σ^B promoters upstream of 82 of the 96 σ^B-dependent genes and one σ^B-dependent ncRNA. The RNA-Seq data also enabled annotation of putative operons and visualization of 5'- and 3'-UTR regions. These results provide compelling evidence that, in combination with bioinformatics tools, RNA-Seq allows quantitative characterization of prokaryotic transcriptomes, thus

providing exciting new strategies for exploring transcriptional regulatory networks in bacteria.

All RNA-Seq data has been submitted to the NCBI GEO Short Read Archives under accession number GSE15651. Four supplemental materials are provided including a comprehensive Access database containing RNA-Seq data, microarray, σ^B promoter and operon annotations, and binomial comparisons reported in this study. Further, a Genbank (gbk) file with ncRNAs identified in this study, a sequecibility file, and coverage file with the normalized RNA-Seq coverage for the 4 RNA-Seq runs are also provided.

INTRODUCTION

The development of powerful new DNA sequencing technologies has yielded new tools with the potential for dramatically revolutionizing scientific approaches to biological questions [1]. These new technologies can be used for a variety of applications, including genome sequencing, identification of DNA-methylation sites, population studies, chromatin precipitation (CHIP-Seq), and transcriptome studies (RNA-Seq). For RNA-Seq, cDNA is generated from an mRNA-enriched total RNA preparation and sequenced using high-throughput technology. Here, we used the Illumina Genome Analyzer to characterize the transcriptome of stationary phase *Listeria monocytogenes* 10403S and its isogenic Δ*sigB* mutant, which lacks the general stress response sigma factor, σ^B.

L. monocytogenes, a Gram-positive foodborne pathogen of the Firmicutes family, is the etiological agent of the disease known as listeriosis. As 20% of listeriosis cases result in death in humans, with an estimated annual human death toll of ~ 500 in the US alone [2], this disease is a considerable public health concern. As a foodborne pathogen (with 99% of human illnesses caused by foodborne infection [2]), this bacterium also presents challenging food safety concerns due to its ability to survive and grow under

many conditions that are typically applied to control bacterial populations in foods, such as low pH, low temperature and high salt conditions [3-5]. The alternative general stress response sigma factor, σ^B, is an essential component of a regulatory mechanism that contributes to the ability of *L. monocytogenes* to respond to and survive exposure to harsh environmental conditions [6].

Sigma factors are dissociable subunits of prokaryotic RNA polymerase responsible for enzyme recognition of a specific DNA sequence encoding a transcriptional promoter site. Promoter recognition specificities of bacterial RNA polymerase are determined by the transient association of an appropriate sigma factor with core polymerase in response to conditions affecting the cell [7]. The regulon of a single alternative sigma factor can include hundreds of transcriptional units, thus sigma factors provide an effective mechanism for simultaneously regulating large numbers of genes under appropriate conditions [7]. Critical phenotypic functions regulated by alternative sigma factors range from bacterial sporulation [8] to stress response systems [6, 9].

Through microarray analyses, the σ^B regulon in *L. monocytogenes* has been reported to encompass more than 200 genes, including both stress response and virulence genes [10]. However, interpretation of microarray analyses is dependent on the quality of existing genome annotations, which are rarely experimentally verified. Further, transcripts that do not correspond to annotated features (e.g., noncoding RNA transcripts) cannot be identified. In addition, the utility of microarrays is limited by the genomic variation that exists among bacterial strains (i.e., ideally, a unique microarray should be constructed for each strain to be analyzed) and by technical biases such as cross-hybridization. Hence, microarray data can be difficult to analyze and occasionally, misleading [11, 12]. Although interpretation of RNA-Seq data also relies on the availability of a genome sequence, it is probe- and annotation-independent and therefore, is free of cross-hybridization and low-hybridization biases, hence enabling genome-wide

identification of all transcripts, including small noncoding RNAs (ncRNAs). Moreover, because RNA-Seq technology can generate multiple reads corresponding to each transcribed nucleotide on the genome, it is usually possible to identify 5' and 3' transcript ends with high resolution [13]. Therefore, in combination with bioinformatics tools, RNA-Seq data can be used to identify transcriptional promoters and terminators. We used *L. monocytogenes* as a model system to explore application of RNA-Seq for the dual purposes of genome-wide transcriptome characterization in a bacterial pathogen and comprehensive quantification of target gene expression for the alternative sigma factor, σ^B.

RESULTS

RNA-Seq provided comprehensive coverage of the *L. monocytogenes* transcriptome

RNA-Seq analyses were performed on two independent replicate RNA samples collected from both the *L. monocytogenes* strain 10403S and an otherwise isogenic Δ*sigB* mutant (FSL A1-254) that had been grown to stationary phase. cDNA was generated from mRNA-enriched total RNA preparations from each strain and sequenced using the Illumina Genome Analyzer to yield a total number of reads for each sample ranging from 3,300,716 to 5,236,748 (Table 3.1). As the 10403S genome has not been completely closed, the sequence reads were aligned to a 10403S pseudochromosome that was created for this study using the completely closed genome of the *L. monocytogenes* strain EGD-e (accession no. AL591824) as a reference (see Material and Methods for details). The total number of reads matching regions other than rRNA and tRNA ranged from 451,548 to 683,746, yielding between 5X and 7.6X coverage of the pseudogenome. Between 87.3% and 92.1% of the reads in a given RNA-Seq run matched uniquely to the 10403S pseudochromosome and thus were used in subsequent analyses. Reads that did not match the 10403S pseudochromosome (i.e., reads that showed > 2 mismatches to the pseudochromosome) represented between 6.7% and 12.6% of the reads sequenced;

Table 3.1 Summary of RNA-Seq coverage data

Statistics	10403S replicate1	10403S replicate 2	Δ*sigB* replicate 1	Δ*sigB* replicate 2
Reads that aligned uniquely with no mismatches (U_0)	2,290,717	3,111,726	2,320,447	3,866,492
Reads that aligned uniquely with 1 mismatch (U_1)	632,173	470,865	544,932	745,360
Reads that aligned uniquely with 2 mismatches (U_2)	234,886	110,882	173,903	181,684
$U_{SUM} = U_0 + U_1 + U_2$	3,157,776	3,693,473	3,039,282	4,793,536
Reads that aligned at more than one location (reads not used; R)	23,485	4,832	38,489	16,103
Reads that did not align to the pseudochromosome (NM)	299,034	533,462	222,945	427,109
Total number of reads in the sample (Total = U_{SUM} + R +NM)	3,480,295	4,231,767	3,300,716	5,236,748
Percentage of unique alignments, i.e. 100*(U_{SUM})/Total	90.73	87.28	92.08	91.54
Reads that aligned to the 16S rRNA gene (16S)	490,381	482,845	434,263	760,863
Reads that aligned to the 23S rRNA gene (23S)	2,160,538	1,860,817	2436325	3,138,329
Reads that aligned to the 16S and 23S rRNA genes (16S + 23S)	2,650,919	2,919,170	2,295,080	3,899,192
Percentage of all reads that aligned to 16S and 23S rRNA genes	83.9	79	75.5	81.3
$U_{TOTAL} = U_{SUM} – (16S + 23S)$	506,857	774,303	744,202	894,344
Normalization factor ($f_{norm} = 894{,}344/U_{TOTAL}$)[a]	1.765	1.155	1.202	1

[a]This indicates the factor that was used for normalization of replicates

another 0.1% to 0.7% of the reads matched to at least two different locations on the pseudochromosome and, therefore, were removed before further analyses. Reads identified as "matching two locations" did not include those matching rRNA genes as the 10403S pseudochromosome created for this study was designed with only one unique rRNA gene sequence.

To allow for quantitative comparisons among genes and runs, the coverage for each run was normalized for the total number of reads in each run and for gene size. The normalized data are presented as the Gene Expression Index (GEI), which is expressed as the number of reads per 100 bases. Although *in silico* analyses suggested that the sequencibility (i.e., the portion of the pseudochromosome that could yield unique 32 nt reads) of the 10403S pseudochromosome was 99.6% (Supplementary Materials S1), approximately 77.5% of the genome was covered by reads from at least one of the four runs, suggesting that more than 20% of the genome is not transcribed or is transcribed at low levels.

RNA-Seq coverage correlated with qRT-PCR transcript levels indicating that RNA-Seq data are quantitative.

We evaluated whether average GEI for specific genes correlated with transcript levels that had been measured using TaqMan qRT-PCR, the current gold standard for quantification of mRNA [14]. Based on transcript levels for 9 and 5 genes in 10403S and Δ*sigB*, respectively, log transformed average GEI and log transformed TaqMan qRT-PCR absolute copy numbers were correlated (*p*-value < 0.001; adj. R^2=0.83; Figure 3.1; Table A2 [S3.1), supporting that RNA-Seq provides reliable quantitative estimates of transcript levels in *L. monocytogenes*. RNA-Seq was previously reported to provide quantitative data on transcript levels in yeast [15], and

Figure 3.1 Correlation between qRT-PCR and RNA-Seq. Correlation between qRT-PCR and RNA-Seq data for selected genes in *L. monocytogenes* 10403S (red) and the Δ*sigB* strain (blue). The selected genes are: *ctc*, *gadA*, *gap*, *opuCA*, *rpoB* (qRT-PCR data from both strains were available for these 5 genes), *flaA*, *inlA*, *plcA* and *sigB* (only qRT-PCR data from 10403S were available for these 4 genes).

Average qRT-PCR log-transformed absolute copy number
2
3
4
5
6
7
8
-1
0
1
2
3
4
y = 4.3444 + 1.1796*x
adjusted R2 = 0.83

more recently, in *Burkholderia cenocepacia* [15], thus, our findings extend this important correlation to a new prokaryotic system.

Stationary phase *L. monocytogenes* transcribed at least 83% of annotated genes

Among the 2888 annotated coding sequences (CDS) in the 10403S pseudochromosome, 2417 (83.7%) showed an average GEI ≥ 0.7 in 10403S (average of two biological replicates) suggesting that at least 83% of the annotated *L. monocytogenes* genes are transcribed in stationary phase (Figure A1 [S3.1]); see Materials and Methods for calculation of coverage, rational for defining transcribed genes, and criteria for classifying transcript levels as low, medium or high). Of these 2417 genes, 654 (22%) had high transcript levels, 586 (20.0%) had medium transcript levels, and 1177 (41.0%) had low transcript levels. A total of 471 genes (17%) had GEI < 0.7 and were considered "not transcribed". RNA-Seq data allowed visual examination of transcript units, aiding in identification of genes that are transcribed monocistronically or as part of an operon (Figure 3.2). A total of 355 transcription units appeared to represent operons; these units were identified and annotated (Supplemental Materials S2). A total of 1107 (38.3%) of the annotated 10403S CDS were located in these putative operons. Further experimental data are necessary to validate our predictions of transcription unit structure as some genes may have rho-dependent terminators that were not identified in this study and, therefore, they may be transcribed monocistronically despite the observation of GEI similar to those of their neighboring genes.

The three genes with the highest average GEI in 10403S all encoded predicted ncRNAs, including tmRNA, 6S and LhrA (Table 3.2). The annotated CDS (as annotated in EGD-e [16]) with the highest average GEI were lmo2257, *fri*, and lmo1847, which encode a hypothetical CDS, iron-binding ferritin, and an ABC

Figure 3.2 View of RNA-Seq data using the Artemis genome browser. This region of the 10403S chromosome includes six coding genes, i.e. LMRG_02429 to LMRG_02435, and the 5' end of LMRG_02436; genes are represented as blue arrows. The top part of the figure shows normalized RNA-Seq coverage (i.e. the number of reads that match an annotated gene after normalization across runs) with red and blue lines representing the two 10403S replicates and the green and black lines representing the Δ*sigB* strain. The horizontal line indicates a normalized RNA-Seq coverage of 49.16 reads. The middle part of the figure shows the three positive frames of translation with the coding regions and vertical black bars representing stop codons. The last line shows putative operons (white bars), a terminator (purple bar) downstream of LMRG_02430 and the chromosome coordinates. Notice the difference in coverage between LMRG_02431 (downstream of the terminator) and the other genes. All genes in the figure have sequencibility of 100% (See Supplemental Materials S1 for a complete sequencibility plot).

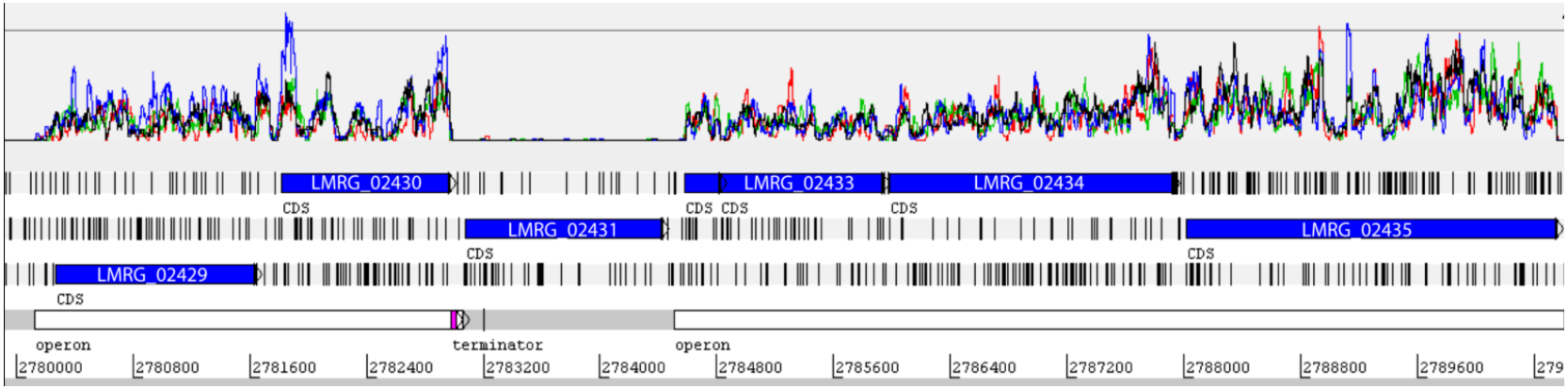

LMRG_02430
LMRG_02433
LMRG_02434
CDS
CDS CDS
CDS
LMRG_02431
LMRG_02435
CDS
CDS
LMRG_02429
CDS
operon
terminator
operon
2780000
2780800
2781600
2782400
2783200
2784000
2784800
2785600
2786400
2787200
2788000
2788800
2789600
279

Table 3.2 Genes with highest GEI

Locus	Gene name[a]	EGD-e locus[b]	Description	10403S Average GEI[c]
LMRG_04519	*ssrA*	NL	transfer-messenger RNA (tmRNA)	8566.2
LMRG_04503	*ssrS*	NL	6S RNA	7921.4
Noncoding	*lhrA*	NL	Hfq-binding RNA	4532.3
Noncoding	*sbrE*	NL	putative ncRNA	2359.9
LMRG_01574[d]	lmo2257	lmo2257	hypothetical CDS	2066.3
LMRG_02041	*fri*	lmo0943	non-heme iron-binding ferritin	1572.6
LMRG_04515	NGN	NL	bacterial signal recognition particle RNA	1462.2
LMRG_02926[e]	NGN	NL	-	1407.0
LMRG_00994	lmo1847	lmo1847	similar to adhesion binding proteins and lipoproteins with multiple specificity for metal cations (ABC transporter)	1378.9
LMRG_00378	*flaA*	lmo0690	flagellin protein	1366.9
LMRG_04523	*rnpB*	NL	bacterial RNAse P class B	1243.8

Table 3.2 (Continued)

Locus	Gene name[a]	EGD-e locus[b]	Description	10403S Average GEI[c]
LMRG_01165	*cspB*	lmo2016	similar to major cold-shock protein	1109.5
Noncoding	NGN	NL	T-box leader	1086.7
LMRG_00891	*sod*	lmo1439	superoxide dismutase	845.4
LMRG_00996	lmo1849	lmo1849	similar to metal cations ABC transporter, ATP-binding proteins	827.4
LMRG_01986	lmo2711	lmo2711	similar to hypothetical proteins	802.1
LMRG_00921	lmo1468	lmo1468	similar to unknown proteins	738.5
LMRG_02618	lmo0196	lmo0196	similar to *B. subtilis* SpoVG protein	702.9
LMRG_00814	*cspL*	lmo1364	similar to cold shock protein	679.4
LMRG_01081	*hup*	lmo1934	similar to non-specific DNA-binding protein HU	631.8
LMRG_00995	lmo1848	lmo1848	similar metal cations ABC transporter (permease protein)	621.2
LMRG_00922	*rpsU*	lmo1469	30S ribosomal protein S21	609.0
LMRG_02619	lmo0197	lmo0197	similar to *B. subtilis* SpoVG protein	577.3

Table 3.2 (Continued)

Locus	Gene name[a]	EGD-e locus[b]	Description	10403S Average GEI[c]
Noncoding	NGN	NL	putative ncRNA	561.9
LMRG_00679	*trxA*	lmo1233	thioredoxin	516.5
LMRG_01674	lmo2158	lmo2158	similar to *B. subtilis* YwmG protein	509.2
LMRG_02633	*ctc*	lmo0211	similar to *B. subtilis* general stress protein	496.4
LMRG_01479	lmo2363	lmo2363	similar to glutamate decarboxylase	491.0
LMRG_00517	*pdhD*	lmo1055	highly similar to dihydrolipoamide dehydrogenase, E3 subunit of pyruvate dehydrogenase complex	483.5
LMRG_00703	lmo1254	lmo1254	similar to alpha,alpha-phosphotrehalase	395.9
LMRG_02718	lmo2373	lmo2373	similar to phosphotransferase system beta-glucoside-specific enzyme IIB component	378.5
LMRG_01737	lmo2511	lmo2511	similar to *B. subtilis* YvyD protein	377.1
LMRG_00515	*pdhB*	lmo1053	highly similar to pyruvate dehydrogenase (E1 beta subunit)	356.4

Table 3.2 (Continued)

Locus	Gene name[a]	EGD-e locus[b]	Description	10403S Average GEI[c]
LMRG_00704	lmo1255	lmo1255	similar to PTS system trehalose-specific enzyme IIBC	353.6
LMRG_00516	*pdhC*	lmo1054	highly similar to pyruvate dehydrogenase (dihydrolipoamide acetyltransferase E2 subunit)	351.3
LMRG_01480	lmo2362	lmo2362	similar to amino acid antiporter (acid resistance)	351.1
LMRG_02239	lmo2692	lmo2692	unknown	344.1
LMRG_00875	lmo1423	lmo1423	unknown	341.2
LMRG_01835	lmo2413	lmo2413	similar to aminotransferase	333.1
LMRG_01429	lmo1541	lmo1541	similar to unknown protein	318.8

[a]NGN=No gene name given;

[b]NL=No EGDe locus;

[c]Average normalized number of reads matching each of the σ^B-dependent genes in the two 10403S datasets divided by the length of the genes times 100 bp;

[d]The high coverage of LMRG_01574 is restricted to the portion that overlaps with *lhrA*. LMRG_01574 may not be a valid coding gene;

[e]LMRG_02926 completely overlaps with the bacterial RNAse P class B noncoding gene. LMRG_02926 may not be a valid coding gene as no Pfam matches were found for the putative protein coded by this gene.

transporter, respectively. Other genes with well defined functions and high average GEI include *flaA*, which encodes a flagellin protein, *sod*, which encodes a superoxide dismutase involved in detoxification, and *cspB* and *cspL*, which encode cold-shock proteins involved in adaptation to atypical conditions (Table 3.2).

Both positive and negative associations were observed between GEI and the TIGR classification of sets of genes to physiological role categories (http://cmr.jcvi.org/cgi-bin/CMR/RoleIds.cgi) (Table 3.3). For example, genes involved in protein synthesis and protein fate showed higher average GEI in stationary phase 10403S as compared to genes involved in other functions, while genes involved in viral functions and amino acid biosynthesis were significantly associated with low average GEI in 10403S. Moreover, a positive significant association was observed between codon bias and the average GEI in 10403S (p-value < 0.001; linear regression analysis).

Identification and annotation of noncoding RNAs (ncRNAs)

Overall, we identified 65 ncRNAs (Table A2 [S3.2]) that showed average GEI ≥ 0.7 in 10403S, indicating that these ncRNAs are transcribed in stationary phase *L. monocytogenes* (see Materials and Methods for more details on ncRNA annotation). Five other ncRNA previously identified (i.e., RliA, RilE, RilF, RliG, SbrC [17,18]) showed low (< 0.7) GEI, ranging from 0 to 0.39, and thus appear not to be transcribed in stationary phase *L. monocytogenes* 10403S. Among the 65 ncRNAs identified as transcribed in the present study, 12 matched ncRNAs previously described in *L. monocytogenes* (Table A2 [S3.2]) [17,18,19]. Among the 53 transcribed ncRNAs not previously described in the literature (Table 3.4), 38 represented ncRNAs predicted by Rfam [20] in the EGD-e genome. These 38 ncRNAs included 6S RNA, tmRNA, several S-box RNA and T-box leader RNA molecules. A total of 15 putative ncRNAs

Table 3.3 Associations between GEI and role categories

	Role categories	Significance[a]
Low average GEI in 10403S	Signal transduction	0.006
	Amino acid biosynthesis	< 0.001
	Transport and binding	0.003
	Viral function	< 0.001
High average GEI in 10403S	Cellular processes	0.011
	DNA metabolism	0.011
	Protein fate	< 0.001
	Protein synthesis	< 0.001
	Purines, pyrimidines, nucleosides, and nucleotides	0.043
	Transcription	< 0.001
	Unknown functions	0.043

[a] Based on one-sided Wilcoxon rank sum test and FDR correction.

Table 3.4 New *L. monocytogenes* ncRNAs identified in this study

Description	Coordinates in 10403S	Length	10403S Average GEI[a]	Δ*sigB*A Average GEI[b]
Putative ncRNAs newly identified in this study				
putative ncRNA	161945..162111	167	32.7	34.1
putative ncRNA	222952..223741	790	1.99	2.17
putative ncRNA	409956..410100	145	43.8	82.82
putative ncRNA	419482..419602	121	269.2	306.99
putative ncRNA	477023..477185	163	7.46	5.46
putative ncRNA	479838..479991	154	56.02	67.52
putative ncRNA	836741..836942	202	15.52	11.34
putative ncRNA	938236..938563	328	14.47	29.94
putative ncRNA	1257547..1257724	178	20.56	23.43
putative ncRNA	1393256..1393496	241	52.11	65.68
putative ncRNA	1884385..1884664	280	25.2	45.98
putative ncRNA	2020305..2020575	271	189.49	224.23
putative ncRNA, *sbrE*	2072821..2073334	514	2359.89	20.95
putative ncRNA	2305436..2305610	175	20.62	49.18

Table 3.4 (Continued)

Description	Coordinates in 10403S	Length	10403S Average GEI[a]	Δ*sigB*A Average GEI[b]
putative ncRNA	2370319..2370547	229	45.73	17.84
ncRNAs in the Rfam database				
putative L10 leader	159701..159845	145	81.31	99.33
putative SAM riboswitch (S-box leader)	204783..204972	190	18.14	61.65
putative TPP riboswitch (THI element)	240868..241057	190	14.06	10.34
putative purine riboswitch	490215..490347	133	650.65	629.28
putative SAM riboswitch (S-box leader)	516988..517156	169	1.89	3.79
putative glucosamine-6-phosphate activated ribozyme	637782..638097	316	51.78	77.18
putative lysine riboswitch	707866..708136	271	57.98	59.25
putative SAM riboswitch (S-box leader)	762904..763066	163	103.57	189.42
putative PreQ1-I riboswitch	788075..788122	48	5.44	7.35
putative yybP-ykoY leader	902340..902520	181	11.41	16.88
putative cobalamin riboswitch	1037938..1038128	191	3.94	1.99
putative cobalamin riboswitch	1074606..1074806	201	2.96	3.97
putative glycine riboswitch	1230912..1231051	140	49.98	29.38
putative TPP riboswitch (THI element)	1319193..1319376	184	20.31	58.86
putative T-box leader	1352163..1352373	211	304.66	296.84

Table 3.4 (Continued)

Description	Coordinates in 10403S	Length	10403S Average GEI[a]	Δ*sigB*A Average GEI[b]
putative T-box leader	1412037..1412289	253	96.34	162.32
putative L21 leader	1435176..1435232	57	252.99	217.91
putative T-box leader	1447016..1447277	262	63.12	69.73
putative T-box leader	1455337..1455592	256	153.6	175.46
putative T-box leader	1500724..1500985	262	118.16	150.49
putative T-box leader	1534507..1534760	254	14.44	35.41
putative T-box leader	1534782..1535053	272	1086.72	1128.9
putative T-box leader	1569120..1569357	238	37.22	37.52
putative SAM riboswitch (S-box leader)	1574285..1574471	187	41.49	86.25
putative SAM riboswitch (S-box leader)	1597226..1597439	214	183.17	264.67
putative T-box leader	1660200..1660479	280	223.69	318.23
putative L19 leader	1707678..1707737	60	87.43	70.5
putative PyrR binding site element	1762762..1762865	104	1.67	1.54
putative PyrR binding site	1763445..1763553	109	5.12	3.11
putative purine riboswitch	1804062..1804230	169	18.56	33.91
putative FMN riboswitch (RFN element)	1865665..1865923	259	72.68	204.4
putative T-box leader	2134253..2134523	271	77.11	84.01

Table 3.4 (Continued)

Description	Coordinates in 10403S	Length	10403S Average GEI[a]	Δ*sigB*A Average GEI[b]
putative SAM riboswitch (S-box leader)	2327827..2328052	226	37.34	34.81
putative T-box leader	2505646..2505928	283	7.38	5.73
putative L13 leader	2524976..2525028	53	26.09	44.14
putative ykoK leader (M-box)	2605679..2605991	313	59.33	70.78
putative bacterial signal recognition particle (SRP)	2623466..2623799	334	1462.2	1295.75
putative T-box leader	2662484..2662739	256	93.57	70.33

[a]Average normalized number of reads matching each of the σ^B-dependent genes in the two 10403S datasets divided by the length of the genes times 100 bp;

[b]Average normalized number of reads matching each of the σ^B-dependent genes in the two Δ*sigB* datasets divided by the length of the genes times 100 bp.

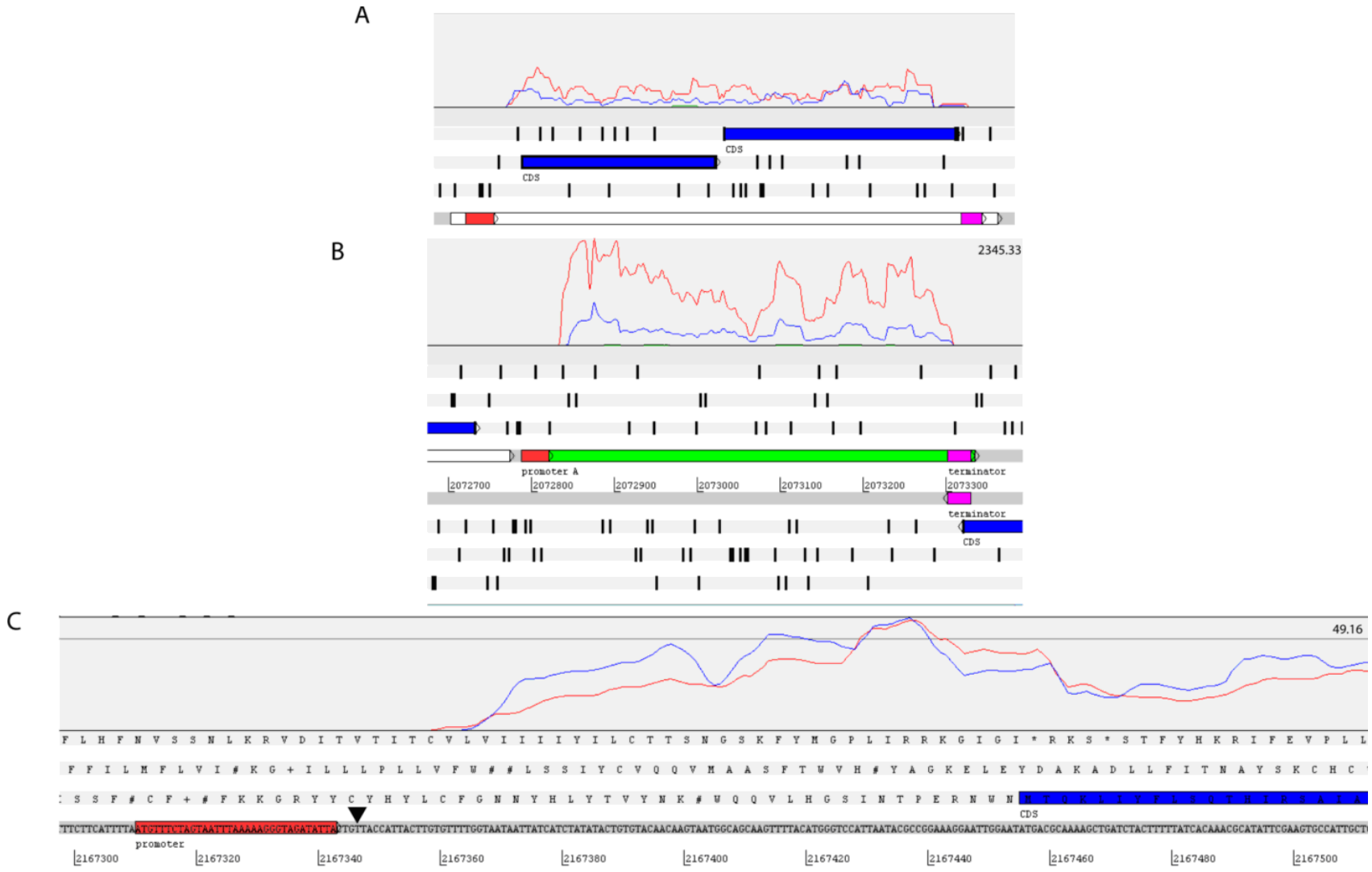
identified here were not previously identified in *L. monocytogenes* and did not match ncRNA entries in Rfam (Table 3.4). These regions showed contiguous coverage by RNA-Seq reads (i.e., at least 100 bp completely covered by RNA-Seq reads), but did not fully match annotated genes.

Three putative ncRNAs with high GEI covered either part or all of each of three annotated CDS, suggesting that ncRNAs overlap with these CDS or that some putative CDS actually encode ncRNAs rather than proteins. Specifically, LMRG_01574 (lmo2257), LMRG_02926 (no homolog in EGD-e), and LMRG_1986 (lmo2711) overlapped with *lhrA* (partial overlap), with the bacterial RNAse P class B ncRNA (full overlap), and with the bacterial signal recognition particle RNA (partial overlap), respectively. In concert with our findings, lmo2257 was previously hypothesized not to be a CDS [19].

RNA-Seq identified 96 annotated CDS and one novel ncRNA as σ^B-dependent and provided comprehensive data on transcript levels of genes in the σ^B regulon

Our RNA-Seq data analyses identified a total of 96 genes as up-regulated by σ^B (Table A2 [3.3]). No annotated genes were identified as significantly down-regulated by σ^B in this study. As illustrated in Figure 3.3A, RNA-Seq data are useful for predicting multi-gene operons controlled by a given regulator such as σ^B. Thirty-eight of the 96 up-regulated genes are organized into a total of 20 operons, including (i) *opuCABCD*, which encodes the subunits of a glycine betaine/carnitine/choline ABC transporter, (ii) lmo0781-lmo0784, which encode the four subunits of a putative mannose-specific phosphotransferase system, (iii) lmo2484-lmo2485, which encode a putative membrane-associated protein and a putative transcriptional regulator similar to PspC, respectively, and (iv) lmo0133 and lmo0134 (Figure 3.3A), which encode proteins similar to *E. coli* YjdI and YjdJ, respectively.

Figure 3.3 Examples of σ^B-dependent transcripts identified by RNA-Seq. In each panel (A, B, and C), red and blue lines representing normalized RNA-Seq coverage (i.e. the number of reads that match an annotated gene after normalization across runs) in the two 10403S replicates and green and black lines represent normalized RNA-Seq coverage in the Δ*sigB* strain replicates; the numbers at the top right in each panel indicates the normalized RNA-Seq coverage represented by the horizontal line shown. Panel (A) depicts LMRG_02382 and LMRG_02383 (shown as blue bars), which form an operon (indicated by a long white bar) with a defined Rho-independent terminator (purple bar) downstream of LMRG_02383; the three positive frames of translation with the coding regions in blue and stop codons shown as vertical black bars are also shown. A σ^B-dependent promoter (red bar) was identified upstream of the operon and the RNA-Seq coverage data clearly shows that the transcription of this operon is positively regulated by σ^B (i.e. almost no coverage was obtained from the Δ*sigB* strain). Panel (B) depicts a putative σ^B-dependent noncoding RNA (ncRNA) with Rho-independent terminator and a σ^B-dependent promoter identified; annotated features as well as positive and negative frames of translation are shown at the bottom with stop codons shown as vertical black bars. Panel (C) shows the 5' end of LMRG_01602 illustrating the position of a σ^B-dependent promoter in relation to the start codon of the gene and the transcriptional start site determined by RNA-Seq. The black triangle indicates the transcriptional start site determined by RACE-PCR as previously described by Kazmierczak et al. [21].


A
CDS
CDS
B
2345.33
promoter A
terminator
2072700
2072800
2072900
2073000
2073100
2073200
2073300
terminator
CDS
C
49.16
CDS
promoter
2167300
2167320
2167340
2167360
2167380
2167400
2167420
2167440
2167460
2167480
2167500

A one-sided Fisher's exact test was used to determine if σ^B-dependent genes are over-represented within specific TIGR role categories. Genes identified as σ^B-dependent were over-represented among genes involved in cellular functions (*q*-value = 0.045, one-sided Fisher's exact test). σ^B-dependent genes in this category include genes involved in pathogenesis (*inlA, inlB, inlH)*, adaptation to atypical conditions (lmo0515, lmo0669, lmo2673, *lrtC*), detoxification (lmo1433, lmo2230), cell division (lmo1624) and an unknown protein that may be involved in toxin production and resistance (lmo0321).

We evaluated RNA-Seq transcript levels for the 96 σ^B-dependent genes identified here (Table A2 [S3.3]). The average fold change (10403S GEI / Δ*sigB* GEI) for the 96 σ^B-dependent genes ranged from 2.6 to 479.4. The σ^B-dependent genes with the highest average GEI in 10403S were lmo2158, lmo1602, and lmo0539, which encode a protein similar to *B. subtilis* YwmG, an unknown protein, and a tagatose-1,6-diphosphate aldolase, respectively (Table 3.5).

An ~ 500 nt σ^B-dependent putative ncRNA was identified between lmo2141 and lmo2142 (3.3B); this ncRNA was designated as *sbrE* (**s**igma **B**-dependent **R**NA). Although BLAST searches and searches against the Pfam database using 6 possible reading frames did not yield significant matches, a σ^B-dependent promoter was identified upstream of the transcript and a Rho-independent terminator was found by TransTermHP (Figure 3.3B). The sequence for this putative ncRNA was also present in 17 other *L. monocytogenes* genomes, including EGD-e (GenBank accession no. NC 003210), F2365 (GenBank accession no. NC 002973), and 15 unfinished genome sequences by the Broad Institute (http://www.broad.mit.edu/annotation/genome/listeria_group/MultiHome.html) as well as in one *L. innocua* (GenBank accession no. NC 003212) and one *L. welshimeri*

Table 3.5 Summary of genes up-regulated by σ^B

Locus	EGD-e locus	Description	Avg. fold change (WT/D*sigB*)[a]	10403S Average GEI[b]	Δ*sigB* Average GEI[c]
σ^B-dependent genes found by RNA-Seq and not previously identified by microarray analyses of stationary phase cells					
LMRG_02371	lmo0122	similar to phage proteins	3.9	2.37	0.6
LMRG_02611	lmo0265	similar to succinyldiaminopimelate desuccinylase	204.5	17.95	0
LMRG_02602	lmo0274	unknown	3.17	2.89	0.91
LMRG_00064	lmo0372	similar to beta-glucosidase	4.26	2.4	0.66
LMRG_00126[d]	lmo0433 (*inlA*)	Internalin A	5.86	6.19	1.06
LMRG_00127[d]	lmo0434 (*inlB*)	Internalin B	6	2.71	0.47
LMRG_02244	lmo0819	unknown	3.01	18.35	6.09
LMRG_00873[d]	lmo1421	similar to glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)	28.44	5.27	0.67
LMRG_00877[d]	lmo1425 (*opuCD*)	similar to betaine/carnitine/choline ABC transporter (membrane p)	3.56	22.59	6.51
LMRG_00878[d]	lmo1426 (*opuCC*)	similar to glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein)	3.77	19.78	5.41
LMRG_01013	lmo1866	similar to conserved hypothetical proteins	2.63	4.87	1.79
LMRG_01151	lmo2003	similar to transcription regulator GntR family	14.67	3.15	0.32
LMRG_01963	lmo2733	similar to PTS system, fructose-specific IIABC component	7.95	1.35	0.32
Noncoding	ND	putative ncRNA, *sbrE*	186.09	2359.89	20.95

Table 3.5 (Continued)

Locus	EGD-e locus	Description	Avg. fold change (WT/D*sigB*)[a]	10403S Average GEI[b]	Δ*sigB* Average GEI[c]
σ^B-dependent genes with Average GEI ≥ 25 in 10403S					
Noncoding	ND	putative ncRNA, *sbrE*	186.09	2359.89	20.95
LMRG_01674	lmo2158	similar to *B. subtilis* YwmG protein	479.39	509.23	22.8
LMRG_01365	lmo1602	similar to unknown proteins	5.47	157.02	30.08
LMRG_00221	lmo0539	similar to tagatose-1,6-diphosphate aldolase	14.54	132.74	9.3
LMRG_01602	lmo2230	similar to arsenate reductase	411	96.43	0
LMRG_02052	lmo0953	unknown	167	73.18	0.48
LMRG_00357	lmo0669	similar to oxidoreductase	75.93	64.6	0.89
LMRG_00358	lmo0670	unknown	105.5	59.6	0.58
LMRG_00341	lmo0654	unknown	7.1	56.61	7.94
LMRG_02219	lmo2674	similar to ribose 5-phosphate epimerase	5.42	52.93	9.94
LMRG_01794	lmo2454	unknown	84.5	50.24	0.76
LMRG_01850	lmo2398 (*ltrC*)	low temperature requirement C protein, also similar to *B. subtilis* YutG protein	2.8	50.03	18.94
LMRG_00745	lmo1295(*hfq*)	similar to host factor-1 protein	4.83	49.77	11.19
LMRG_01948	lmo2748	similar to *B. subtilis* stress protein YdaG	207.5	49.37	0

Table 3.5 (Continued)

Locus	EGD-e locus	Description	Avg. fold change (WT/D*sigB*)[a]	10403S Average GEI[b]	Δ*sigB* Average GEI[c]
LMRG_00583	lmo1140	unknown	11.93	47.84	4.28
LMRG_02036	lmo0937	unknown	54.38	44.68	0.91
LMRG_00484	lmo0796	conserved hypothetical protein	4.21	43.88	10.61
LMRG_02772	lmo1698	similar to ribosomal-protein-alanine N-acetyltransferase	4.1	42.94	10.92
LMRG_02736	lmo2391	conserved hypothetical protein similar to *B. subtilis* YhfK protein	11.76	39.48	4.54
LMRG_02011	lmo0911	unknown	4.04	33.9	8.58
LMRG_01763	lmo2485	similar to *B. subtilis* yvlC protein	3.93	32.87	8.47
LMRG_00482	lmo0794	similar to B. subtilis YwnB protein	67.02	32.5	0.72
LMRG_00278	lmo0596	similar to unknown proteins	170.5	32.33	0.09
LMRG_02218	lmo2673	conserved hypothetical protein	150.5	31.92	0.11
LMRG_02013	lmo0913	similar to succinate semialdehyde dehydrogenase	330.38	30.05	0.11
LMRG_00469	lmo0781	similar to mannose-specific phosphotransferase system (PTS) component IID	59.58	29.59	0.65
LMRG_00470	lmo0782	similar to mannose-specific phosphotransferase system (PTS) component IIC	18.99	29.59	1.58
LMRG_01360	lmo1606	similar to DNA translocase	7.88	29.5	3.97

Table 3.5 (Continued)

Locus	EGD-e locus	Description	Avg. fold change (WT/D*sigB*)[a]	10403S Average GEI[b]	Δ*sigB* Average GEI[c]
LMRG_02696	lmo2572	similar to Chain A, Dihydrofolate Reductase	8.05	29.05	3.59
LMRG_02768	lmo1694	similar to CDP-abequose synthase	155.31	27.51	0.2
LMRG_02216	lmo2671	unknown	3.13	27.29	8.82
LMRG_02695	lmo2573	similar to zinc-binding dehydrogenase	7.52	25.91	3.83
LMRG_00472	lmo0784	similar to mannose-specific phosphotransferase system (PTS) component IIA	88.5	25.25	0.21
LMRG_02215	lmo2670	conserved hypothetical protein	3	25.23	8.58
LMRG_02697	lmo2571	similar to nicotinamidase	9.84	25.15	2.99

[a]Average fold changes from the 10403S and Δ*sigB.* Genes with no matching reads in Δ*sigB* had their GEI set to 1 to allow for calculation of the fold change;
[b]Average normalized number of reads matching each of the σ^B-dependent genes in the two 10403S datasets relative to the length of the genes times 100 bp;
[c]Average normalized number of reads matching each of the σ^B-dependent genes in the two Δ*sigB* datasets relative to the length of the genes times 100 bp.
[d]Genes previously identified as σ^B-dependent under salt stress in *L. monocytogenes*10403S by Raengpradub et al., 2008.

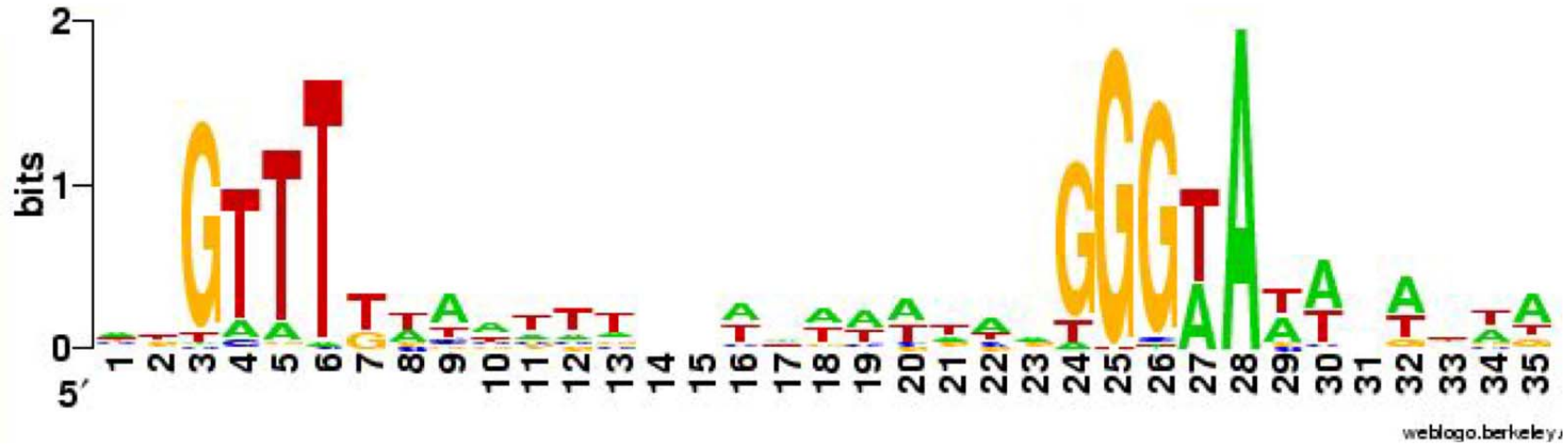
(GenBank accession no. NC 008555) genome. The 514 nt *sbrE* sequence was 96.6 % conserved among the 18 *L. monocytogenes* genomes.

HMM showed that 84% of σ^B-dependent genes and operons identified by RNA-Seq are preceded by σ^B promoters and therefore, appear to be directly regulated by σ^B

An HMM representing *L. monocytogenes* σ^B-dependent promoters was dynamically created by using an initial training set of experimentally verified *L. monocytogenes* σ^B-dependent promoters to search the RNA-Seq data. The final model yielded a total of 5387 motifs with scores > 5.00 bits throughout the pseudochromosome sequence. Among these motifs, we identified 65 possible σ^B-dependent promoter sequences upstream of genes and operons identified as σ^B-dependent based on RNA-Seq data (see Figure 3.4 for the *L. monocytogenes* σ^B promoter sequence logo). Because some of the genes with experimentally validated σ^B promoters were not found to be significantly up-regulated by σ^B in our study (e.g. *prfA* and the *rsbV* operon) and because the *ltrC* promoter, which was in the initial training set, had a score below our threshold of 5.00 bits in the final search, our annotation does not include all promoters present in the training set (i.e., only promoters identified upstream of genes that were significantly up-regulated by σ^B in the present study were annotated). Specifically, σ^B-dependent promoter sequences were found upstream of 15 of the 20 putative σ^B-dependent operons, 49 of the 58 monocistronic σ^B-dependent genes, and the one σ^B-dependent ncRNA identified here (Figure 3.3B). We compared RNA-Seq defined transcriptional start sites for 8 genes with σ^B promoters to transcriptional start sites determined by Rapid Amplification of cDNA Ends PCR (RACE-PCR) in a previous study [21]. Transcriptional start sites identified with RNA-Seq were located between 0 to 29 bases down-stream of start

Figure 3.4. Logo of the σ^B promoter. This logo was created from the alignment of 65 σ^B promoters identified in this study.

.



sites determined by RACE-PCR (see Figure 3.3C for LMRG_01602 transcriptional start site mapped by RACE-PCR and RNA-Seq), indicating that RNA-Seq successfully approximates transcriptional start sites. Some transcriptional start sites could not be specifically mapped to a σ^B promoter site using RNA-Seq as some genes (e.g. *opuCA*) have multiple promoters. A dendrogram of the putative σ^B promoter sequences showed no apparent clustering of these promoter sequences by either average GEI in 10403S or by σ^B-dependence (average fold change). These results suggest that additional regulatory elements or mechanisms other than promoter sequence *per se* (e.g., RNA stability) also influence transcript levels and/or σ^B-dependence for these genes (data not shown).

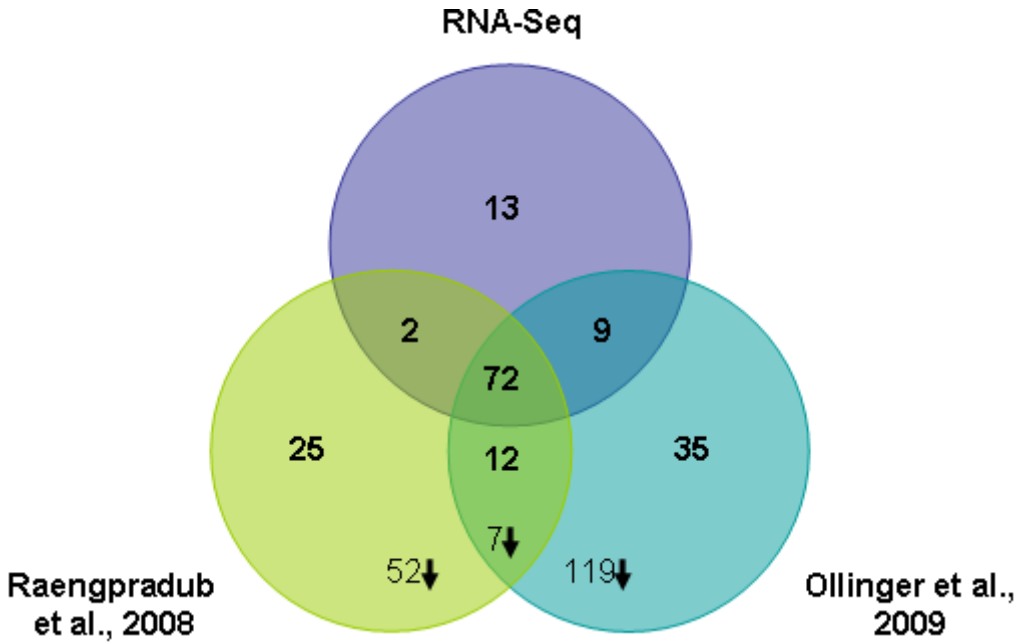
RNA-Seq successfully identifies a number of previously identified as well as novel σ^B-dependent genes.

To evaluate the ability of RNA-Seq to identify *L. monocytogenes* σ^B-dependent genes, we compared the σ^B-dependent genes identified here with those identified in two independent microarray studies by our research group. Specifically, we compared our results with microarray data reported by (i) Raengpradub et al. [10], who identified σ^B-dependent genes using *L. monocytogenes* strains and growth conditions identical to those in this study, and by (ii) Ollinger et al. [22], who identified σ^B-dependent genes by comparing transcripts from *L. monocytogenes* 10403S with a PrfA* (G155S) allele [23], which constitutively expresses the PrfA-regulated virulence genes [23,24,25], with those from an isogenic Δ*sigB* mutant grown to stationary phase under the same conditions. Further, we compared our results with those from a microarray study using another *L. monocytogenes* strain (EGD-e) and its isogenic Δ*sigB* mutant, grown under similar conditions (i.e., growth to early stationary phase [26]). Among the 96 σ^B-dependent annotated CDS identified in the present study, 72 were also identified as

σ^B-dependent in previous microarray studies of stationary phase *L. monocytogenes* 10403S cells [10,22] (Figure 3.5). In addition, 64 (66.7%) of the 96 σ^B-dependent genes identified here were identified as positively regulated by σ^B in *L. monocytogenes* strain EGD-e cells grown to early stationary phase (8 h growth in BHI) [26]. Overall, 12 genes identified as σ^B-dependent in both previous microarray studies by our group [10,22], were not identified as σ^B-dependent by the RNA-Seq experiments reported here (Figure 3.5); 9 of these genes showed a σ^B-dependent promoter based on the HMM analyses in this study and are likely to be directly regulated by σ^B (see Table A2 [S3.4] for further details on these genes).

Finally, a total of 13 annotated CDS identified as σ^B-dependent by RNA-Seq (including 9 genes that also showed a σ^B-dependent promoter in our HMM analysis) had not been identified as σ^B-dependent in either of the previous microarray studies with strain 10403S grown to stationary phase [10,22]. Among these 13 genes not previously identified as σ^B-dependent in stationary phase *L. monocytoegnes* 10403S, five had previously been identified as σ^B-dependent in salt-stressed cells [10]. In addition, two of these 13 genes had been identified as positively regulated by σ^B in *L. monocytogenes* strain EGD-e [26]. For one gene (i.e. lmo0265) identified as σ^B-dependent by RNA-Seq, but not in previous microarray studies of strain 10403S [10,22], the microarray probe (designed based on the genome of *L. monocytogenes* strain EGD-e) showed a low hybridization index (HI; % match between strain-specific sequence and oligonucleotide probe) to 10403S (< 80%). Further supporting the hypothesis that low HIs led to false negatives in our previous microarray studies with strain 10403S, Hain et al. [26] found that lmo0265 was σ^B-dependent in their microarray study using EGD-e. Interestingly, lmo2003, which encodes a transcription regulator similar to the GntR family, was identified as σ^B-dependent by RNA-Seq, but had not been previously identified as σ^B-dependent.

Figure 3.5 σ^B-dependent genes identified by RNA-Seq and microarray analyses. Venn diagram of σ^B-dependent genes identified in stationary phase cells in this study and in previous microarray studies of stationary phase *L. monocytogenes* [10, 22]. Numbers in bold are the number of up-regulated annotated CDS identified as σ^B-dependent in each study; numbers followed by down arrows are down-regulated σ^B-dependent genes. No down-regulated σ^B-dependent genes were identified by RNA-Seq. The 13 genes identified as σ^B-dependent in stationary phase only by RNA-Seq, but not by previous microarray studies of *L. monocytogenes* 10403S, include 5 genes that had been found to be σ^B-dependent, by microarray studies [10] in salt stressed cells (see Table 3.5). In a number of instances, (e.g. *opuCB*, *rsbX*; See Table A2 [S3.4]) genes with significantly different transcript levels in both microarrays [10, 22] had significant binomial probabilities ($q < 0.05$) and a fold change ≥ 2.0 for most of the possible combinations (i.e. 10403S replicate 1 vs Δ*sigB* replicate 1; 10403S replicate 1 vs Δ*sigB* replicate 2; 10403S replicate 2 vs Δ*sigB* replicate 1; 10403S replicate 2 vs Δ*sigB* replicate 2), but not for all four comparisons and these genes were, therefore, not identified as showing significant differences in normalized RNA-Seq coverage (based on our conservative definition of genes with significant differences in normalized RNA-Seq coverage); see Supplemental Table S4 for detailed RNA-Seq data for genes identified as σ^B-dependent by microarrays, but not by RNA-Seq.


RNA-Seq
13
2
9
72
25
12
35
7
52
119
Raengpradub et al., 2008
Ollinger et al., 2009

DISCUSSION

In this study, we used RNA sequencing to define and characterize the transcriptomes of *L. monocytogenes* strain 10403S and an otherwise isogenic Δ*sigB* mutant, which does not express the general stress-response sigma factor, σ^B. The data generated using this approach showed that (i) at least 83% of annotated *L. monocytogenes* genes are transcribed in stationary phase cells; and (ii) stationary phase *L. monocytogenes* transcribes 65 ncRNAs, including 52 ncRNA that, to our knowledge, have not previously been experimentally validated in *L. monocytogenes* and one newly identified σ^B-dependent ncRNA. Additionally, use of a novel, iterative, dynamic HMM in combination with RNA-Seq data allowed us to independently validate the *L. monocytogenes* σ^B regulon, quantify transcript levels for all σ^B-dependent genes, and identify putative σ^B-dependent promoters and the approximate location of transcriptional start sites on a genome scale.

The majority of annotated *L. monocytogenes* genes are transcribed in stationary phase cells.

While genome sequencing and microarray approaches have provided important insight into the biology of prokaryotic organisms, including a number of human bacterial pathogens, identification of all genes and their transcriptional patterns remains a major challenge in all areas of biology. Our results support the notion that global probe-independent approaches for transcriptome characterization are valuable tools for analyzing bacterial transcriptomes. RNA-Seq has recently been used to characterize two yeast [13,27] and two bacterial transcriptomes ([15,28]). Mao et al. [28] used the GS FLX system (Roche and 454 Life Sciences) to identify novel genes in *Sinorhizobium meliloti*, but the low coverage of transcripts under the conditions used resulted in limited quantitative data. More recently, Yoder-Himes et al. [15] used the

Illumina Genome Analyzer for comparative transcriptomics of *B. cenocepacia* grown under conditions mimicking pathogen and endemic soil environments to identify transcripts unique to each environment as well as novel ncRNAs.

A major challenge that currently hinders analysis of transcriptomic data generated by approaches such as RNA-Seq is the ability to differentiate between genes with low levels of transcription and background levels of coverage. Several approaches have been used to define cut-off values between background GEI and GEI indicative of low transcript levels (e.g., [13,27,29]). We chose a comparative analysis of *L. monocytogenes* 10403S transcript levels with those of a mutant strain that does not express a transcription factor (i.e., the alternative sigma factor σ^B) as a novel approach for robustly defining background RNA-Seq coverage. Our results show that a number of σ^B-dependent genes were solely σ^B-dependent (at least under the conditions used here), as supported by the lack of detectable RNA-Seq coverage in the Δ*sigB* strain, despite considerable RNA-Seq coverage of the same genes in the isogenic parent strain 10403S. This is an important observation as a number of σ^B-dependent *L. monocytogenes* genes are also activated by other sigma factors (e.g., σ^A [30,31]). Using the average GEI for *L. monocytogenes* genes that were solely σ^B-dependent in the Δ*sigB* strain as a cut-off value for our RNA-Seq data, we found that approximately 83% (2417/2888) of *L. monocytogenes* 10403S annotated CDS were transcribed in stationary phase cells. These transcribed genes include 355 putative operons, which cover a total of 1107 genes, indicating that a considerable proportion of *L. monocytogenes* genes appear to be transcribed polycistronically.

Overall, our data indicate that the majority of annotated *L. monocytogenes* genes are transcribed, even with application of a conservative cut-off value and use of bacteria grown under a single condition. This conclusion is consistent with results from a whole-genome tiled microarray transcriptome study of *E. coli* MG1655.

Tjaden et al. [32] found evidence of transcription of 4052 *E. coli* MG1655 genes in bacteria grown under different conditions. As the *E. coli* MG1655 genome in GenBank (NC 000913) is annotated as containing 4131 protein coding genes, these data suggest that, collectively, 98.1% of the *E. coli* MG1655 genes are transcribed in cells grown under different conditions.

Consistent with some previous studies (e.g., in yeast [13,27], human cell lines [33], murine tissue [29]), which have suggested that RNA-Seq generates data that allow for quantitative analysis of transcript levels, we found that RNA-Seq coverage levels correlate well with quantitative RT-PCR-based transcript level data. Previous reports include high correlations between quantitative transcript data generated using RNA-Seq (with the Illumina Genome Analyzer System) and (i) quantitative transcript data based on tiled microarrays (e.g., in yeast [13,27]), and (ii) quantitative transcript data based on qRT-PCR (e.g., for human tissue samples [34]). We also found a positive correlation between RNA-Seq-based transcript levels and codon bias, consistent with the well-documented observation that genes with high codon bias are often highly expressed [35,36,37]. Genes in four role categories, including (i) signal transduction, (ii) viral functions, (iii) amino acid biosynthesis, and (iv) transport and binding, were significantly associated with lower transcript levels. These categories likely include a number of genes that encode proteins predominantly required for growth and survival under specialized environmental conditions (e.g., viral replication genes) or conditions other than stationary phase (e.g., amino acid biosynthesis may not be required in stationary phase as sufficient amino acids from dead bacteria may be available for scavenging), and/or proteins that may only be required in small amounts. On the other hand, we found that genes in seven role categories, including (i) cellular processes, (ii) DNA metabolism, (iii) protein fate, (iv) protein synthesis, (v) purines, pyrimidines, nucleosides, and nucleotides, (vi) transcription, and (vii) genes encoding

proteins with unknown functions, showed, on average, higher transcript levels in stationary phase *L. monocytogenes*. These findings are consistent with an apparent role of proteins encoded by these genes in bacterial growth and metabolism.

Overall, the *L. monocytogenes* genes with the highest transcript levels were ncRNAs, specifically the transfer-messenger RNA (tmRNA) and 6S RNA (Table 3.2), consistent with the observation that tmRNAs are involved with bacterial recovery from a variety of stresses including entry into stationary phase, amino acid starvation, and heat shock [38]. 6S RNA has been shown to accumulate in cells during stationary phase; cells lacking 6S RNA have reduced fitness relative to wildtype cells during stationary phase [39]. In addition to down-regulating some housekeeping genes, 6S RNA may also up-regulate expression of some σ^S-dependent genes in Gram-negative bacteria [39]. σ^S is the stationary phase stress response alternative sigma factor in *E. coli* [40]. Overall, these previously reported data are consistent with our findings that 6S RNA transcript levels are high in stationary phase *L. monocytogenes* and suggest that 6S RNA plays a critical role in the ability of *L. monocytogenes* to survive stationary-phase associated stress conditions.

Specific protein-encoding genes with very high transcript levels in stationary phase *L. monocytogenes* include *fri, sod, cspB,* and *cspL*, all genes with some previous evidence for contributions to *L. monocytogenes* stationary phase and stress survival. For example, *L. monocytogenes* SOD activity has been shown to be highest in early-stationary phase at 37°C [41]. A Δ*sod* strain was found to be hypersensitive to increased intracellular levels of superoxides [42] and showed reduced growth in spleens and livers of intravenously infected mice, as compared to the wildtype strain [42]. A recent study found that Csp proteins were not necessary for optimal growth at 37°C, but were required for efficient growth under cold and osmotic stress in *L. monocytogenes* [43]. Chan et al. [44] found that *cspL* was highly transcribed in *L.*

monocytogenes 10403S grown at 37°C and 4°C. Further, the cold shock proteins CspB and CspC were highly expressed in stationary phase *Bacillus subtilis* based on 2D-gel protein analysis [45]. Jin et al. [46] previously observed that a *L. monocytogenes* EGD-e Δ*fri* strain grew more slowly and was impaired in its ability to regulate intracellular iron availability, while another study [47] found that a *L. monocytogenes* EGD Δ*fri* strain showed increased sensitivity to H_2O_2 and was impaired in its ability to proliferate inside a macrophage cell line. Our finding that *flaA*, which encodes a flagellin protein, was also highly transcribed in stationary phase cells was somewhat surprising as *L. monocytogenes* typically only shows flagellar motility when grown at ≤ 30°C [48,49]. However, strain 10403S, which was used here, has been shown to express flagellin at 37°C [49]. Interestingly, we also found some annotated CDS without known function to be extremely highly transcribed, including lmo1847 and lmo1849, which encode putative ABC transporters based on BLAST and Pfam [50] searches, respectively, and lmo1468, which encodes an unknown protein.

RNA-Seq identifies novel ncRNA molecules in *L. monocytogenes*, including a σ^B-dependent ncRNA, in 10403S

Using RNA-Seq, we found 65 previously identified or putative ncRNA that were transcribed in stationary phase *L. monocytogenes*. Of these, 53 represent ncRNAs that have not been identified previously as transcribed in *L. monocytogenes*, including 38 with similarities to ncRNAs identified in other bacterial organisms (based on Rfam searches) and 15 novel putative ncRNA molecules with no homologies to ncRNA entries in Rfam. Twelve of the ncRNAs identified here (including 10 that have been experimentally validated) have previously been reported by Nielsen et al. [18], Mandin et al. [17] and Christiansen et al.[19]. Interestingly, 16 *L.*

monocytogenes ncRNAs with similarities to ncRNAs identified in other bacterial organisms are putative riboswitches. Further, one novel putative ncRNA with no homologies to ncRNA entries in Rfam appears to be σ^B-dependent based on (i) significantly higher transcript levels (186 fold) in the parent strain as compared to the *sigB*-null mutant, and (ii) prediction of a σ^B promoter immediately upstream of the ncRNA transcriptional start site by the HMM used in this study. As the RNA isolation procedure used here selected against small RNA molecules (see Materials and Methods for details), it is likely that additional small ncRNAs not detected here also are transcribed in stationary phase *L. monocytogenes*.

Prior to this study, *L. monocytogenes* ncRNAs, including potential σ^B-dependent ncRNAs [18], had been identified using *in silico* modeling [17,18] or co-precipitation with the RNA-binding protein Hfq [19], followed by subsequent characterization. Nielsen et al. [18] found that among 4 putative σ^B-dependent ncRNAs genes identified *in silico* in *L. monocytogenes* strain EGD-e (i.e. *sbrA, sbrB, sbrC, sbrD*), only SbrA could be biologically confirmed as transcribed and σ^B-dependent; Nielsen et al. [18] did not show data confirming transcription of the *in silico* predicted *sbrB*, *sbrC*, and *sbrD*. While we did not detect SbrC transcripts using RNA-Seq, we did identify SbrA, SbrB, and SbrD transcripts, although they were not σ^B-dependent under the conditions used in our study. The fact that SbrA was not found to be σ^B-dependent here may be due to differences in strains or growth conditions used (e.g., Nielsen et al. [18] used strain EGD-e, while we used strain 10403S).

Identification of a large number of transcribed *L. monocytogenes* ncRNAs, including ncRNAs with no similarities to previously identified ncRNAs, clearly shows the power of RNA-Seq for characterizing bacterial transcriptomes, thus providing novel insights into transcriptional regulation. Our results, taken together with previous

studies that have identified numerous novel transcripts with RNA-Seq in *S. meliloti* [28], *B. cenocepacia* [15], yeast [13,27], mouse [29], Arabidopsis [51], human cell lines [33,52], and human tissue [34] underscores the power of the emerging high throughput sequencing techniques for improving our understanding of genome-wide transcription [1], including transcription of ncRNAs, among a wide variety of living organisms. -

The *L. monocytogenes* σ^B regulon is composed of at least 96 genes, including 1 putative ncRNA and 82 genes directly regulated by 65 σ^B promoters.

As alternative sigma factors, such as σ^B, are known to play critical roles in gene regulation across bacterial genera [31], we used *L. monocytogenes* 10403S and an isogenic Δ*sigB* null mutant as a model system for exploring the use of RNA-Seq, in combination with *in silico* analyses, for characterization of transcriptional blueprints associated with bacterial regulatory elements. In our study, RNA-Seq identified 96 annotated CDS and 1 putative ncRNA (SbrE) that showed significantly lower transcript levels in the Δ*sigB* strains, indicating that transcription of these genes is up-regulated by σ^B. Quantitative RT-PCR experiments also confirmed σ^B-dependent transcript levels of SbrE (Mujahid et al., unpublished). A number of the σ^B-dependent protein coding genes had also been identified in previous studies [10,22,26] using genome-wide microarrays. Specifically, among the 96 σ^B-dependent annotated CDS identified in this study, 74 (77.1%) [10] and 81 (84.4%) [22] were also identified as σ^B-dependent in stationary phase cells in two previous microarray studies using the same strain background. Also, 63 of the 96 σ^B-dependent genes identified here were reported as positively regulated by σ^B in another *L. monocytogenes* strain (EGD-e) grown to early stationary phase [26]. Twelve genes were identified as σ^B-dependent in both previous microarray studies performed with the same *L. monocytogenes* strain

background and the same conditions used here, but were not identified as σ^B-dependent by RNA-Seq in this study. This disparity is likely due to the fact that the thresholds and statistical cut-offs used to define σ^B-dependent genes were very stringent in the present study (e.g., a *q*-value < 0.05 in all four comparisons).

Overall, in addition to identifying a novel σ^B-dependent ncRNA, RNA-Seq identified 13 genes that had not been defined as σ^B-dependent in previous microarray studies of stationary phase *L. monocytogenes* 10403S cells [10,22], including 5 genes that had been identified as σ^B-dependent in salt stressed cells, but not in stationary phase cells. One gene not previously identified as σ^B-dependent was lmo2003, which encodes a transcription regulator similar to the GntR family. The GntR family of regulators has been characterized as global regulators of primary metabolism in a number of bacteria [53,54,55] and this finding further supports that *L. monocytogenes* σ^B appears to be involved in a number of transcriptional regulatory networks. The well-characterized virulence genes *inlA* and *inlB* (which have been shown by qRT-PCR and promoter mapping to be directly regulated by σ^B [56]) were identified, by RNA-Seq, as σ^B-dependent in stationary phase, even though they had not been identified as σ^B-dependent under stationary phase conditions in microarray studies by Raengpradub et al. [10], Hain et al. [26], or Ollinger et al. [22]. These genes encode proteins that contribute to the ability of *L. monocytogenes* to cross the intestinal barrier (InlA) [57,58] and to cross the placental barrier (InlA and InlB) [59].

While the observation that a number of genes found to be σ^B-dependent by RNA-Seq but not in previous microarray studies may be attributable to a number of factors, including (i) mismatches between oligonucleotide probes and 10403S sequences, and (ii) absolute transcript levels that were outside the dynamic range of microarrays, our data clearly show that RNA-Seq is a powerful tool that enables identification of differentially regulated genes. Importantly, RNA-Seq not only

allowed for identification of differentially transcribed genes, but also allowed (i) quantification of transcript levels, and (ii) mapping of the approximate location of transcriptional start and termination sites, particularly when RNA-Seq data are used in conjunction with appropriate bioinformatics tools, such as the iterative, dynamic HMM that we developed here, to identify putative σ^B promoters. The quantitative nature of RNA-Seq allowed us to identify highly transcribed σ^B-dependent genes, including lmo2158 (which encodes a protein similar to the *B. subtilis* YwmG), lmo1602 (which encodes an unknown protein), and lmo0539 (which encodes a tagatose-1,6-diphosphate aldolase). Interestingly, none of these genes encode proteins that appear to contribute to any of the presently recognized σ^B-dependent phenotypes in *L. monocytogenes*, such as acid resistance [9,60], oxidative stress resistance [60,61], or virulence [31,56,62,63]. There are no published reports of construction and characterization of null mutations in these highly transcribed σ^B-dependent genes. Our data clearly suggest that σ^B and the σ^B regulon make additional important contributions to *L. monocytogenes* physiology that remain to be characterized.

RNA-Seq data also allowed identification of approximate 5' and 3' transcript ends, which contributed to identification of putative promoters, which are usually located at a defined distance upstream of the transcription initiation site (and therefore at a defined distance upstream of the 5'ends of transcripts). In particular, a combination of RNA-Seq data and a novel iterative HMM approach identified putative σ^B promoters upstream of (i) 49 monocistronic σ^B-dependent genes, (ii) 15 σ^B-dependent operons (covering a total of 40 genes), and (iii) 1 σ^B-dependent ncRNA. By comparison, a previous study that solely relied on HMM and genome sequence data (without genome wide transcriptional start site data) only identified putative σ^B-dependent promoters upstream of 40 σ^B-dependent genes identified by microarray [10] in stationary phase. Our data reported here show that the majority of σ^B-dependent

genes are directly regulated by σ^B and illustrate the power of combining RNA-Seq data and bioinformatics approaches for characterization of transcriptional regulatory systems. Specifically, combining transcriptional start site information with an HMM that identifies promoter motifs (e.g., the motif for σ^B-dependent promoters) provides a powerful approach for identifying genes directly regulated by a given transcription factor. This approach allows rapid genome-wide identification of putative transcriptional start sites, which currently represents a critical bottleneck in genome-wide characterization of transcriptional regulation and regulatory networks, as many current strategies for promoter mapping are time- and labor-intensive (e.g., primer extension, rapid amplification of cDNA ends (RACE-PCR), RNAse protection assays).

CONCLUSIONS

Using the human foodborne pathogen *L. monocytogenes* as a model system, we have shown that RNA-Seq provides an extremely powerful method, particularly if combined with appropriate bioinformatics tools to (i) rapidly, comprehensively, and quantitatively characterize prokaryotic genome-wide transcription profiles without hybridization bias, and (ii) characterize putative transcriptional start sites and operon structures. We also show that RNA-Seq transcriptomic evaluation of a bacterial strain bearing a deletion in a transcriptional regulator in comparison with its parent strain can provide rapid, comprehensive insights into the blueprints of prokaryotic transcriptional regulation. These tools and approaches will revolutionize our ability to characterize genome-wide transcriptional regulatory networks, with wide ranging applications from medicine to ecology, e.g., by providing a means to quickly characterize transcriptional networks contributing to pathogen transmission and virulence as well as environmental growth and gene expression in bacteria used for specific purposes, such

as bio-remediation. When applied to both genome and transcriptome sequencing, novel high throughput sequencing approaches can also provide rapid and comprehensive characterization of bacterial genomes, representing an important tool for initial rapid characterization of novel and emerging bacterial pathogens.

MATERIALS AND METHODS

Strains and growth conditions. RNA-Seq was performed on the *L. monocytogenes* parent strain 10403S and a previously described [9] isogenic mutant (Δ*sigB*, FSL A1-254) with an internal non-polar deletion of *sigB*, which encodes the stress response alternative sigma factor σ^B.

Prior to RNA isolation, bacteria were grown in 5 ml Brain Heart Infusion (BHI) broth (BD Difco, Franklin Lakes, NJ) at 37ºC with shaking (230 rpm) for 15 h, followed by transfer of a 1% inoculum to 5 ml pre-warmed BHI. After growth to OD_{600} ~ 0.4, a 1% inoculum was transferred to a 300 ml nephelo flask (Bellco, Vineland, NJ) containing 50 ml pre-warmed BHI. This culture was incubated at 37ºC with shaking until cells reached stationary phase (defined as growth to OD_{600} = 1.0, followed by incubation for an additional 3 h). Two independent growth replicates and RNA isolations were performed for each strain.

RNA isolation, integrity and quality assessment. RNA isolation was performed as previously described [10]. Briefly, RNAProtect bacterial reagent (Qiagen, Valencia, CA) was added according to the manufacturer's instructions to the cultures grown to stationary phase; treated cells were stored at -80ºC (for no longer than 24 h) until RNA isolation was performed. Bacterial cells were treated with lysozyme followed by 6 sonication cycles at 18W on ice for 30s. Total RNA was isolated and purified using the RNeasy Midi kit (Qiagen) according to the manufacturer's protocol; RNA molecules <200 nt in length are not recovered well with this procedure, according to

the manufacturer. RNA was eluted from the column using RNase-free water. Total RNA was incubated with RQ1 DNase (Promega, Madison, WI) in the presence of RNasin (Promega) to remove remaining DNA. Subsequently, RNA was purified using two phenol-chloroform extractions and one chloroform extraction, followed by RNA precipitation and resuspension of the RNA in RNAse free TE (10 mM Tris, 1 mM EDTA; pH 8.0; Ambion, Austin, TX). UV spectrophotometry (Nanodrop, Wilmington, DE) was used to quantify and assess purity of the RNA.

Efficacy of the DNase treatment was assessed by TaqMan qPCR analysis of DNA levels for two housekeeping genes, *rpoB* [64] and *gap* [31]. qPCR was performed using TaqMan One-Step RT-PCR Master Mix Reagent and the ABI Prism 7000 Sequence Detection System (all from Applied Biosystems, Foster City, CA). Each RNA sample was run in duplicate and standard curves for each target gene were included for each assay to allow for absolute quantification of residual DNA. Data were analyzed using the ABI Prism 7000 Sequence Detection System software as previously described [65] Normalization and log transformation were performed as described by Kazmierczak et al. [21]. All samples showed log copy numbers ≤ 1.5 and C_t values > 35 for both *rpoB* and *gap*, indicating negligible levels of DNA contamination. As a final step, RNA integrity was assessed using the 2100 Bioanalzyer (Agilent, Foster City, CA).

mRNA enrichment. Removal of 16S and 23S rRNA from total RNA was performed using MicrobExpress™ Bacterial mRNA Purification Kit (Ambion) according to the manufacturer's protocol with the exception that no more than 5 μg total RNA was treated per enrichment reaction. Each RNA sample was divided into multiple aliquots of $\leq$ 5 μg RNA and separate enrichment reactions were performed for each sample. Enriched mRNA samples were pooled and run on the 2100 Bioanalzyer (Agilent) to

confirm reduction of 16S and 23S rRNA prior to preparation of cDNA fragment libraries.

Preparation of cDNA fragment libraries. Ambion RNA fragmentation reagents were used to generate 60-200 nucleotide RNA fragments with an input of 100 ng of mRNA. Following precipitation of fragmented RNA, first strand cDNA synthesis was performed using random N_6 primers and Superscript II Reverse Transcriptase, followed by second strand cDNA synthesis using RNaseH and DNA pol I (Invitrogen, CA). Double-stranded cDNA was purified using Qiaquick PCR spin columns according to the manufacturer's protocol (Qiagen).

RNA-Seq using the Illumina Genome Analyzer. The Illumina Genomic DNA Sample Prep kit (Illumina, Inc., San Diego, CA) was used according to the manufacturer's protocol to process double-stranded cDNA for RNA-Seq, including end repair, A-tailing, adapter ligation, size selection, and pre-amplification. Amplified material was loaded onto independent flow cells; sequencing was carried out by running 36 cycles on the Illumina Genome Analyzer.

The quality of the RNA-Seq reads was analyzed by assessing the relationship between the quality score and error probability; these analyses were performed on Illumina RNA-Seq quality scores that were converted to phred format [http://www.phrap.com/phred/]. Quality scores are reported in Figure A1 [S3.2].

RNA-Seq data will be available in the NCBI GEO Short Read Archives.

RNA-Seq alignment and coverage. The program nucmer, which is part of the MUMmer package (http://mummer.sourceforge.net/), was used to align the 10403S unfinished genome sequences (available at http://www.broad.mit.edu/annotation/genome/listeria_group/MultiHome.html as supercontigs 5.1 to 5.21) against the finished genome sequence of the *L. monocytogenes* reference strain EGD-e [16] to create a pseudochromosome for

10403S. Creation of the 10403S pseudochromosome was performed using the order and orientation of the 10403S supercontigs provided by the alignment with EGD-e. The annotation of the genes in the individual 10403S supercontigs, as provided by the Broad Institute (http://www.broad.mit.edu/annotation/genome/listeria_group/MultiHome.html) was then mapped to the 10403S pseudochromosome (Supplementary Materials S3). The 5S, 16S and 23S rRNA genes as well as the various tRNA genes in 10403S were identified using blastn and the EGD-e annotated rRNA and tRNA genes as a reference (Genbank ID: AL591824).

Based on quantitative analyses of RNA-Seq data, throughout this manuscript, transcript levels of a given gene are reported as the Gene Expression Index (GEI), which is expressed as number of reads per 100 bases. To obtain the GEI, the 10403S pseudochromosome was used to align Illumina RNA-Seq reads. These alignments were performed using the whole genome alignment software Eland (Illumina), which reports unique alignments of the first 32 bases of each read, allowing up to 2 mismatches. Coverage at each base position along the pseudochromosome was calculated by enumerating the number of reads that align to a given base. The coverage for each base from the first to last nt in an annotated CDS was summed then divided by 32 (i.e., the length of each aligned read) to obtain the RNA-Seq coverage for that gene before normalization. The following data were discarded prior to further analyses: (i) reads with more than 2 mismatches, (ii) reads that matched to multiple locations, (iii) reads that did not map to the chromosome, and (iv) reads that mapped to the 16S or 23S genes (Table 3.1). Reads identified as "matching two locations" did not include those matching rRNA genes as the 10403S pseudochromosome created for this study was designed with only one unique rRNA gene sequence. Reads matching the 16S and 23S genes were removed prior to normalizing the total number of aligned

reads across the four samples because of the technical bias introduced by our deliberate partial removal of 16S and 23S transcripts from the samples. Despite removal of 16S and 23S rRNA, in a given run, between 1,860,817 and 3,138,329 reads aligned to the 23S gene and between 434,263 and 760,863 reads aligned to the 16S gene. In a given run, between 101,419 and 242,246 reads matched the 5S rRNA gene and between 7,778 and 62,699 reads matched the various tRNA genes present in the pseudochromosome.

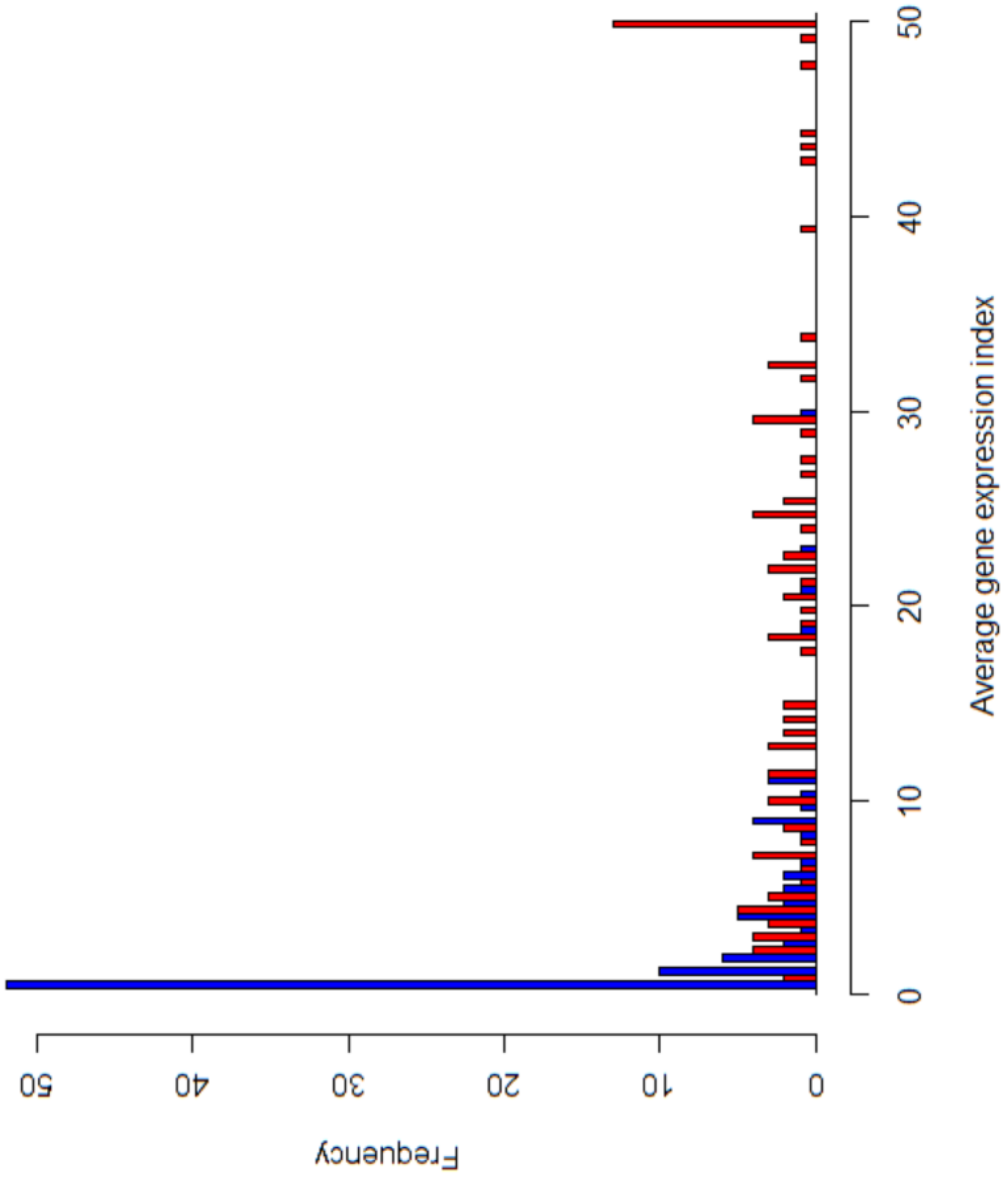
Because of the inherent differences in the total number of reads among the four runs, the total number of reads for each run was normalized to the run with the highest coverage (i.e. Δ*sigB* replicate 2, Table 3.1). The ratio of total number of reads for Δ*sigB* replicate 2 to the total number of reads for 10403S replicate 1, 10403S replicate 2, or Δ*sigB* replicate 2 was used as a multiplier to normalize the approximate number of reads matching a given gene (Table 3.1). The GEI was then obtained by dividing the normalized number of reads matching each gene by the gene length. The average GEI was the number of reads that match each nt in a given gene after normalization; this value represented the average of the 2 biological replicates for a given strain and is presented as reads per 100 bases (as opposed to reads per 1 base) to simplify identification of differences. The distribution of the coefficient of variation for each gene between replicates is depicted in Figure A1 [S3.3].

Identification of transcribed annotated CDS. Sequence reads matching annotated CDS in the 10403S genome were used to identify those annotated CDS that were transcribed under the experimental conditions used. As our RNA-Seq analyses included both a wildtype strain and an isogenic mutant with a deletion in a transcriptional regulator (i.e., the alternative sigma factor σ^B), our data also provide a novel approach for characterizing background RNA-Seq coverage for genes that are not transcribed, similar to a previous approach that used background RNA-Seq

coverage of so-called "gene deserts" in human chromosomes to characterize background average GEI [66]. The observations that (i) eight genes that showed average GEI between 8.64 reads and 96.43 reads per 100 bases in the parent strain showed 0 reads per 100 bases in the Δ*sigB* strain; (ii) 42 genes with average GEI of 1.21 to 73.81 reads per 100 bases in the parent strain showed between 0.01 and 0.7 reads per 100 bases in the Δ*sigB* strain; and (iii) 0.7 reads per 100 bases is the approximate median of the average GEI in σ^B-dependent genes in the Δ*sigB* strain, clearly indicate that extremely low background RNA-Seq coverage is expected for genes that are not transcribed. Overall, 50/96 σ^B-dependent genes show an average GEI < 0.7 in the Δ*sigB* strain (Table A2 [S3.3]); genes with GEI < 0.7 reads are overrepresented in the Δ*sigB* strain (Figure 3.6). It is not unexpected that some σ^B-dependent genes showed average GEI ≥ 0.7 as a number of genes are not solely dependent on σ^B and will still be transcribed in the absence of σ^B (e.g., *opuCABCD* operon [30,67,68]). Based on these observations, we set an average GEI ≥ 0.7 as a conservative cut-off to identify genes that are transcribed (i.e., we define genes with average GEI ≥ 0.7 as being transcribed as the RNA-Seq data indicate that non-specific reads [e.g., from DNA] are highly unlikely to provide average GEI ≥ 0.7).

Depending on RNA-Seq coverage, genes were classified into four categories, including (i) not transcribed (average GEI < 0.7), (ii) low transcript levels (average GEI ≥ 0.7 and < 10), (iii) medium transcript levels (average GEI ≥ 10 and < 25), and (iv) high transcript levels (average GEI ≥ 25). While cut-offs between low, medium, and high transcript level categories were somewhat arbitrary, they were chosen to yield a relative distribution of genes into these categories similar to the distribution of yeast genes into low, medium, and high expression categories reported previously by Nagalakshimi et al. [13].

Figure 3.6 Average gene expression indices for σ^B-dependent genes. The histogram shows the average GEI of σ^B-dependent genes in 10403S (red) and the Δ*sigB* (blue) strains. GEIs were grouped in intervals of 0.7, i.e., the first bar represents genes with GEIs between 0 and 0.7; the second bar represents GEIs between > 0.7 and ≤ 1.4, etc. Genes with average GEI ≥ 50 were grouped together.


Frequency
0
10
20
30
40
50
Average gene expression index
0
10
20
30
40
50

.

Annotation of Rho-independent terminators and putative operons. Potential operons were manually annotated based on the continuity of a similar level of RNA-Seq coverage across consecutive genes and the (i) absence of putative Rho-independent terminators between genes, and/or (ii) presence of a putative Rho-independent terminator at the end of a putative operon. Putative Rho-independent terminators in the 10403S pseudochromosome were identified using the program TransTermHP v2.04 [69].

Discovery and annotation of regions transcribing ncRNAs. To aid in identification of transcribed ncRNAs, ncRNAs previously identified in *L. monocytogenes* EGD-e were mapped onto the 10403S pseudochromosome; 16 of the 18 ncRNA previously identified in *L. monocytogenes* EGD-e [17,18,19] matched the 10403S pseudochromosome; the RliC ncRNAs identified in EGD-e did not have a homolog in the 10403S pseudochromosome while LhrC also identified in EGD-e falls in a repetitive region [19] that has not been completely sequenced and closed in 10403S. An additional 38 ncRNAs identified by Rfam in the EGD-e genome were also mapped onto the 10403S pseudochromosome.

New putative ncRNAs (i.e., ncRNAs not previously reported or previously identified by Rfam) were manually identified using the genome browser Artemis [70]. Specifically, regions not matching annotated genes, but showing contiguous coverage by RNA-Seq reads (i.e., regions that contain at least 100 bp completely covered by RNA-Seq reads) were designated putative ncRNAs. Further, RNA-Seq reads that did not cover an entire annotated CDS, but showed partial contiguous coverage within a CDS, were also designated as putative ncRNAs. All ncRNAs, including those reported in previous publications [17,18,19], those identified by Rfam, and those with no matches to the Rfam database were annotated into a Genbank (gbk) file that is available as Supplemental Material S3. ncRNAs identified by RNA-Seq, but with no

matches to the Rfam database were designated "putative ncRNA". The presence of rho-independent transcriptional terminators was used to assign the strand of putative ncRNAs. For two instances where terminators were not observed, the ncRNAs were annotated on both strands. ncRNAs previously described in the literature were annotated even if no RNA-Seq coverage was observed for these ncRNAs in this study.

Differential expression analysis. To identify genes that showed significantly different transcript levels in the parent strain (10403S) and the Δ*sigB* strain, statistical analyses were performed using the normalized RNA-Seq coverage of each coding gene (as annotated by the Broad Institute). Normalized RNA-Seq coverage (i.e. the number of reads that match an annotated CDS after normalization across runs) was used in lieu of the GEI (in which the normalized RNA-Seq coverage number is divided by the gene length) for statistical analyses. Corresponding analyses were also performed for each region encoding a putative ncRNA transcript identified as described above. A coverage file of normalized RNA-Seq coverage is available in Supplemental Materials S4.

For each gene, a binomial probability was calculated for the normalized RNA-Seq coverage, using each of the four possible comparisons between the 10403S and Δ*sigB* transcripts (i.e. 10403S replicate 1 vs Δ*sigB* replicate 1; 10403S replicate 1 vs Δ*sigB* replicate 2; 10403S replicate 2 vs Δ*sigB* replicate 1; 10403S replicate 2 vs Δ*sigB* replicate 2). The binomial probability was calculated under the hypothesis that genes that are not regulated by σ^B will show the same normalized number of reads in the two strains ($p = 0.5$ and $q = 0.5$). For a gene to be considered up-regulated by σ^B, the binomial probability of observing as many reads in the Δ*sigB* strain as those observed for 10403S had to be < 0.05 for each of the four possible combinations. Conversely, for a gene to be considered down-regulated by σ^B, the binomial probability of observing as many reads as those observed for Δ*sigB* had to have *q*-

values < 0.05 for each of the four possible combinations. To control for multiple comparisons, a False Discovery Rate (FDR) approach was used. *q*-values (representing the FDR) were calculated using the program *Q*-Value [71] for R. Only genes with *q*-values < 0.05 and fold change ≥ 2 or ≤ 0.5 among all four possible comparisons between 10403S and Δ*sigB* were considered significantly up-regulated or down-regulated by σ^B.

Iterative HMM-based promoter identification. An initial training set containing 17 experimentally validated σ^B-dependent promoter motifs was used to build a Hidden Markov Model (HMM) of these motifs (Table A3 [S3.5]). HMM construction and searches were performed using the program hmmer version 1.8.5. The HMM was constructed from unaligned sequences (using hmmt) and then used to search the 10403S pseudochromosome (using the hmmls tool). The null frequencies of each nucleotide used were those observed in the *L. monocytogenes* genome (i.e., A/T = 0.31 and G/C = 0.19).

To identify new promoter motifs that could be added to the training set, we used an iterative HMM approach. In each given HMM iteration, the only hits added to the training set were those that met four conservative criteria, including (i) location within 100 bp upstream of the start codon of an annotated CDS (or 100 bp upstream the first nt for the manually annotated noncoding genes), (ii) *q*-values < 0.05 (from the binomial probabilities) for σ^B dependence of a given gene (based on RNA-Seq data), and (iii) fold change ≥ 2 among all possible comparisons between 10403S and Δ*sigB*, and (iv) a score higher than the lowest score for which 50% of the motifs fall in noncoding regions (i.e. for each iteration, we adaptively chose a threshold score such that 50% of the motifs that score higher than this threshold lie in noncoding regions). After adding all hits that met these criteria (in a given iteration) to the training set, a new model was built and used to search the 10403S pseudochromosome. This process

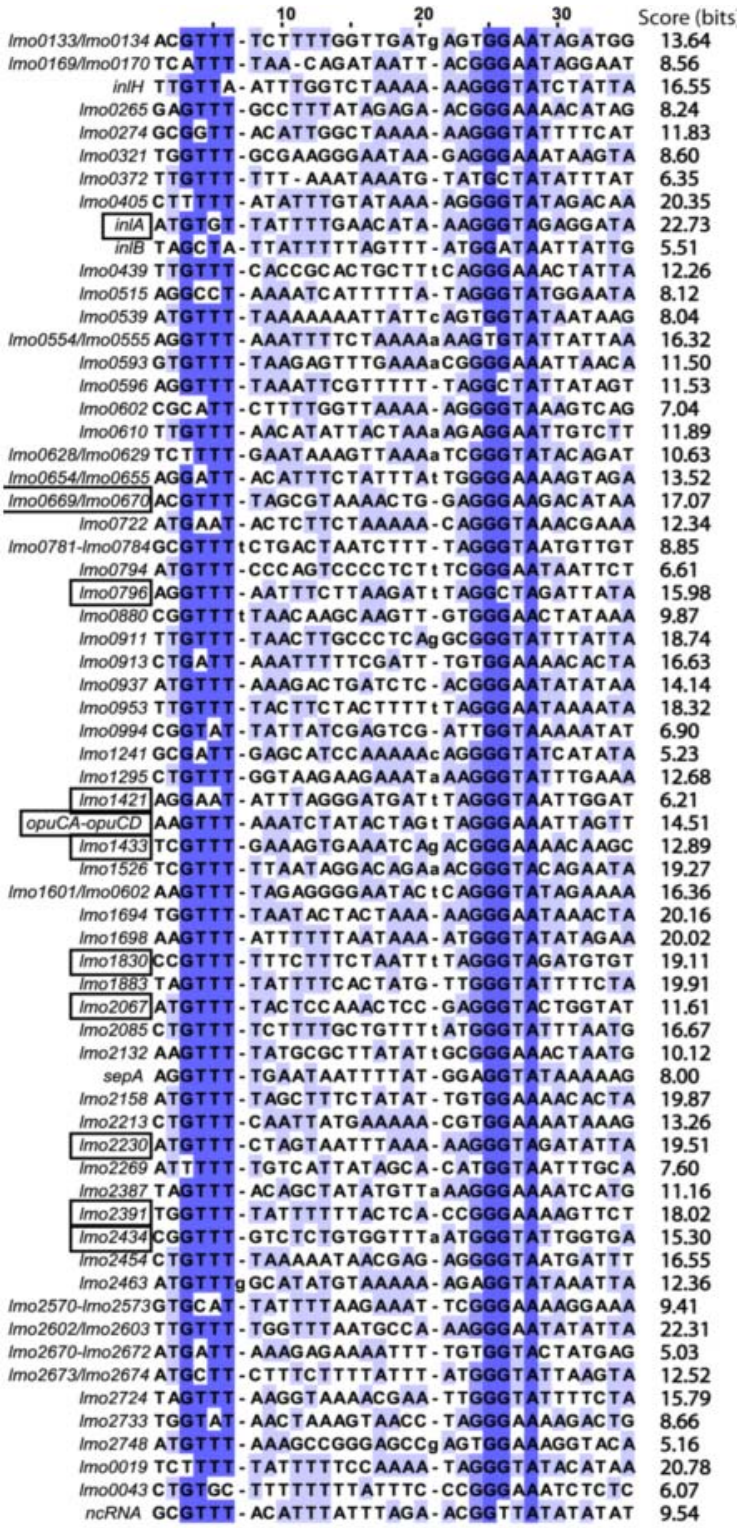
was repeated until no new motifs could be added to the training set; the final training set can be found in Table A3 [S3.5]. When no new motifs that matched our criteria were discovered, the model was considered complete and the results from the last search were used for promoter identification. The final model was used to search the 10403S pseudochromosome for potential σ^B promoters. Potential σ^B promoters identified by this HMM upstream of σ^B-dependent genes and the σ^B-dependent putative ncRNA were visually evaluated. Potential σ^B promoters identified by HMM were considered probable σ^B promoters if the promoter was within 50 bp upstream of the transcriptional start site (as identified by RNA-Seq). In some instances, the transcriptional start site was not discernable due to an upstream gene transcript that overlapped with a σ^B-dependent gene transcript or because the gene had a low average relative normalized RNA-Seq coverage. For these instances, putative promoters were considered if they were located within 200 bp from the start codon of the σ^B-dependent gene. σ^B-dependent genes with probable σ^B promoters are described in Figure 3.7; the σ^B promoter sequence logo is presented in Figure 3.4 (http://weblogo.berkeley.edu/) [72].

Correlation of RNA-Seq relative coverage (GEI) with TaqMan absolute transcript copy number. Average GEI was correlated with absolute transcript copy numbers quantified by TaqMan qRT-PCR. qRT-PCR-based transcript level data obtained for selected genes in *L. monocytogenes* grown under the same conditions used here (i.e., stationary phase) were obtained from previous studies and unpublished work (see Table A2 [S3.1]); qRT-PCR methods are detailed in Raengpradub et al. [10]. qRT-PCR data from these studies were used to calculate absolute transcript copy numbers (using a standard curve as described by Sue et al. [65]); values were log transformed.

Statistical Analyses. One-sided Wilcoxon rank sum tests were used to assess whether genes in certain role categories showed lower or higher average GEI in 10403S than genes in other role categories. One-sided Fisher's exact tests were used to assess whether σ^B-dependent genes were overrepresented in certain TIGR role categories (http://cmr.jcvi.org/cgi-bin/CMR/RoleIds.cgi). Linear regression analysis was used to assess correlations between average GEI and qRT-PCR data as well as between codon bias and average GEI in 10403S. The effective number of codons used in a gene (Nc), a measure of the codon bias, was assessed using the program "chips" implemented in the EMBOSS package [73]. All tests were carried out in R (version 2.7.0; http://www.r-project.org/). Correction for multiple testing was performed using the procedure reported by Benjamini & Hochberg [74], as implemented in the program *Q*-Value [71]. Significance was set at 5%.

Data access. RNA-Seq data will be available in the NCBI GEO Short Read Archives. All RNA-Seq data are provided in an Access database file (Supplemental Materials S2). This database contains information on the annotated CDS and ncRNAs with their 10403S locus name, 10403S start and end coordinates, lengths, strand, EGD-e locus, EGD-e gene name, EGD-e common name, EGD-e role category, codon bias, GEI, average GEI in 10403S and Δ*sigB* strains, fold change for the four possible comparisons involving the two replicates with 10403S and the Δ*sigB* strains, *q*-values of the binomial tests, operon annotation, promoter annotation, list of σ^B-dependent genes identified in this study, and data from 3 other studies of the σ^B regulon in *L. monocytogenes* using microarrays including Ollinger et al. [22], Hain et al. [26] , and Raengpradub et al. [10].

Figure 3.7 Alignment of the 65 putative σ^B-dependent promoters identified in this study. EGD-e homologs of genes or operons downstream of a given promoters are indicated on the left. Positions 3 to 6 in the alignment represent the -35 region while positions 24 to 29 represent the -10 region. Darker nucleotides are more conserved than lighter nucleotides in the alignment. Gene names that are boxed indicated promoters that have been experimentally validated (e.g., by RACE-PCR).


10
20
30
Score (bits
lmo0133/lmo0134 ACGTTT-TCTTTTGGTTGATg AGTGGAATAGATGG 13.64
lmo0169/lmo0170 TCATTT-TAA-CAGATAATT-ACGGGAATAGGAAT 8.56
inlH TTGTTA-ATTTGGTCTAAAA-AAGGGTATCTATTA 16.55
lmo0265 GAGTTT-GCCTTTATAGAGA-ACGGGAAAACATAG 8.24
lmo0274 GCGGTT-ACATTGGCTAAAA-AAGGGTATTTTCAT 11.83
lmo0321 TGGTTT-GCGAAGGGAATAA-GAGGGAAATAAGTA 8.60
lmo0372 TTGTTT-TTT-AAATAAATG-TATGCTATATTTAT 6.35
lmo0405 CTTTTT-ATATTTGTATAAA-AGGGGTATAGACAA 20.35
inlA ATGTGT-TATTTTGAACATA-AAGGGTAGAGGATA 22.73
inlB TAGCTA-TTATTTTTAGTTT-ATGGATAATTATTG 5.51
lmo0439 TTGTTT-CACCGCACTGCTTtCAGGGAAACTATTA 12.26
lmo0515 AGGCCT-AAAATCATTTTTA-TAGGGTATGGAATA 8.12
lmo0539 ATGTTT-TAAAAAAATTATTcAGTGGTATAATAAG 8.04
lmo0554/lmo0555 AGGTTT-AAATTTTCTAAAAaAAGTGTATTATTAA 16.32
lmo0593 GTGTTT-TAAGAGTTTGAAAaCGGGGAAATTAACA 11.50
lmo0596 AGGTTT-TAAATTCGTTTTT-TAGGCTATTATAGT 11.53
lmo0602 CGCATT-CTTTTGGTTAAAA-AGGGGTAAAGTCAG 7.04
lmo0610 TTGTTT-AACATATTACTAAaAGAGGAATTGTCTT 11.89
lmo0628/lmo0629 TCTTTT-GAATAAAGTTAAAaTCGGGTATACAGAT 10.63
lmo0654/lmo0655 AGGATT-ACATTTCTATTTAtTGGGGAAAAGTAGA 13.52
lmo0669/lmo0670 ACGTTT-TAGCGTAAAACTG-GAGGGAAGACATAA 17.07
lmo0722 ATGAAT-ACTCTTCTAAAAA-CAGGGTAAACGAAA 12.34
lmo0781-lmo0784 GCGTTTtCTGACTAATCTTT-TAGGGTAATGTTGT 8.85
lmo0794 ATGTTT-CCCAGTCCCCTCTtTCGGGAATAATTCT 6.61
lmo0796 AGGTTT-AATTTCTTAAGATtTAGGCTAGATTATA 15.98
lmo0880 CGGTTTtTAACAAGCAAGTT-GTGGGAACTATAAA 9.87
lmo0911 TTGTTT-TAACTTGCCCTCAgGCGGGTATTTATTA 18.74
lmo0913 CTGATT-AAATTTTTCGATT-TGTGGAAAACACTA 16.63
lmo0937 ATGTTT-AAAGACTGATCTC-ACGGGAATATATAA 14.14
lmo0953 TTGTTT-TACTTCTACTTTTtTAGGGAATAAAATA 18.32
lmo0994 CGGTAT-TATTATCGAGTCG-ATTGGTAAAAATAT 6.90
lmo1241 GCGATT-GAGCATCCAAAAAcAGGGGTATCATATA 5.23
lmo1295 CTGTTT-GGTAAGAAGAAATaAAGGGTATTTGAAA 12.68
lmo1421 AGGAAT-ATTTAGGGATGATtTAGGGTAATTGGAT 6.21
opuCA-opuCD AAGTTT-AAATCTATACTAGtTAGGGAAATTAGTT 14.51
lmo1433 TCGTTT-GAAAGTGAAATCAgACGGGAAAACAAGC 12.89
lmo1526 TCGTTT-TTAATAGGACAGAaACGGGTACAGAATA 19.27
lmo1601/lmo0602 AAGTTT-TAGAGGGGAATACtCAGGGTATAGAAAA 16.36
lmo1694 TGGTTT-TAATACTACTAAA-AAGGGAATAAACTA 20.16
lmo1698 AAGTTT-ATTTTTTAATAAA-ATGGGTATATAGAA 20.02
lmo1830 CCGTTT-TTTCTTTCTAATTtTAGGGTAGATGTGT 19.11
lmo1883 TAGTTT-TATTTTCACTATG-TTGGGTATTTTCTA 19.91
lmo2067 ATGTTT-TACTCCAAACTCC-GAGGGTACTGGTAT 11.61
lmo2085 CTGTTT-TCTTTTGCTGTTTtATGGGTATTTAATG 16.67
lmo2132 AAGTTT-TATGCGCTTATATtGCGGGAAACTAATG 10.12
sepA AGGTTT-TGAATAATTTTAT-GGAGGTATAAAAAG 8.00
lmo2158 ATGTTT-TAGCTTTCTATAT-TGTGGAAAACACTA 19.87
lmo2213 CTGTTT-CAATTATGAAAAA-CGTGGAAAATAAAG 13.26
lmo2230 ATGTTT-CTAGTAATTTAAA-AAGGGTAGATATTA 19.51
lmo2269 ATTTTT-TGTCATTATAGCA-CATGGTAATTTGCA 7.60
lmo2387 TAGTTT-ACAGCTATATGTTaAAGGGAAAATCATG 11.16
lmo2391 TGGTTT-TATTTTTTACTCA-CCGGGAAAAGTTCT 18.02
lmo2434 CGGTTT-GTCTCTGTGGTTTaATGGGTATTGGTGA 15.30
lmo2454 CTGTTT-TAAAAATAACGAG-AGGGGTAATGATTT 16.55
lmo2463 ATGTTTgGCATATGTAAAAA-AGAGGTATAAATTA 12.36
lmo2570-lmo2573 GTGCAT-TATTTTAAGAAAT-TCGGGAAAAGGAAA 9.41
lmo2602/lmo2603 TTGTTT-TGGTTTAATGCCA-AAGGGAATATATTA 22.31
lmo2670-lmo2672 ATGATT-AAAGAGAAAATTT-TGTGGTACTATGAG 5.03
lmo2673/lmo2674 ATGCTT-CTTTCTTTTATTT-ATGGGTATTAAGTA 12.52
lmo2724 TAGTTT-AAGGTAAAACGAA-TTGGGTATTTTCTA 15.79
lmo2733 TGGTAT-AACTAAAGTAACC-TAGGGAAAAGACTG 8.66
lmo2748 ATGTTT-AAAGCCGGGAGCCgAGTGGAAAGGTACA 5.16
lmo0019 TCTTTT-TATTTTTCCAAAA-TAGGGTATACATAA 20.78
lmo0043 CTGTGC-TTTTTTTTATTTC-CCGGGAAATCTCTC 6.07
ncRNA GCGTTT-ACATTTATTTAGA-ACGGTTATATATAT 9.54

http://www.r-project.org/). Correction for multiple testing was performed using the procedure reported by Benjamini & Hochberg [74], as implemented in the program *Q*-Value [71]. Significance was set at 5%.

Data access. RNA-Seq data will be available in the NCBI GEO Short Read Archives. All RNA-Seq data are provided in an Access database file (Supplemental Materials S2). This database contains information on the annotated CDS and ncRNAs with their 10403S locus name, 10403S start and end coordinates, lengths, strand, EGD-e locus, EGD-e gene name, EGD-e common name, EGD-e role category, codon bias, GEI, average GEI in 10403S and Δ*sigB* strains, fold change for the four possible comparisons involving the two replicates with 10403S and the Δ*sigB* strains, *q*-values of the binomial tests, operon annotation, promoter annotation, list of σ^B-dependent genes identified in this study, and data from 3 other studies of the σ^B regulon in *L. monocytogenes* using microarrays including Ollinger et al. [22], Hain et al. [26] , and Raengpradub et al. [10].

SUPPORTING MATERIALS

All supporting materials are available at http://www.foodscience.cornell.edu/cals/foodsci/research/labs/wiedmann/links/.

Supplemental Material S1: Sequencibility text file. The resulting plot, when used in conjunction with the Artemis genome browser, shows the regions that can (0) and cannot (1) be sequenced in the 10403S pseudochromosome with the Illumina Genome Analyzer. Regions that cannot be sequenced appear as high peaks.

Supplemental Material S2: Access database. All RNA-Seq data are provided in an Access database file. This database contains information on the annotated CDS and ncRNAs with their 10403S locus name, 10403S start and end coordinates, lengths, strand, EGD-e locus, EGD-e gene name, EGD-e common name, EGD-e role category, codon bias, GEI, average GEI in 10403S and Δ*sigB* strains, fold change for the four

possible comparisons involving the two replicates with 10403S and the Δ*sigB* strains, *q*-values of the binomial tests, operon annotation, promoter annotation, list of σ^B-dependent genes identified in this study, and data from the other 3 studies of the σ^B regulon in *L. monocytogenes* using microarrays including Ollinger et al. [22], Hain et al. [26], and Raengpradub et al. [10].

Supplemental Material S3: Genbank (gbk) file with ncRNAs identified here.

Supplemental Material S4: Coverage file with the normalized RNA-Seq coverage for the 4 RNA-Seq runs.

ACKNOWLEDGEMENTS: We thank P. Schweitzer and the staff at the Cornell DNA Sequencing and Genotyping Core Facility for sample preparation and sequencing and A. G. Clark and T. B. Sackton for helpful discussion.

FINANCIAL DISCLOSURE: This work was funded by NIH-NIAID (R01 AI052151 to K.J.B.). U.K. was supported by NSF (award no. 0644136).

REFERENCES

1. Wang Z, Gerstein M, Snyder M: **RNA-Seq: a revolutionary tool for transcriptomics.** *Nat Rev Genet* 2009, **10:**57-63.
2. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, Griffin PM, Tauxe RV: **Food-related illness and death in the United States.** *Emerg Infect Dis* 1999, **5:**607-625.
3. Begley M, Gahan CG, Hill C: **Bile stress response in *Listeria monocytogenes* LO28: adaptation, cross-protection, and identification of genetic loci involved in bile resistance.** *Appl Environ Microbiol* 2002, **68:**6005-6012.
4. Phan-Thanh L, Gormon T: **Analysis of heat and cold shock proteins in *Listeria* by two-dimensional electrophoresis.** *Electrophoresis* 1995, **16:**444-450.
5. Watkins J, Sleath KP: **Isolation and enumeration of *Listeria monocytogenes* from sewage, sewage sludge and river water.** *J Appl Bacteriol* 1981, **50:**1-9.
6. Chaturongakul S, Raengpradub S, Wiedmann M, Boor KJ: **Modulation of stress and virulence in *Listeria monocytogenes*.** *Trends Microbiol* 2008, **16:**388-396.
7. Kazmierczak MJ, Wiedmann M, Boor KJ: **Alternative sigma factors and their roles in bacterial virulence.** *Microbiol Mol Biol Rev* 2005, **69:**527-543.
8. Piggot PJ, Hilbert DW: **Sporulation of *Bacillus subtilis*.** *Curr Opin Microbiol* 2004, **7:**579-586.
9. Wiedmann M, Arvik TJ, Hurley RJ, Boor KJ: **General stress transcription factor σ^B and its role in acid tolerance and virulence of *Listeria monocytogenes*.** *J Bacteriol* 1998, **180:**3650-3656.

10. Raengpradub S, Wiedmann M, Boor KJ: **Comparative analysis of the σ^B-dependent stress responses in *Listeria monocytogenes* and *Listeria innocua* strains exposed to selected stress conditions.** *Appl Environ Microbiol* 2008, **74:**158-171.

11. Asmann YW, Wallace MB, Thompson EA: **Transcriptome profiling using next-generation sequencing.** *Gastroenterology* 2008, **135:**1466-1468.

12. Mockler TC, Ecker JR: **Applications of DNA tiling arrays for whole-genome analysis.** *Genomics* 2005, **85:**1-15.

13. Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M, Snyder M: **The transcriptional landscape of the yeast genome defined by RNA sequencing.** *Science* 2008, **320:**1344-1349.

14. Schmittgen TD, Lee EJ, Jiang J, Sarkar A, Yang L, Elton TS, Chen C: **Real-time PCR quantification of precursor and mature microRNA.** *Methods* 2008, **44:**31-38.

15. Yoder-Himes DR, Chain PSG, Zhu Y, Wurtzel O, Rubin EM, Tiedje JM, Sorek R: **Mapping the *Burkholderia cenocepacia* niche response via high-throughput sequencing.** *Proc Natl Acad Sci U S A* 2009, **106:**3976-3981.

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

17. Mandin P, Repoila F, Vergassola M, Geissmann T, Cossart P: **Identification of new noncoding RNAs in *Listeria monocytogenes* and prediction of mRNA targets.** *Nuc Acids Res* 2007, **35:**962-974.

18. Nielsen JS, Olsen AS, Bonde M, Valentin-Hansen P, Kallipolitis BH: **Identification of a σ^B-dependent small noncoding RNA in *Listeria monocytogenes*.** *J Bacteriol* 2008, **190:**6264-6270.

19. Christiansen JK, Nielsen JS, Ebersbach T, Valentin-Hansen P, Sogaard-Andersen L, Kallipolitis BH: **Identification of small Hfq-binding RNAs in *Listeria monocytogenes*.** *RNA (N Y)* 2006, **12:**1383-1396.

20. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A: **Rfam: annotating non-coding RNAs in complete genomes.** *Nucl Acids Res* 2005, **33:**D121-124.

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

22. Ollinger J, Bowen B, Wiedmann M, Boor KJ, Bergholtz TM: ***Listeria monocytogenes* σ^B modulates PrfA-mediated virulence factor expression.** *Infect Immun* 2009, **77:**2113-2124.

23. Shetron-Rama LM, Mueller K, Bravo JM, Bouwer HG, Way SS, Freitag NE: **Isolation of *Listeria monocytogenes* mutants with high-level in vitro expression of host cytosol-induced gene products.** *Mol Microbiol* 2003, **48:**1537-1551.

24. McGann P, Raengpradub S, Ivanek R, Wiedmann M, Boor KJ: **Differential regulation of *Listeria monocytogenes* internalin and internalin-like genes by σ^B and PrfA as revealed by subgenomic microarray analyses.** *Food Path Dis* 2008, **5:**417-435.

25. Mueller KJ, Freitag NE: **Pleiotropic enhancement of bacterial pathogenesis resulting from the constitutive activation of the *Listeria monocytogenes* regulatory factor PrfA.** *Infect Immun* 2005, **73:**1917-1926.

26. Hain T, Hossain H, Chatterjee SS, Machata S, Volk U, Wagner S, Brors B, Haas S, Kuenne CT, Billion A, et al: **Temporal transcriptomic analysis of the *Listeria monocytogenes* EGD-e σ^B regulon.** *BMC Microbiol* 2008, **8:**20.

27. Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bahler J: **Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution.** *Nature* 2008, **453:**1239-1243.

28. Mao C, Evans C, Jensen RV, Sobral BW: **Identification of new genes in *Sinorhizobium meliloti* using the Genome Sequencer FLX system.** *BMC Microbiol* 2008, **8:**72.

29. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq.** *Nat Meth* 2008, **5:**621-628.

30. Chan YC, Boor KJ, Wiedmann M: **σ^B-dependent and σ^B-independent mechanisms contribute to transcription of *Listeria monocytogenes* cold stress genes during cold shock and cold growth.** *Appl Environ Microbiol* 2007, **73:**6019-6029.

31. Kazmierczak MJ, Wiedmann M, Boor KJ: **Contributions of *Listeria monocytogenes* σ^B and PrfA to expression of virulence and stress response genes during extra- and intracellular growth.** *Microbiology (Read)* 2006, **152:**1827-1838.

32. Tjaden B, Saxena RM, Stolyar S, Haynor DR, Kolker E, Rosenow C: **Transcriptome analysis of *Escherichia coli* using high-density oligonucleotide probe arrays.** *Nuc Acids Res* 2002, **30:**3732-3738.

33. Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, Scherf M, Seifert M, Borodina T, Soldatov A, Parkhomchuk D, et al: **A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome.** *Science* 2008, **321:**956-960.

34. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y: **RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays.** *Genome Res* 2008, **18:**1509-1517.

35. Gouy M, Gautier C: **Codon usage in bacteria: correlation with gene expressivity.** *Nucl Acids Res* 1982, **10:**7055-7074.

36. Ikemura T: **Codon usage and tRNA content in unicellular and multicellular organisms.** *Mol Biol Evol* 1985, **2:**13-34.

37. Kanaya S, Yamada Y, Kudo Y, Ikemura T: **Studies of codon usage and tRNA genes of 18 unicellular organisms and quantification of *Bacillus subtilis* tRNAs: gene expression level and species-specific diversity of codon usage based on multivariate analysis.** *Gene* 1999, **238:**143-155.

38. Keiler KC: **Biology of trans-translation.** *Ann Rev Microbiol* 2008, **62:**133-151.

39. Trotochaud AE, Wassarman KM: **6S RNA function enhances long-term cell survival.** *J Bacteriol* 2004, **186:**4978-4985.

40. Loewen PC, Hengge-Aronis R: **The role of the sigma factor sigma S (KatF) in bacterial global regulation.** *Ann Rev Microbiol* 1994, **48:**53-80.

41. Vasconcelos JA, Deneer HG: **Expression of superoxide dismutase in *Listeria monocytogenes*.** *Appl Environ Microbiol* 1994, **60:**2360-2366.

42. Archambaud C, Nahori MA, Pizarro-Cerda J, Cossart P, Dussurget O: **Control of *Listeria* superoxide dismutase by phosphorylation.** *J Biol Chem* 2006, **281:**31812-31822.

43. Schmid B, Klumpp J, Raimann E, Loessner MJ, Stephan R, Tasara T: **Role of cold shock proteins (Csp) for growth of *Listeria monocytogenes* under cold and osmotic stress conditions.** *Appl Environ Microbiol* 2009, **75:**1621-1627.

44. Chan YC, Raengpradub S, Boor KJ, Wiedmann M: **Microarray-based characterization of the *Listeria monocytogenes* cold regulon in log- and stationary-phase cells.** *Appl Environ Microbiol* 2007, **73:**6484-6498.

45. Graumann PL, Marahiel MA: **Cold shock proteins CspB and CspC are major stationary-phase-induced proteins in *Bacillus subtilis*.** *Arch Microbiol* 1999, **171:**135-138.

46. Jin B, Newton SM, Shao Y, Jiang X, Charbit A, Klebba PE: **Iron acquisition systems for ferric hydroxamates, haemin and haemoglobin in *Listeria monocytogenes*.** *Mol Microbiol* 2006, **59:**1185-1198.

47. Olsen KN, Larsen MH, Gahan CGM, Kallipolitis B, Wolf XA, Rea R, Hill C, Ingmer H: **The Dps-like protein Fri of *Listeria monocytogenes* promotes stress tolerance and intracellular multiplication in macrophage-like cells** *Microbiology (Read)* 2005, **151:**925-933.

48. Bigot A, Pagniez H, Botton E, Frehel C, Dubail I, Jacquet C, Charbit A, Raynaud C: **Role of FliF and FliI of *Listeria monocytogenes* in flagellar assembly and pathogenicity.** *Infect Immun* 2005, **73:**5530-5539.

49. Way SS, Thompson LJ, Lopes JE, Hajjar AM, Kollmann TR, Freitag NE, Wilson CB: **Characterization of flagellin expression and its role in *Listeria monocytogenes* infection and immunity.** *Cell Microbiol* 2004, **6:**235-242.

50. Finn RD, Tate J, Mistry J, Coggill PC, Sammut SJ, Hotz H-R, Ceric G, Forslund K, Eddy SR, Sonnhammer ELL, Bateman A: **The Pfam protein families database.** *Nucl Acids Res* 2008, **36:**D281-288.

51. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR: **Highly integrated single-base resolution maps of the epigenome in Arabidopsis.** *Cell* 2008, **133:**523-536.

52. Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, McDonald H, Varhol R, Jones S, Marra M: **Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing.** *BioTechniques* 2008, **45:**81-94.

53. Chai Y, Kolter R, Losick R: **A widely conserved gene cluster required for lactate utilization in *Bacillus subtilis* and its involvement in biofilm formation.** *J Bacteriol* 2009, **191:**2423-2430.

54. Hillerich B, Westpheling J: **A new GntR family transcriptional regulator in *Streptomyces coelicolor* is required for morphogenesis and antibiotic production and controls transcription of an ABC transporter in response to carbon source.** *J Bacteriol* 2006, **188:**7477-7487.

55. Ogasawara H, Ishida Y, Yamada K, Yamamoto K, Ishihama A: **PdhR (pyruvatedehydrogenase complex regulator) controls the respiratory electron transport system in *Escherichia coli*.** *J Bacteriol* 2007, **189:**5534-5541.

56. Kim H, Marquis H, Boor KJ: **σ^B contributes to *Listeria monocytogenes* invasion by controlling expression of *inlA* and *inlB*.** *Microbiology (Read)* 2005, **151:**3215-3222.

57. Lecuit M, Ohayon H, Braun L, Mengaud J, Cossart P: **Internalin of *Listeria monocytogenes* with an intact leucine-rich repeat region is sufficient to promote internalization.** *Infect Immun* 1997, **65:**5309-5319.

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

59. Disson O, Grayo S, Huillet E, Nikitas G, Langa-Vives F, Dussurget O, Ragon M, Le Monnier A, Babinet C, Cossart P, Lecuit M: **Conjugated action of two**

species-specific invasion proteins for fetoplacental listeriosis. *Nature* 2008, **455:**1114-1118.

60. Ferreira A, O'Byrne CP, Boor KJ: **Role of σ^B in heat, ethanol, acid, and oxidative stress resistance and during carbon starvation in *Listeria monocytogenes*.** *Appl Environ Microbiol* 2001, **67:**4454-4457.

61. Moorhead SM, Dykes GA: **The role of the *sigB* gene in the general stress response of *Listeria monocytogenes* varies between a strain of serotype 1/2a and a strain of serotype 4c.** *Curr Microbiol* 2003, **46:**461-466.

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

63. Sleator RD, Clifford T, Hill C: **Gut osmolarity: A key environmental cue initiating the gastrointestinal phase of *Listeria monocytogenes* infection?** *Med Hypoth* 2007, **69:**1090-1092.

64. Sue D, Boor KJ, Wiedmann M: **σ^B-dependent expression patterns of compatible solute transporter genes *opuCA* and lmo1421 and the conjugated bile salt hydrolase gene *bsh* in *Listeria monocytogenes*.** *Microbiology (Read)* 2003, **149:**3247-3256.

65. Sue D, Fink D, Wiedmann M, Boor KJ: **σ^B-dependent gene induction and expression in *Listeria monocytogenes* during osmotic and acid stress conditions simulating the intestinal environment.** *Microbiology (Read)* 2004, **150:**3843-3855.

66. Core LJ, Waterfall JJ, Lis JT: **Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters.** *Science* 2008, **322:**1845-1848.

67. Cetin MS, Zhang C, Hutkins RW, Benson AK: **Regulation of transcription of compatible solute transporters by the general stress sigma factor, σ^B, in *Listeria monocytogenes*.** *J Bacteriol* 2004, **186:**794-802.

68. Fraser KR, Sue D, Wiedmann M, Boor K, O'Byrne CP: **Role of σ^B in regulating the compatible solute uptake systems of *Listeria monocytogenes*: osmotic induction of *opuC* is σ^B dependent.** *Appl Environ Microbiol* 2003, **69:**2015-2022.

69. Kingsford C, Ayanbule K, Salzberg S: **Rapid, accurate, computational discovery of Rho-independent transcription terminators illuminates their relationship to DNA uptake.** *Genome Biol* 2007, **8:**R22.

70. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B: **Artemis: sequence visualization and annotation.** *Bioinformatics (Oxf)* 2000, **16:**944-945.

71. Storey JD, Tibshirani R: **Statistical significance for genomewide studies.** *P N A S U S A* 2003, **100:**9440-9445.

72. Crooks GE, Hon G, Chandonia JM, Brenner SE: **WebLogo: a sequence logo generator.** *Genome Res* 2004, **14:**1188-1190.

73. Rice P, Longden I, Bleasby A: **EMBOSS: the european molecular biology open software suite.** *Trends Genet* 2000, **16:**276-277.

74. Benjamini Y, Hochberg Y: Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc B* 1995, 57:289-300.

CHAPTER FOUR

Conclusions

The widespread presence of *Listeria monocytogenes* in diverse environments, including those that are natural (i.e., non-agricultural), agricultural, and food-associated, suggests that these environments may serve as sources or reservoirs of *L. monocytogenes* that can be transmitted to various hosts, including humans. Phylogenetic analysis of *L. monocytogenes* strains has identified three distinct lineages which form distinct, but overlapping populations. While lineages I and II are both common among human clinical and food isolates, lineage I strains are overrepresented among human clinical isolates, and lineage II strains are overrepresented among food and environmental isolates. Lineage III, which includes subgroups IIIA and IIIB, is rare and predominantly associated with animal disease. σ^B, encoded by *sigB*, is a sigma factor previously demonstrated to critically contribute to stress response and virulence.

Prior to this study, our understanding of the role of σ^B in stress response and virulence in *L. monocytogenes* was almost exclusively limited to lineage II strains leaving a significant disparity in our understanding of stress response and virulence mechanisms modulated by σ^B and our knowledge of *L. monocytogenes* lineages propensity to cause disease. We used transcriptomic and phenotypic analyses to characterize the role of σ^B in *L. monocytogenes* strains representing lineages I, II, IIIA, and IIIB. Specifically, whole-genome expression microarrays, acid and oxidative stress resistance assays, a Caco-2 invasion efficiency model, and the guinea pig gastrointestinal model for listeriosis were used to characterize the role of σ^B stationary

phase wildtype and isogenic Δ*sigB* mutants representing *L. monocytogenes* diversity. While σ^B only contributes to *in vitro* intestinal epithelial cell invasion and environmental stress in some strains, it contributes to guinea pig virulence in all *L. monocytogenes* strains tested, further supporting strain specific contributions of σ^B to gene regulation in *L. monocytogenes* virulence. Phenotypic diversification provides species with a capacity to survive environmental adversity and is a key player in niche adaptation (18) and therefore may partially contribute to differences in *L. monocytogenes* strains' abilities to cause disease.

The development of powerful new DNA sequencing technologies has recently yielded new tools which have dramatically revolutionized scientific approaches to biological questions, including transcriptomics. In an effort to harness high-throughput, deep DNA sequencing technology for prokaryotic transcriptomics, we used *L. monocytogenes* as a model system to explore the application of RNA-Seq for the dual purposes of genome-wide transcriptome characterization in a bacterial pathogen and comprehensive quantification of target gene expression for the alternative sigma factor, σ^B. Specifically, transcriptomes were compared between stationary phase cells of *L. monocytogenes* wildtype 10403S and an otherwise isogenic Δ*sigB* mutant, which does not express the alternative sigma factor σ^B, a major regulator of genes contributing to stress response. Overall, 83% of all genes were transcribed under these conditions including 96 genes which had significantly higher transcript levels in 10403S than in Δ*sigB*, indicating σ^B-dependent transcription of these genes. RNA-Seq analyses identified a number of noncoding RNA molecules (ncRNAs) including 53 which had not been previously described in *L. monocytogenes*.

Holistically, the results from these studies suggest that σ^B contributes to a complex network of transcriptional regulators which helps *L. monocytogenes* survive stress and subsequently cause disease. As the σ^B regulon appears to differ among

strains representing *L. monocytogenes* diversity, technologies such as RNA-Seq will allow us to further explore these differences, both quantitatively and qualitatively. Further, high-throughput transcriptomic strategies such as RNA-Seq will be essential to our understanding of transcriptional regulatory networks in bacteria and will play a role in initial rapid characterization of novel and emerging bacterial pathogens.

APPENDIX ONE

Supplemental figures

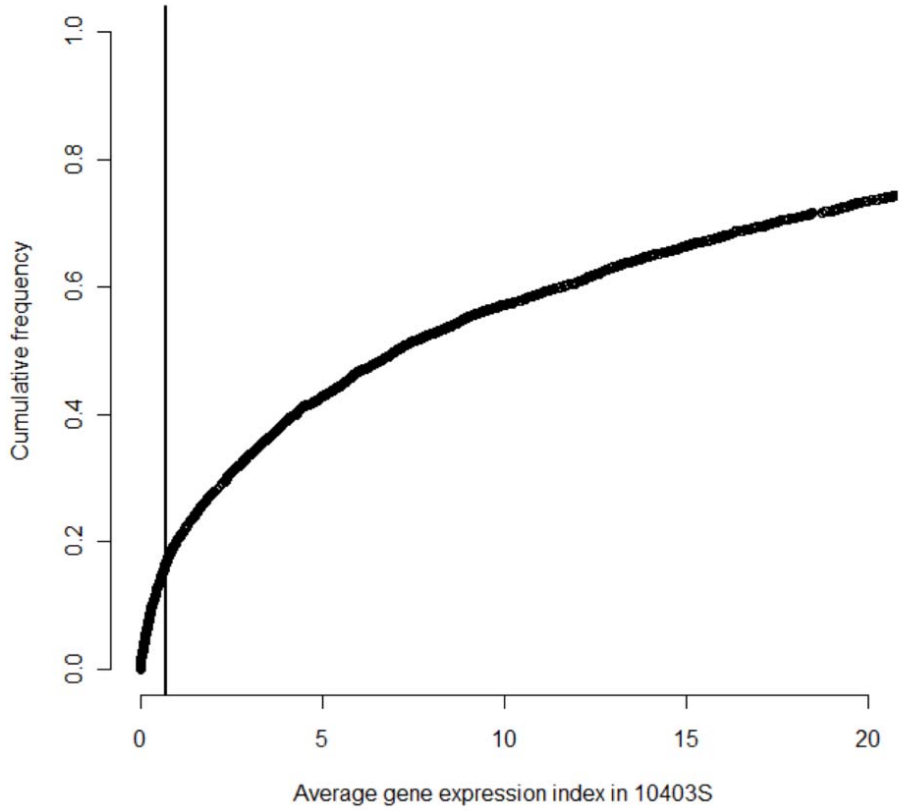


Figure A1 [S3.1]. Cumulative frequency of average GEI in *L. monocytogenes* 10403S. The vertical line indicates an average GEI of 0.7 reads, which is the cut-off used to identify transcription. The graph shows that about 83% of the genes fall at the right of the average GEI cut-off of 0.7 reads and were therefore considered transcribed.

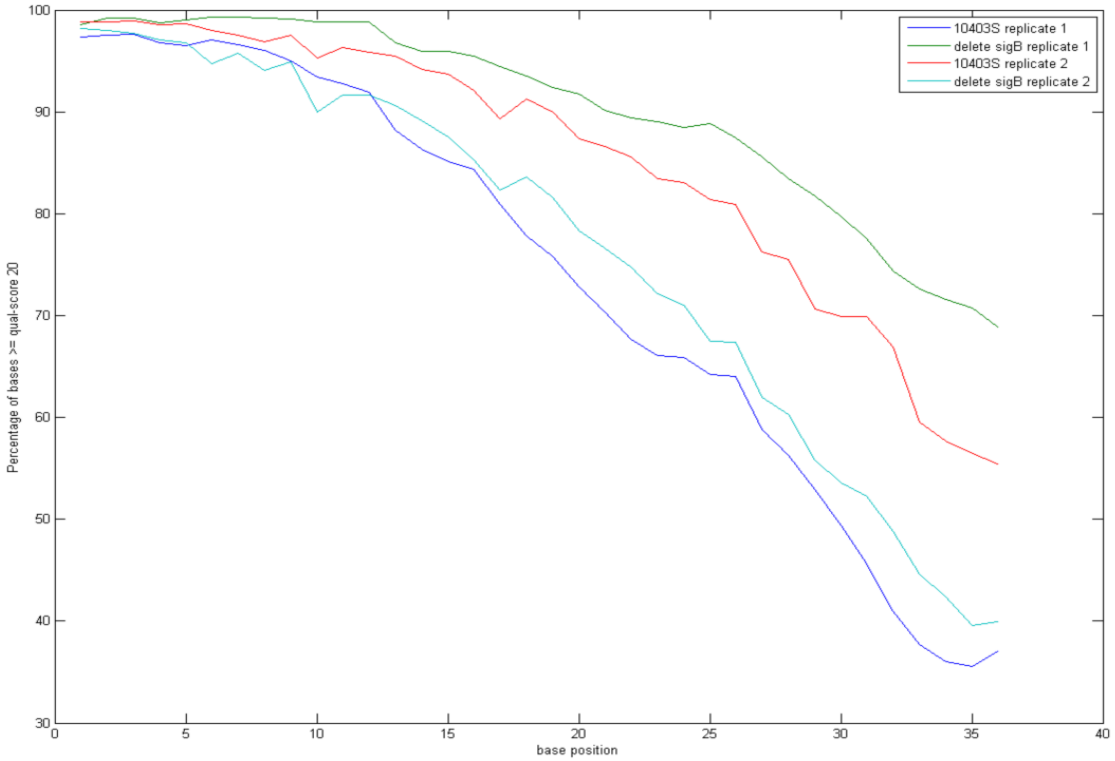


Figure A1 [S3.2] Distribution of quality scores for all RNA-Seq runs. The quality of the RNA-Seq reads was analyzed using the correspondence between the quality score and error probability; these analyses were performed on Illumina RNA-Seq quality scores that were converted to phred format [http://www.phrap.com/phred/].

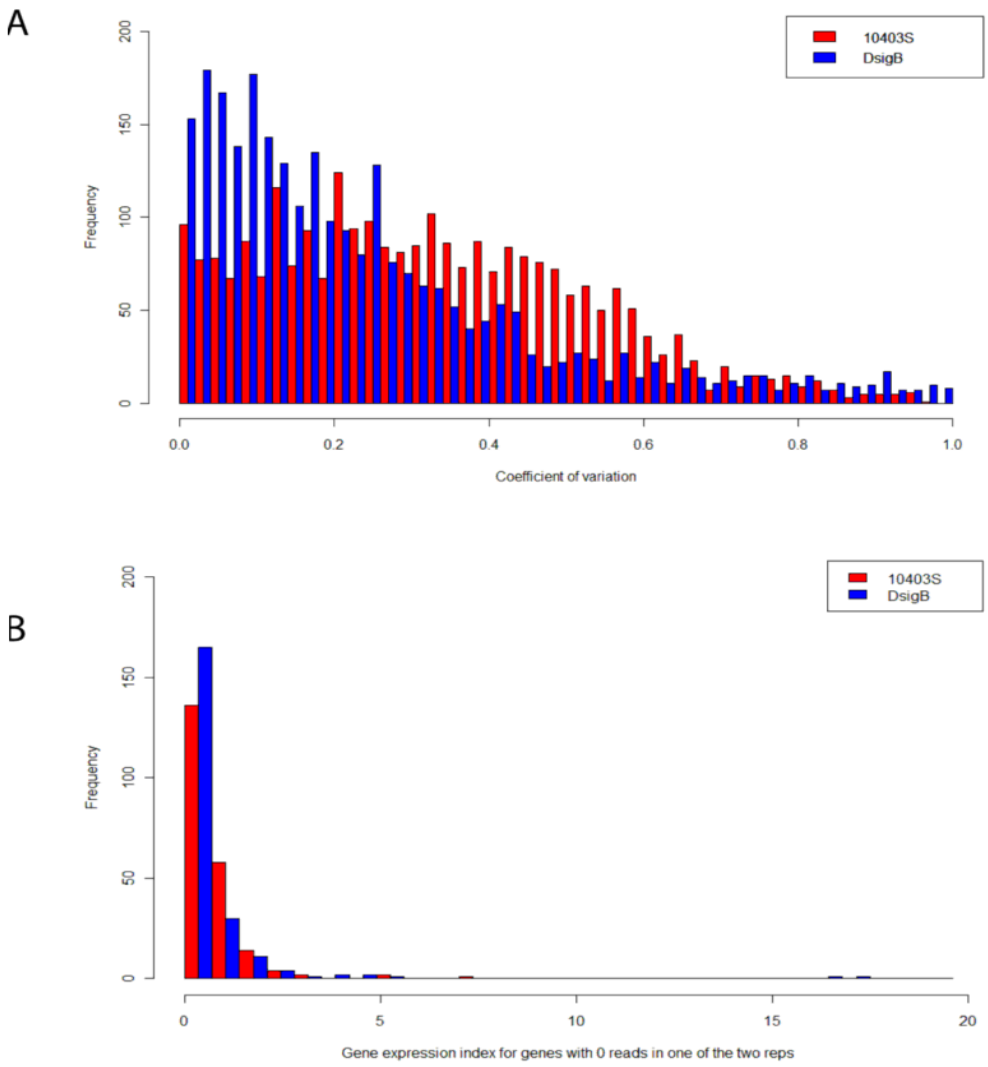


Figure A1 [S3.3]. Coefficient of variation among RNA-Seq replicates by strain. (A) Histogram of the coefficient of variation (standard deviation/mean) for genes with GEI > 0 in both replicates for 10403S and Δ*sigB* strain. There is less variation between Δ*sigB* replicates compared to the 10403S replicates, but very few genes have a coefficient > 0.6. (B) Histogram depicting the GEI of one replicate for genes where the other replicate GEI = 0. The replicate GEI of the gene for which the other replicate is 0 (zero) is typically very low (GEI < 0.7).

APPENDIX TWO

Supplemental Tables

Table A2 [S2.1] Splice by Overlap Extension PCR (SOEing-PCR) primers for mutant creation

Primer	Sequence (5'→3')
HO-01 SoeA	CGG GAT CCA GGT AGA CTT TCA TTA TCA GG
HO-02 SoeB	ATA GAC TTT TTC TTT CGC CTC
HO-03 SoeC	GAG GCG AAA GAA AAA CTC TAT TTG CAG AAT GAG GAA GTG GA
HO-04 SoeD	GGA AAT CAC CGA TAT TTC TGT TTT CGA C
HO-05 sigB external F	TAC ATT ACA ACT TCC TGC CAA
HO-06 sigB external R	GCA TTA AAA ACT TAC TGC CTG

Table A2 [S2.2] TaqMan qRT-PCR primers and probes

Gene	Common Name	Forward Primer (5'→3')	MGB Probe (6Fam 5'→3' NFQ)	Reverse Primer (5'→3')
rpoB	RNAP subunit	CCG GAC GTC ACG GTA ACA A	TTA TCT CCC GTA TTT TAC C	CAG GTG TTC CGT CTG GCA TA
opuCA	carnitine transport facilitator	ACA TCG ATA AAG GAG AAT TTG TTT GTT	TCG TTT TCC CAC AAC CA	GCC GGT TAA TCA TCT TCA TTG TT
lmo0433	Internalin A	GAC AAA TGC TCA GGC AGC TAC A	CAG CTC TAG CGG AAA A	TTT GCG AGA CCG TGT CTG TTA
lmo0398	similar to phosphotransferase system enzyme IIA	CGC GAA AAA GAA TAC CGA ACA	CAG TAC AAC AAT TAA TCG	TGA CTC CCA ATC AAT CGG TTC
lmo1539	similar to glycerol uptake facilitator	CGG GAT GTC TGT TGG TGG AA	TCA ACC CAG CTC GTG	AAA TCG GCC AAA CAA AAT GC
lmo2668	similar to transcriptional antiterminator BglG family	CCG TGC ATT ACA CG TGA AAG A	TCG TTA GCA CCA TTA GAA	TTC ATT TTC AGG AAC TTT GCT ACC A

Table A2 [S2.3]. σ^B dependent genes differentially expressed in at least one lineage representative

lmo[a]	Common Name[b]	Lineage I fold change[c]	Lineage I adj p-value[d]	Lineage II fold change[c]	Lineage II adj p-value[d]	Lineage IIIA fold change[c]	Lineage IIIA adj p-value[d]	Lineage IIIB fold change[c]	Lineage IIIB adj p-value[d]
lmo0019	unknown	1.4	0.1649	2.9	0.0000	3	0.0016	3.1	0.0005
lmo0024	similar to PTS system, mannose-specific IID component	0.9	0.9770	1.1	0.7034	1.9	0.2258	2	0.0336
lmo0027	similar to PTS system, beta-glucosides specific enzyme IIABC	2	0.0108	2	0.0028	1.3	0.4911	1.2	0.5825
lmo0033	similar to endoglucanase	0.5	0.4029	0.7	0.0377	0.8	0.8310	49.8	0.0171
lmo0043	similar to arginine deiminase	1.9	0.3567	3.7	0.0000	0.9	0.9592	1.3	0.7670
lmo0044	ribosomal protein S6	1.4	0.1275	1	0.9867	1.7	0.0075	1.6	0.0362
lmo0046	ribosomal protein S18	1.1	0.8775	0.8	0.0999	1.5	0.0465	1.3	0.3494
lmo0093	similar to ATP synthase epsilon chain	.	.	1.8	0.0008	3.7	0.0000	3.7	0.0001
lmo0100	unknown	1.2	0.7976	1.5	0.0155	1	0.9851	0.7	0.2295
lmo0109	similar to transcriptional regulatory proteins, AraC family	1.3	0.3916	1.3	0.0234	1.1	0.8045	1.8	0.0056
lmo0130	similar to 5-nucleotidase, putative peptidoglycan bound protein (LPXTG motif)	2.1	0.0046	1.3	0.0764	0.6	0.1548	0.9	0.7956
lmo0133	similar to E. coli YjdI protein	4.5	0.0003	5.1	0.0000	9	0.0000	14.5	0.0000
lmo0134	similar to E. coli YjdJ protein	6.9	0.0001	5.6	0.0000	14.3	0.0000	7.7	0.0001
lmo0135	similar to oligopeptide ABC transport system substrate-binding proteins	0.8	0.2953	2	0.0000	0.6	0.0216	0.5	0.0197
lmo0136	similar to oligopeptide ABC transporter, permease protein	0.6	0.1425	1.5	0.0000	0.6	0.0646	0.6	0.0196

lmo0142	unknown	1	0.9873	1.1	0.7775	1.9	0.0361	1.7	0.0877
lmo0169	glucose uptake protein	5.5	0.0002	3.4	0.0000	5.1	0.0000	6.9	0.0000
lmo0170	unknown	3.6	0.0006	3	0.0008	5.6	0.0000	7.4	0.0000
lmo0188	dimethyladenosine transferase (16S rRNA dimethylase)	1.7	0.0348	1.1	0.3556	1	0.9848	0.8	0.3281
lmo0207	hypothetical lipoprotein	0.9	0.7441	1.7	0.0004	1.2	0.4646	1.4	0.2036
lmo0210	similar to L-lactate dehydrogenase	4.5	0.0000	2.6	0.0000	3.8	0.0001	4.6	0.0058
lmo0211	similar to B. subtilis general stress protein	1.7	0.0335	1.5	0.0001	2.1	0.0021	2.4	0.0001
lmo0217	similar to B. subtilis DivIC protein	1.8	0.0167	1.2	0.0230	1.1	0.8404	1.1	0.8456
lmo0232	endopeptidase Clp ATP-binding chain C	1.3	0.1100	1	0.8241	2	0.0009	1.9	0.0430
lmo0239	cysteinyl-tRNA synthetase	1.9	0.0244	1.1	0.4018	1	0.9776	1	0.9461
lmo0264	internalin C2	1.6	0.3567	5.3	0.0000	15.4	0.0013	1.3	0.3629
lmo0265	peptidase, M20/M25/M40 family	1.5	0.7760	8.2	0.0000	23.1	0.0000	28.5	0.0000
lmo0274	unknown	1.7	0.0786	2	0.0000	1.4	0.6115	.	.
lmo0292	similar to heat-shock protein htrA serine protease	1.4	0.1117	1.6	0.0017	1.3	0.0315	1.5	0.0140
lmo0315	similar to thiamin biosynthesis protein	2.5	0.0395	1	0.9741	.	.	0.5	0.7863
lmo0321	similar to unknown proteins	3.5	0.1010	5.3	0.0000	3.3	0.0649	5.7	0.0228
lmo0341	unknown	0.8	0.7323	1.5	0.0119	2.2	0.1111	1.6	0.2453
lmo0342	similar to transketolase	1.5	0.5712	1.6	0.0185	2.6	0.0094	4.3	0.0071
lmo0343	similar to transaldolase	1.1	0.9782	2	0.0011	3.2	0.0009	5.6	0.0058
lmo0344	similar to dehydrogenase/reductase	1.3	0.8013	1.8	0.0083	2.6	0.0075	5.6	0.0129
lmo0345	similar to sugar-phosphate isomerase	1.3	0.8159	1.5	0.0002	1.6	0.0380	2.8	0.0041
lmo0346	similar to triosephosphate isomerase	1.3	0.8578	1.9	0.0062	8.1	0.0040	1.6	0.3419

lmo0347	similar to dihydroxyacetone kinase	1.6	0.5761	2.1	0.0025	1	0.9848	5.3	0.0402
lmo0348	similar to dihydroxyacetone kinase	1.2	0.9272	1.8	0.0098	3.8	0.0095	3.3	0.0111
lmo0373	similar to PTS betaglucoside-specific enzyme IIC component	1.1	0.9418	1.2	0.3941	1.2	0.6630	1.6	0.0226
lmo0382	transcriptional regulator, DeoR family	1.5	0.2146	1.2	0.1674	1.6	0.0185	1.4	0.0390
lmo0397	similar to unknown proteins	1.2	0.6979	1	0.9498	1.3	0.1027	1.6	0.0235
lmo0398	similar to phosphotransferase system enzyme IIA	0.7	0.1649	13.5	0.0000	3.3	0.0761	3.6	0.0102
lmo0399	similar to fructose-specific phosphotransferase enzyme IIB	0.9	0.8049	10.5	0.0000	3.2	0.1515	2.5	0.0578
lmo0400	PTS system, fructose subfamily, IIC component subfamily	1	0.9922	20.5	0.0000	3.4	0.0123	1.9	0.0148
lmo0401	highly similar to E. col YbgG protein, a putative sugar hydrolase	0.8	0.3601	13.4	0.0000	2.4	0.0223	2.9	0.0182
lmo0402	similar to transcriptional antiterminator (BglG family)	1.5	0.1404	15.7	0.0000	1.4	0.5206	1.2	0.5420
lmo0405	phosphate transporter family protein	2.7	0.0410	1.7	0.0106	2.1	0.0053	2.1	0.0036
lmo0406	similar to B. subtilis YyaH protein	1.1	0.9866	1.6	0.0014	1.2	0.7135	1.5	0.2314
lmo0408	unknown	1.9	0.0305	1.5	0.0003	1.4	0.1089	1.6	0.0569
lmo0411	similar to phosphoenolpyruvate synthase (N-terminal part)	1.2	0.7215	1.7	0.0001	0.9	0.6474	0.9	0.8888
lmo0429	similar to sugar hydrolase	0.8	0.4988	0.6	0.0008	1.5	0.0487	1.3	0.5493
lmo0433	Internalin A	3.6	0.0004	3.1	0.0000	4.8	0.0000	7	0.0000
lmo0434	internalin B	1.7	0.0459	1.5	0.0361	1.3	0.5999	1.4	0.5644
lmo0438	conserved hypothetical protein	0.5	0.2095	1.1	0.8162	1.6	0.0399	1.5	0.5342
lmo0439	weakly similar to a module of peptide synthetase	3.8	0.1291	5.6	0.0000	1.1	0.9158	3.3	0.1302

lmo0445	similar to transcription regulator	.	.	5	0.0000	6.9	0.0277	12.6	0.0042
lmo0449	unknown	1.7	0.4564	1.5	0.0238	0.3	0.0038	0.9	0.7279
lmo0459	similar to transcription regulator (VirR from Streptococcus pyogenes)	2.5	0.0101	0.9	0.8646	0.8	0.7361	1	0.9918
lmo0468	unknown	0.5	0.7695	1.2	0.6799	1.7	0.0153	2	0.1525
lmo0503	similar to PTS fructose-specific enzyme IIA component	0.9	0.9624	1.2	0.4249	1.9	0.0841	1.9	0.0448
lmo0507	similar to PTS system, Galactitol-specific IIB component	0.9	0.8544	1.2	0.0945	2.3	0.0191	2.2	0.0380
lmo0515	conserved hypothetical protein	3.3	0.0116	3.4	0.0002	3.5	0.0000	5.4	0.0000
lmo0524	similar to putative sulfate transporter	2.5	0.0037	2.3	0.0000	2.7	0.1648	7.3	0.0027
lmo0539	putative tagatose 1,6-diphosphate aldolase	14.4	0.0000	7.5	0.0000	19.5	0.0000	24.2	0.0000
lmo0554	similar to NADH-dependent butanol dehydrogenase	2.3	0.0380	6	0.0000	1.2	0.6752	5.2	0.0128
lmo0555	proton-dependent oligopeptide transporter	7.2	0.0002	4.1	0.0000	7.1	0.0000	9.3	0.0000
lmo0576	hypothetical cell wall associated protein	5.1	0.0005	1.1	0.7666	1.4	0.5942	1.1	0.9377
lmo0579	similar to unknown protein	2	0.2506	1.5	0.0049	1.7	0.0064	1.7	0.0122
lmo0580	weakly similar to carboxylesterase	1.5	0.1056	1.6	0.0013	0.9	0.8473	1.1	0.9130
lmo0584	conserved hypothetical membrane protein	1.2	0.3709	1.7	0.0001	0.9	0.8715	1.7	0.0196
lmo0589	unknown	1.7	0.1977	1.7	0.0079	1.7	0.0051	1.9	0.0181
lmo0590	similar to a fusion of two types of conserved hypothetical proteinconserved hypothetical	1.1	0.9228	1.7	0.0003	1.3	0.7512	1.5	0.2520

lmo0591	similar to unknown membrane proteins	1.2	0.4479	1.5	0.0166	1.6	0.0404	2	0.0014
lmo0593	similar to transport proteins (formate?)	6	0.0032	5.7	0.0000	9	0.0003	18.2	0.0000
lmo0596	similar to unknown proteins	14.2	0.0001	22.7	0.0000	33.8	0.0000	36	0.0000
lmo0602	weakly similar to transcription regulator	2.6	0.0023	3.7	0.0000	2.8	0.0003	2.5	0.0121
lmo0610	similar to internalin proteins, putative peptidoglycan bound protein (LPXTG motif)	1.9	0.0380	3.7	0.0000	5.5	0.0000	8.2	0.0000
lmo0620	conserved hypothetical protein	0.9	0.9415	1.6	0.0005	0.9	0.9494	1.1	0.8307
lmo0626	similar to unknown protein	1.4	0.8736	2	0.0005	1	0.9776	1.2	0.5948
lmo0628	unknown	2.9	0.2091	3.3	0.0000	6.3	0.0000	9.4	0.0000
lmo0629	unknown	1.8	0.1730	1.6	0.0000	4.4	0.0000	5.1	0.0000
lmo0640	similar to oxidoreductase	1.8	0.0081	1.2	0.0748	1	0.8897	1.2	0.4460
lmo0641	similar to heavy metal-transporting ATPase	0.7	0.1613	1	0.9196	1.5	0.0174	1.2	0.4602
lmo0642	membrane protein, putative	3.4	0.0028	2	0.0002	2	0.0029	3.1	0.0005
lmo0647	unknown	1.4	0.1241	2	0.0003	0.6	0.6432	1	0.9710
lmo0648	similar to membrane proteins	1.7	0.2015	1.7	0.0014	1	0.9771	1.8	0.1468
lmo0649	similar to transcription regulators	2.7	0.1298	1.7	0.0004	1.5	0.0024	2	0.0070
lmo0650	conserved membrane protein	1.8	0.0763	1.9	0.0000	1.6	0.0075	2	0.0023
lmo0654	unknown	3.5	0.0109	3.8	0.0000	1.1	0.7016	1.4	0.6052
lmo0655	similar to phosphoprotein phosphatases	4.4	0.0000	2.9	0.0000	2.5	0.0014	3.1	0.0007
lmo0669	similar to oxidoreductase	2.4	0.0189	15.8	0.0000	1.9	0.1753	6.6	0.0005
lmo0670	conserved hypothetical protein	5.1	0.0001	11.1	0.0000	2.1	0.1022	3.5	0.0959
lmo0722	similar to pyruvate oxidase	4.6	0.0002	5.4	0.0000	8.8	0.0000	3.8	0.0070
lmo0758	unknown	0.8	0.5846	1.1	0.9181	1.6	0.0133	1.9	0.0162

lmo0759	unknown	0.8	0.5762	1.6	0.0049	2	0.0148	2.1	0.0256
lmo0760	unknown	0.9	0.5681	1.6	0.0016	1.5	0.0418	1.7	0.0071
lmo0761	similar to unknown proteins	0.9	0.9563	1.3	0.0494	2.1	0.0001	2.2	0.0014
lmo0775	unknown	0.5	0.0019	0.8	0.0942	1.6	0.0027	1.5	0.0804
lmo0781	similar to mannose-specific phosphotransferase system (PTS) component IID	10	0.0000	15.6	0.0000	18.2	0.0000	25.4	0.0000
lmo0782	similar to mannose-specific phosphotransferase system (PTS) component IIC	12.9	0.0000	13.5	0.0000	20.3	0.0000	22	0.0000
lmo0783	similar to mannose-specific phosphotransferase system (PTS) component IIB	6.4	0.0000	12	0.0000	14.8	0.0000	18	0.0000
lmo0784	similar to mannose-specific phosphotransferase system (PTS) component IIA	2	0.0341	5.7	0.0000	5.5	0.0000	6.4	0.0000
lmo0791	unknown	1.3	0.7916	1.5	0.0197	1.6	0.1791	1.6	0.0577
lmo0794	similar to B. subtilis YwnB protein	6.1	0.0007	12.8	0.0000	10.3	0.0000	25.4	0.0000
lmo0796	conserved hypothetical protein	1.9	0.0029	4	0.0000	8.3	0.0000	12.3	0.0000
lmo0811	similar to carbonic anhydrase	1.7	0.1107	1.7	0.0002	0.8	0.5999	0.8	0.1598
lmo0818	similar to cation transporting ATPase	0.7	0.4525	1.7	0.0023	1.1	0.8391	1.2	0.4886
lmo0819	unknown	1.5	0.0833	1.7	0.0010	0.9	0.8892	1.4	0.3776
lmo0850	hypothetical protein	1	0.9959	1.3	0.0156	1.9	0.0155	1.5	0.1957
lmo0869	unknown	1.4	0.2325	1.3	0.0754	2.4	0.0471	2.7	0.0778
lmo0870	unknown	1.3	0.8077	1.5	0.0410	1.5	0.6078	1.7	0.2325
lmo0871	similar to B. subtilis YtcD protein	1	0.9951	1.1	0.2294	1.3	0.1045	1.5	0.0254
lmo0874	similar to PTS system enzyme IIA component	0.4	0.2219	1.3	0.4824	2.4	0.0040	2	0.0274

lmo0880	similar to wall associated protein precursor (LPXTG motif)	7.4	0.0000	6.7	0.0000	16.4	0.0000	6.8	0.0006
lmo0893	anti-anti-sigma factor (antagonist of RsbW)	0.9	0.8458	1.5	0.0024	1.1	0.8655	1.3	0.4856
lmo0894	sigma-B activity negative regulator RsbW	1.5	0.0947	1.4	0.0184	2	0.0404	2.4	0.0005
lmo0896	Indirect negative regulation of sigma B dependant gene expression (serine phosphatase)	1.5	0.0244	2.3	0.0000	1.7	0.0197	2.7	0.0017
lmo0905	unknown	1.6	0.0703	1	0.9176	1.5	0.0118	1.2	0.6502
lmo0907	similar to phosphoglycerate mutase	0.9	0.7117	1	0.8845	1.3	0.2384	1.7	0.0061
lmo0911	unknown	2.1	0.0152	2.1	0.0000	9.2	0.0000	1.8	0.0200
lmo0913	succinate-semialdehyde dehydrogenase	6.6	0.0001	13.4	0.0000	16	0.0000	22.3	0.0000
lmo0927	hypothetical transmembrane protein	1.4	0.2060	1.6	0.0009	1.2	0.5328	0.9	0.8382
lmo0928	similar to 3-methyladenine DNA glycosylase	2	0.1498	1.5	0.0150	1.9	0.1381	1.7	0.1968
lmo0929	similar to sortase	1.8	0.0663	1.6	0.0032	1.9	0.0174	1.7	0.1598
lmo0937	unknown	6.4	0.0001	10.4	0.0000	16.2	0.0000	18.9	0.0000
lmo0942	histidine kinase domain protein	4.4	0.0146	1	0.9595	0.9	0.8325	0.9	0.7890
lmo0944	similar to B. subtilis YneR protein	1.7	0.0533	1.5	0.0004	2	0.0000	2.5	0.0000
lmo0953	unknown	3.2	0.0156	6.5	0.0000	12	0.0000	16.3	0.0000
lmo0956	N-acetylglucosamine-6-phosphate deacetylase	2.6	0.0073	2.1	0.0000	4.4	0.0000	4.3	0.0000
lmo0957	similar to glucosamine-6-Phoasphate isomerase (EC 5.3.1.10)	2.3	0.0006	1.6	0.0027	3.1	0.0027	2.4	0.0441
lmo0958	similar to transcription regulator	1.7	0.0082	1.4	0.0100	2.3	0.0001	2.3	0.0002

	(GntR family)								
lmo0959	similar to undacaprenyl-phosphate N-acetylglucosaminyltransferase	1.8	0.0398	1.1	0.3736	1.4	0.1609	1.2	0.8463
lmo0994	unknown	24.4	0.0000	14.1	0.0000	54.2	0.0000	79.2	0.0000
lmo0995	membrane protein, putative	1.1	0.9752	3.4	0.0001	1.6	0.7781	0.7	0.8014
lmo1025	unknown	1.3	0.7606	1.2	0.7530	2	0.0476	0.8	0.6849
lmo1037	membrane protein, putative	1	0.9873	1.6	0.0064	0.7	0.1667	1.2	0.7450
lmo1064	hypothetical protein	1.4	0.7066	1.7	0.0002	0.9	0.8325	0.9	0.8020
lmo1068	unknown	1.1	0.8445	1.4	0.0060	1.8	0.0002	1.9	0.0008
lmo1072	pyruvate carboxylase	1.2	0.5247	1.5	0.0062	0.6	0.0439	0.5	0.0012
lmo1076	similar to autolysin (EC 3.5.1.28) (N-acetylmuramoyl-L-alanine amidase)	1.6	0.0264	1.2	0.2002	1.1	0.9592	.	.
lmo1121	unknown	1.9	0.0539	2.4	0.0000	6.2	0.0000	7.2	0.0006
lmo1122	unknown	0.9	0.9450	0.9	0.7816	0.6	0.7698	1.8	0.0441
lmo1123	conserved hypothetical protein	1.3	0.7392	0.9	0.5761	1.8	0.0235	1.6	0.0848
lmo1139	unknown	0.9	0.8641	1.5	0.0188	1.9	0.2767	1.5	0.2985
lmo1140	unknown	4	0.0084	3.5	0.0000	4.9	0.0001	4.9	0.0000
lmo1189	similar to transcriptional regulator	0.5	0.0136	0.6	0.0021	1.4	0.1271	1.6	0.0133
lmo1194	similar to cobalamin biosynthesis protein CbiD	0.7	0.8687	1.1	0.7917	1.9	0.0480	1.3	0.4672
lmo1220	similar to unknown protein	0.7	0.1252	1	0.9237	1.4	0.0248	1.6	0.0091
lmo1226	similar to transporter, (to B. subtilis YdgH protein)	1.2	0.5341	1.5	0.0023	0.7	0.3845	1	0.9654
lmo1237	glutamate racemase	1.7	0.0324	1	0.8658	0.9	0.8324	0.8	0.2966
lmo1241	unknown	2.7	0.0012	2.2	0.0002	4.2	0.0000	5	0.0000
lmo1242	conserved hypothetical protein	1.1	0.8038	1.6	0.0004	1.1	0.5067	1	0.8588
lmo1243	unknown	0.9	0.8989	1.5	0.0117	0.7	0.3816	0.7	0.5351
lmo1254	similar to alpha,alpha-	1.8	0.0257	1.2	0.0124	0.8	0.1592	1.8	0.0020

	phosphotrehalase								
lmo1255	PTS system, trehalose-specific, IIBC component	2.7	0.0227	1.3	0.0495	1.2	0.4237	1.2	0.2841
lmo1261	unknown	4.2	0.0116	3	0.0000	1.3	0.5213	2.4	0.0906
lmo1291	similar to acyltransferase (to B. subtilis YrhL protein)	1.1	0.7978	1.1	0.5918	1.7	0.0011	1.7	0.4517
lmo1293	similar to glycerol 3 phosphate dehydrogenase	1.8	0.0200	0.5	0.0026	5.7	0.0000	10.8	0.0000
lmo1295	similar to host factor-1 protein	1.8	0.0120	3.4	0.0000	5.4	0.0000	8	0.0000
lmo1340	similar to B. subtilis YqgU protein	1.4	0.3531	1.3	0.0140	2.1	0.0005	2.3	0.0019
lmo1348	similar to aminomethyltransferase	1.7	0.0447	0.9	0.1346	0.5	0.0061	1.4	0.6250
lmo1355	highly similar to elongation factor P (EF-P)	1	0.9624	1	0.8227	1.5	0.0111	1.2	0.6339
lmo1356	acetyl-CoA carboxylase, biotin carboxyl carrier protein	2.2	0.0350	1.7	0.0014	0.9	0.5408	0.8	0.2257
lmo1357	acetyl-CoA carboxylase subunit (biotin carboxylase subunit)	1.9	0.0290	1.3	0.0092	0.9	0.8212	0.8	0.2208
lmo1360	methylenetetrahydrofolate dehydrogenase/methenyltetrahydrofolate cyclohydrolase	1.6	0.1421	1.6	0.0017	0.7	0.0337	0.7	0.0951
lmo1361	similar to exodeoxyribonuclease VII (large subunit)	1	0.9873	1	0.8256	2	0.0127	1.3	0.6594
lmo1367	similar to arginine repressor	1.1	0.9629	1.2	0.0857	1.3	0.1470	1.5	0.0249
lmo1372	similar to branched-chain alpha-keto acid dehydrogenase E1 subunit (2-oxoisovalerate dehydrogenase alpha subunit)	1.7	0.0394	1.1	0.6516	0.8	0.2843	1.1	0.7593
lmo1375	similar to aminotripeptidase	2.4	0.0168	3	0.0000	2.4	0.0025	2.1	0.0381
lmo1376	similar to 6-phosphogluconate	2.3	0.0146	1.7	0.0001	1.4	0.1139	1	0.9382

	dehydrogenase								
lmo1388	CD4 T cell-stimulating antigen, lipoprotein	1.6	0.1146	1.7	0.0066	0.6	0.0597	0.8	0.6706
lmo1389	similar to sugar ABC transporter, ATP-binding protein	2.2	0.0020	1.2	0.0868	0.6	0.0280	1	0.9506
lmo1390	similar to ABC transporter (permease proteins)	2.2	0.0041	1.2	0.0399	0.6	0.0276	1	0.9853
lmo1391	similar to sugar ABC transporter, permease protein	1.8	0.0187	1.2	0.2869	0.6	0.0661	1.1	0.9390
lmo1394	similar to 3-ketoacyl-acyl carrier protein reductase	1	0.9873	1.1	0.6047	1.5	0.0155	1	0.9779
lmo1421	similar to glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)	1.7	0.1477	2.2	0.0003	2	0.0785	2.4	0.0027
lmo1425	similar to betaine/carnitine/choline ABC transporter (membrane p)	1.9	0.0062	3.4	0.0000	3.5	0.0000	4.7	0.0000
lmo1426	similar to glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein)	1.8	0.0541	3.1	0.0000	3.3	0.0000	4.3	0.0000
lmo1427	similar to glycine betaine/carnitine/choline ABC transporter (membrane protein)	1.3	0.4386	3	0.0000	1.9	0.0150	3.8	0.0004
lmo1428	similar to glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)	2.2	0.0008	2.9	0.0000	2.7	0.0001	3.5	0.0004
lmo1432	unknown	2.8	0.0049	1.9	0.0000	1.4	0.4042	1.8	0.2733
lmo1433	similar to glutathione reductase	4.7	0.0006	4.2	0.0000	3	0.0378	3.7	0.0309
lmo1454	RNA polymerase sigma factor	1.4	0.0698	1.1	0.4693	2.3	0.0609	3.6	0.0042

	RpoD								
lmo1495	similar to unknown proteins	1.7	0.0694	1.4	0.0113	1.5	0.0289	1.1	0.7125
lmo1496	similar to transcription elongation factor GreA	2	0.0728	1.4	0.0081	1.8	0.0277	1.2	0.7261
lmo1526	similar to unknown proteins	2.3	0.0882	4	0.0000	5.2	0.0009	12	0.0000
lmo1527	similar to protein-export membrane protein SecDF	1.8	0.1750	1.5	0.0006	0.8	0.5944	0.7	0.5522
lmo1528	similar to unknown proteins	2.1	0.0574	1.8	0.0002	0.5	0.0085	0.7	0.1088
lmo1532	highly similar to Holliday junction DNA helicase RuvB	1.8	0.0494	1.5	0.0022	0.6	0.0306	0.7	0.1190
lmo1534	similar to L-lactate dehydrogenase	1.5	0.1649	1.6	0.0005	0.9	0.8791	1.3	0.1243
lmo1538	similar to glycerol kinase	2.9	0.0001	0.7	0.0195	2.4	0.0003	4.1	0.0001
lmo1539	similar to glycerol uptake facilitator	3.9	0.0000	0.7	0.0746	3.8	0.0000	5.6	0.0000
lmo1542	ribosomal protein L21	1.5	0.0354	1.4	0.0355	1	0.9708	0.8	0.4844
lmo1545	similar to cell-division inhibition (septum placement) protein MinC	1.4	0.8622	1.5	0.0006	0.6	0.1135	0.7	0.3184
lmo1557	highly similar to glutamyl-tRNA reductase	1.6	0.3681	1.3	0.0060	2.4	0.0300	1.6	0.2059
lmo1570	highly similar to pyruvate kinases	1.8	0.0227	1.1	0.7807	1.3	0.2436	1.3	0.2235
lmo1571	6-phosphofructokinase	1.4	0.3590	1.5	0.0001	1.4	0.2725	0.9	0.8954
lmo1580	similar to unknown protein	1.3	0.2022	2	0.0000	1.2	0.6724	4.3	0.0001
lmo1582	weakly similar to site specific DNA-methyltransferase	1.2	0.3609	0.8	0.0317	1.5	0.0448	1	0.9553
lmo1601	similar to general stress protein	3.2	0.0000	4	0.0000	1.4	0.0213	2.1	0.0010
lmo1602	similar to unknown proteins	4.2	0.0000	4.6	0.0000	1.7	0.0417	2.7	0.0001
lmo1605	similar to UDP-N-acetyl muramate-alanine ligases	7.4	0.0001	2.2	0.0000	2.4	0.0047	2.8	0.0001
lmo1606	FtsK/SpoIIIE family protein	9.1	0.0000	5.6	0.0000	3.5	0.0000	4.1	0.0000

lmo1607	tRNA-binding domain protein	1.1	0.7959	1	0.8060	1	0.9516	1.9	0.0346
lmo1616	similar to unknown proteins	1.6	0.0533	1.1	0.3443	2.2	0.0187	2.3	0.0098
lmo1622	similar to unknown proteins	1.2	0.5714	1.6	0.0011	1.5	0.0076	1.4	0.1507
lmo1635	similar to unknown proteins	2.1	0.1056	1.7	0.0037	1	0.8871	1	0.9192
lmo1636	similar to similar to ABC transporter (ATP-binding protein)	2	0.0936	1.9	0.0001	0.7	0.1312	0.6	0.0269
lmo1637	similar to membrane proteins	1.4	0.2716	1.6	0.0002	0.7	0.0423	0.7	0.0550
lmo1657	translation elongation factor Ts	2.1	0.0042	1.5	0.0014	1.8	0.0009	1.4	0.1043
lmo1658	30S ribosomal protein S2	2.2	0.0029	1.3	0.0530	1.7	0.0011	1.5	0.0591
lmo1663	similar to asparagine synthetase	1.7	0.0826	1.1	0.7726	12	0.0448	0.9	0.4916
lmo1664	similar to S-methionine adenosyltransferase	2	0.0103	1.6	0.0004	1.2	0.3430	2.7	0.0404
lmo1666	peptidoglycan linked protein (LPxTG)	1.4	0.7464	1.6	0.0002	1.1	0.9234	1.5	0.4404
lmo1675	2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylic acid synthase/2-oxoglutarate decarboxylase	2.1	0.0379	1	0.9228	0.6	0.0424	0.5	0.0721
lmo1681	similar to cobalamin-independent methionine synthase	1.2	0.7274	1.5	0.0061	0.7	0.0509	0.6	0.2051
lmo1694	similar to CDP-abequose synthase	8	0.0000	8	0.0000	26.2	0.0000	1.8	0.0162
lmo1696	putative membrane protein	1.2	0.4402	1.5	0.0073	1.2	0.7602	0.8	0.6933
lmo1698	similar to ribosomal-protein-alanine N-acetyltransferase	1.1	0.8156	3.5	0.0001	1.1	0.6149	1	0.9046
lmo1702	glyoxalase family protein	1	0.9402	1.6	0.0369	1.5	0.0258	1.3	0.2435
lmo1703	23S rRNA (uracil-5-)-methyltransferase RumA	0.9	0.7139	1.4	0.0149	1.5	0.0416	1.5	0.0372
lmo1704	similar to conserved hypothetical proteins	0.9	0.9078	1.4	0.0057	1.5	0.0471	1.8	0.0231

lmo1706	similar to transport proteins	1.1	0.8387	1.1	0.4251	1.4	0.0323	1.6	0.0162
lmo1708	aminoglycoside N3-acetyltransferase	1.5	0.0980	0.9	0.4128	1.5	0.1116	1.7	0.0362
lmo1713	cell shape-determining protein	1.5	0.4402	1.5	0.0244	0.9	0.7995	0.9	0.8729
lmo1724	similar to ABC transporter, ATP-binding protein	.	.	1.6	0.0003	1.2	0.4160	0.9	0.8629
lmo1725	similar to transcriptional regulator (GntR family)	1	0.9866	1.5	0.0035	1	0.9872	1.2	0.6786
lmo1730	similar to sugar ABC transporter binding protein	1.2	0.6616	0.6	0.0000	1.7	0.0060	2.2	0.0116
lmo1731	ABC transporter, permease protein	1	0.9873	0.6	0.0001	1.3	0.2257	2.2	0.0071
lmo1742	highly similar to adenine deaminases	1.1	0.8962	1	0.8892	1.4	0.0988	1.5	0.0484
lmo1749	similar to shikimate kinase	1.5	0.5157	1.9	0.0003	1	0.9227	1	0.9856
lmo1763	similar to unknown protein	1.2	0.9033	1	0.8571	1.6	0.0120	1.2	0.5434
lmo1766	phosphoribosylglycinamide formyltransferase	0.6	0.0220	1	0.9037	2.1	0.0148	1.8	0.0408
lmo1788	transcriptional regulator, MerR family	0.9	0.8766	1.7	0.0003	1.3	0.8600	1.6	0.0140
lmo1789	weakly similar to Nad(P)h Oxidoreductase chain B	1.4	0.2874	1.4	0.0111	1.6	0.0014	2.1	0.0014
lmo1790	similar to unknown proteins	1.5	0.2399	1.7	0.0001	1.7	0.0057	2.3	0.0024
lmo1791	unknown	0.9	0.7837	1	0.7805	1.2	0.0959	1.7	0.0021
lmo1798	similar to unknown protein	0.7	0.7304	1.1	0.6057	1.9	0.0208	2.2	0.0041
lmo1799	cell wall surface anchor family protein, authentic frameshift	2.8	0.0017	1.7	0.0002	4.8	0.0000	4.7	0.0000
lmo1806	highly similar to acyl carrier proteins	1.2	0.7644	1.9	0.0000	0.5	0.0116	0.6	0.0254
lmo1807	similar to 3-ketoacyl-acyl carrier	1.7	0.1908	1.5	0.0210	0.6	0.0378	0.6	0.0937

	protein reductase								
lmo1808	similar to malonyl CoA-acyl carrier protein transacylase	1.6	0.4251	1.6	0.0001	.	.	.	.
lmo1810	conserved hypothetical protein	2.1	0.0268	1.9	0.0002	0.7	0.0664	0.9	0.8465
lmo1830	similar to conserved hypotheticl proteins	3.3	0.0041	7.2	0.0000	2.4	0.1753	10.6	0.0039
lmo1849	similar to metal cations ABC transporter, ATP-binding proteins	2.3	0.0224	1	0.8362	1.3	0.0977	1.3	0.2640
lmo1868	similar to conserved hypothetical proteins	1.3	0.8325	1.1	0.8074	1.6	0.0340	1.4	0.4164
lmo1877	formate--tetrahydrofolate ligase	2.3	0.0060	1.6	0.0021	0.8	0.0975	0.9	0.6981
lmo1878	similar o transcriptional regulators	3	0.0066	1.6	0.0003	1	0.9451	1.3	0.4460
lmo1879	similar to cold shock protein	1.6	0.0989	0.8	0.1117	1.3	0.1468	1.7	0.0009
lmo1883	similar to chitinases	2.6	0.3466	3.2	0.0000	10	0.0000	16.8	0.0000
lmo1902	similar to ketopantoate hydroxymethyltransferases	1.5	0.0615	0.7	0.0005	1.4	0.0088	1.5	0.0053
lmo1908	similar to unknown proteins	1	0.9873	1.4	0.0037	0.9	0.6899	1.9	0.0126
lmo1929	similar to nucleoside diphosphate kinase	1.1	0.9873	1.8	0.0001	0.6	0.6767	1.2	0.6420
lmo1930	heptaprenyl diphosphate syntase component II [imported]	2.2	0.1056	1.6	0.0003	0.9	0.6611	1.2	0.5614
lmo1931	similar to 2-heptaprenyl-1,4-naphthoquinone methyltransferase	1.9	0.0763	1.8	0.0002	0.8	0.3166	0.9	0.8944
lmo1932	heptaprenyl diphosphate synthase component I, putative	1.8	0.1616	1.5	0.0272	1.9	0.0127	1.1	0.9297
lmo1933	similar to GTP cyclohydrolase I	2.2	0.5761	1.9	0.0002	2.1	0.0023	2.4	0.0004
lmo1935	similar to protein-tyrosine/serine phosphatase	1.8	0.0411	1.8	0.0014	0.9	0.4553	1	0.9378
lmo1956	similar to transcriptional regulator	2	0.0148	1	0.9788	0.9	0.8386	0.9	0.8544

	(Fur family)								
lmo1964	similar to ABC transporter, ATP-binding protein	0.9	0.5802	1	0.9795	1.2	0.3105	1.5	0.0490
lmo1993	similar to pyrimidine-nucleoside phosphorylase	1.7	0.0197	1.5	0.0051	1.1	0.6102	1.4	0.2611
lmo1997	similar to PTS mannose-specific enzyme IIA component	3.5	0.1153	0.8	0.0875	3.5	0.0270	3.3	0.0074
lmo1998	similar to opine catabolism protein	5.3	0.0467	0.7	0.0471	4.6	0.0095	5.2	0.0005
lmo1999	weakly similar to glucosamine-fructose-6-phosphate aminotransferase	4.3	0.1659	0.6	0.0128	3.9	0.1895	4.8	0.0098
lmo2000	similar to PTS mannose-specific enzyme IID component	13.5	0.0019	0.9	0.7744	5	0.0417	3.6	0.0129
lmo2001	similar to PTS mannose-specific enzyme IIC component	10.4	0.0089	0.7	0.0102	5	0.0198	13.1	0.0013
lmo2002	similar to PTS mannose-specific enzyme IIB component	2.4	0.1798	0.8	0.3997	7.4	0.0004	3.4	0.0192
lmo2003	transcriptional regulator, GntR family	2.3	0.3845	0.9	0.7192	5.1	0.0078	4.4	0.0215
lmo2004	transcriptional regulator, GntR family	3.8	0.0379	0.9	0.4411	4.3	0.0145	3.9	0.0329
lmo2020	similar to cell-division initiation protein (septum placement)	2.5	0.0096	1.4	0.0014	0.8	0.0761	0.9	0.5607
lmo2031	conserved hypothetical protein TIGR00044	1.2	0.4204	1.6	0.0111	0.9	0.4813	1	0.8617
lmo2033	highly similar to cell-division protein FtsA	1.4	0.2091	1.6	0.0015	0.8	0.0982	0.6	0.0404
lmo2038	similar to UDP-N-acetylmuramoylalanyl-D-	2.9	0.0124	1.2	0.0404	0.9	0.5445	0.7	0.2694

	glutamate-2,6-diaminopimelate ligase								
lmo2039	similar to penicillin-binding protein 2B	1.2	0.4477	1.1	0.6835	1.1	0.8391	1.6	0.0363
lmo2040	similar to cell-division protein FtsL	3.4	0.0066	1.5	0.0095	0.8	0.1536	0.8	0.2657
lmo2041	similar to unknown proteins	3.2	0.0187	1.8	0.0000	0.9	0.2491	0.8	0.4306
lmo2042	similar to unknown proteins	3	0.0092	1.8	0.0004	1	0.8560	1.1	0.7794
lmo2057	highly similar to heme A farnesyltransferase	0.5	0.0267	0.6	0.0014	1.5	0.0277	1.6	0.0056
lmo2058	similar to heme O oxygenase	2.1	0.0462	1.2	0.1411	.	.	.	.
lmo2067	similar to conjugated bile acid hydrolase	2.4	0.2565	3.5	0.0001	4.1	0.0007	6.4	0.0005
lmo2083	unknown	1.1	0.9430	1.1	0.3933	1.2	0.4428	1.6	0.0308
lmo2085	putative peptidoglycan bound protein (LPXTG motif)	11	0.0000	12.2	0.0000	14.1	0.0007	16.8	0.0067
lmo2092	glycine betaine transporter BetL	1.7	0.0494	1.5	0.0272	1.1	0.9583	1.7	0.0502
lmo2101	pyridoxine biosynthesis protein	2.4	0.0011	1.1	0.5643	1	0.9799	1	0.9837
lmo2102	glutamine amidotransferase, SNO family	1.9	0.0066	1.2	0.0845	0.9	0.4423	0.9	0.8221
lmo2118	similar to phosphoglucomutase	1.5	0.0367	0.9	0.6917	1	0.9037	0.9	0.6819
lmo2130	similar to unknown protein	2.1	0.0112	2.6	0.0000	1.9	0.0104	2.3	0.0006
lmo2132	unknown	2.3	0.0020	5.1	0.0000	2.8	0.2110	26.5	0.0012
lmo2157	sepA	7.8	0.0000	15	0.0000	12.1	0.0018	3	0.4231
lmo2158	similar to B. subtilis YwmG protein	1.9	0.0471	7.6	0.0000	1.4	0.1350	2.5	0.0071
lmo2159	similar to oxidoreductase	1.3	0.3295	3.4	0.0000	0.9	0.3815	1	0.9603
lmo2160	similar to unknown proteins	1.5	0.0988	2.8	0.0000	0.7	0.1229	1	0.9837
lmo2161	unknown	1.7	0.1642	2.2	0.0000	0.6	0.0024	0.9	0.6265
lmo2162	similar to unknown proteins	1.6	0.1513	2.1	0.0001	0.6	0.0011	1.2	0.6967

lmo2163	similar to oxidoreductase	1.9	0.0426	2.3	0.0000	0.7	0.0076	0.9	0.7272
lmo2167	similar to unknown proteins	2	0.0201	1.2	0.1227	1.6	0.2661	1.3	0.6501
lmo2168	similar to glyoxalase I	1.4	0.3124	1.7	0.0000	0.6	0.1038	0.8	0.3681
lmo2169	unknown	1.4	0.2512	1.7	0.0000	1	0.9847	1.1	0.9191
lmo2186	unknown	1.1	0.8937	0.7	0.0445	1	0.8930	1.8	0.0293
lmo2191	similar to unknown proteins	2.2	0.0066	3	0.0000	2.3	0.0000	2.3	0.0000
lmo2196	similar to pheromone ABC transporter (binding protein)	0.9	0.9436	1.6	0.0006	0.9	0.6604	0.7	0.0956
lmo2205	similar to phosphoglyceromutase 1	1.6	0.1304	2	0.0000	1.3	0.6091	2.6	0.0008
lmo2208	similar to unknown protein	2	0.0153	1	0.9761	0.9	0.8119	0.6	0.2350
lmo2213	similar to unknown protein	0.8	0.7963	2.3	0.0005	1.2	0.5468	1.1	0.9100
lmo2216	similar to histidine triad (HIT) protein	1.7	0.0444	1.3	0.0018	0.7	0.0386	0.9	0.9098
lmo2217	similar to unknown protein	1.5	0.0329	1.4	0.0010	0.6	0.0082	0.8	0.4273
lmo2223	similar to unknown proteins	2	0.0136	1.4	0.0056	2	0.0005	1.8	0.0090
lmo2230	similar to arsenate reductase	10.5	0.0947	18.7	0.0000	44.8	0.0000	61	0.0000
lmo2231	similar to unknown proteins	0.9	0.9720	3.1	0.0009	1.2	0.8828	1.3	0.7071
lmo2232	similar to unknown proteins	2.5	0.0123	1.2	0.1146	1.7	0.0019	1	0.9999
lmo2240	similar to ABC transporter (ATP-binding protein)	1.7	0.0137	1.1	0.4621	0.9	0.9298	1.4	0.4005
lmo2242	similar to O6-methylguanine-DNA methyltransferase	0.6	0.1730	0.8	0.1713	1.8	0.0408	1.2	0.6826
lmo2263	similar to unkown proteins	1.1	0.8678	0.9	0.5144	1.5	0.0195	1.2	0.6578
lmo2269	unknown	4.3	0.0096	5.7	0.0000	5.6	0.0000	6.2	0.0000
lmo2281	protein gp22 [Bacteriophage A118]	1.2	0.8360	1.2	0.5576	2	0.5209	2.2	0.0447
lmo2311	unknown	1	0.9763	1.1	0.8102	1.2	0.5213	1.5	0.0381
lmo2312	conserved hypothetical protein	0.8	0.8440	3.2	0.0325	3.8	0.0371	4	0.0976
lmo2324	prophage LambdaSa1, antirepressor, putative	0.8	0.8354	1.1	0.7613	1.7	0.0374	1.6	0.3839

lmo2331	weakly similar to gp32_Bacteriophage A118 protein	0.8	0.8746	1.2	0.5865	2	0.0384	1.4	0.2881
lmo2358	similar to N-acetylglucosamine-6-phosphate isomerase	1.1	0.9644	1.1	0.3686	1.5	0.7299	1.8	0.0388
lmo2367	glucose-6-phosphate isomerase	1.7	0.0424	1.5	0.0053	0.8	0.2180	0.6	0.0138
lmo2368	unknown	1.6	0.0577	1.5	0.0054	1.3	0.5531	1.4	0.6319
lmo2373	PTS system, cellobiose-specific, IIB component	0.7	0.1295	1.2	0.2524	1.4	0.0176	1.6	0.0063
lmo2386	similar to B. subtilis YuiD protein	1.1	0.9436	1.9	0.0000	2.3	0.0105	2	0.0395
lmo2387	conserved hypothetical protein	1.4	0.3857	4	0.0000	1	0.9804	1.5	0.1996
lmo2389	similar to NADH dehydrogenase	1.9	0.1056	1.8	0.0000	1.2	0.5807	1.2	0.1882
lmo2391	conserved hypothetical protein similar to B. subtilis YhfK protein	6.2	0.0000	9.1	0.0000	21.6	0.0000	28.5	0.0000
lmo2397	similar to NifU protein	1.8	0.0244	1.2	0.3112	1.5	0.0396	1.3	0.4215
lmo2398	low temperature requirement C protein, also similar to B. subtilis YutG protein	1.4	0.4240	2.4	0.0000	2.2	0.0280	1.2	0.9064
lmo2399	CBS domain protein	1.5	0.2577	1.8	0.0004	1	0.9331	1.2	0.6527
lmo2415	similar to ABC transporter, ATP-binding protein	2	0.0254	1	0.8350	1	0.8238	0.7	0.1064
lmo2425	similar to glycine cleavage system protein H	1.2	0.8038	1.1	0.7806	1.3	0.0701	1.5	0.0176
lmo2428	similar to cell division proteins RodA, FtsW	1.2	0.8989	1.1	0.7550	12.5	0.0435	.	.
lmo2434	highly similar to glutamate decarboxylases	3.2	0.0193	2.7	0.0000	4	0.0097	3.8	0.0000
lmo2436	similar to transcription antiterminator	2.1	0.0035	1.5	0.0063	1.3	0.5531	1.3	0.6952
lmo2437	unknown	1	0.9866	2	0.0000	1.7	0.0014	1.5	0.0413

lmo2454	unknown	3.6	0.0003	4.6	0.0000	6.6	0.0000	8.2	0.0000
lmo2455	highly similar to enolase	1.2	0.3981	1.2	0.0868	1.6	0.0074	1.6	0.0175
lmo2456	highly similar to phosphoglycerate mutase	1.1	0.9417	0.9	0.6339	2.9	0.0002	2.9	0.0010
lmo2457	highly similar to triose phosphate isomerase	1	0.9842	1	0.9538	2.6	0.0000	3.5	0.0000
lmo2458	highly similar to phosphoglycerate kinase	1	1.0000	0.9	0.7550	2.6	0.0000	3.3	0.0002
lmo2459	highly similar to glyceraldehyde 3-phosphate dehydrogenase	0.9	0.9512	0.9	0.6242	2.8	0.0000	4	0.0000
lmo2460	transcriptional regulator, putative	1	0.9873	1.2	0.0867	3.5	0.0000	4.4	0.0000
lmo2462	similar to dipeptidases	0.9	0.6680	1.3	0.0094	2.1	0.0001	2.2	0.0037
lmo2463	similar to transport protein	3	0.0426	3.9	0.0000	6.2	0.0000	4.3	0.0266
lmo2465	unknown	1.5	0.3884	1.8	0.0000	1.7	0.6178	2	0.0020
lmo2484	similar to B. subtilis YvlD protein	2	0.0728	5.1	0.0000	1.8	0.0055	1.9	0.0485
lmo2485	similar to B. subtilis yvlC protein	1.9	0.0113	4.4	0.0000	1.8	0.0006	2	0.0021
lmo2494	similar to negative regulator of phosphate regulon	2.3	0.0179	2.3	0.0027	2.9	0.1623	2.1	0.2611
lmo2507	highly similar to the cell-division ATP-binding protein FtsE	2.7	0.0290	1.3	0.0181	0.6	0.0997	0.5	0.0071
lmo2511	similar to conserved hypothetical proteins like to B. subtilis YvyD protein	1.2	0.5123	1.9	0.0001	1.3	0.5054	1.7	0.0231
lmo2520	N-acylamino acid racemase	0.8	0.6061	1.8	0.0000	0.8	0.7981	1.3	0.5028
lmo2522	LysM domain protein	1	0.9825	1.6	0.0028	1.1	0.8212	0.7	0.1624
lmo2533	highly similar to H -transporting ATP synthase chain b	1.3	0.3745	1.5	0.0025	0.9	0.4413	1	0.9378
lmo2534	highly similar to H -transporting ATP synthase chain c	1.4	0.1426	1.5	0.0006	1	0.8217	1.1	0.7593

lmo2536	ATP synthase protein I	1.4	0.2900	1.6	0.0009	0.8	0.4695	1.1	0.8465
lmo2539	highly similar to glycine hydroxymethyltransferase	2.2	0.0141	1.5	0.0040	0.8	0.0624	1	0.9043
lmo2547	homoserine dehydrogenase	1.5	0.0341	0.9	0.6523	0.8	0.3248	0.7	0.1535
lmo2568	unknown	1.4	0.6145	1.6	0.0236	1.3	0.6206	1	0.9600
lmo2570	unknown	5.5	0.0000	4.5	0.0000	7.3	0.0000	8.9	0.0000
lmo2571	similar to nicotinamidase	4.4	0.0000	5.8	0.0000	7.3	0.0000	8.3	0.0000
lmo2572	conserved hypothetical protein	4.4	0.0002	1.6	0.0018	7.5	0.0000	3.5	0.0000
lmo2573	similar to zinc-binding dehydrogenase	3.4	0.0002	4.6	0.0000	5.6	0.0000	7.5	0.0000
lmo2586	similar to formate dehydrogenase alpha chain	1.8	0.0267	1.3	0.0052	1.1	0.7557	1.6	0.2447
lmo2593	transcriptional regulator, MerR family	0.6	0.0268	1	0.8902	1.8	0.0302	1.5	0.2955
lmo2599	highly similar to B. subtilis YbaF protein	0.8	0.6145	1	0.9685	2.9	0.0219	0.8	0.4330
lmo2602	conserved hypothetical protein	9	0.0002	7	0.0000	6.7	0.0214	3.6	0.1571
lmo2603	unknown	2.3	0.0035	6.8	0.0000	1.2	0.8573	1.2	0.6160
lmo2605	ribosomal protein L17	1.3	0.3448	1	0.8362	1.5	0.0061	1.5	0.0309
lmo2611	adenylate kinase	1	0.9866	1.5	0.0143	1.9	0.0147	1.9	0.0054
lmo2633	ribosomal protein S10	1.6	0.0187	1	0.8983	1.1	0.8689	0.9	0.8244
lmo2638	similar to NADH dehydrogenase	1	0.9914	1.5	0.0025	1.5	0.0036	1.9	0.0004
lmo2660	similar to transketolase	1.7	0.0305	0.9	0.1802	0.9	0.8042	1.4	0.2805
lmo2662	ribose 5-phosphate isomerase B	2	0.0398	0.8	0.0416	1.1	0.6326	1.6	0.0494
lmo2663	similar to polyol dehydrogenase	2.1	0.0011	1.7	0.0001	1	0.8951	1.7	0.1618
lmo2664	similar to sorbitol dehydrogenase	2.2	0.0027	1.1	0.7538	1.1	0.8475	1.8	0.1064
lmo2665	similar to PTS system galactitol-specific enzyme IIC component	2.9	0.0064	1	0.9200	1.1	0.5942	1.8	0.0108
lmo2666	similar to PTS system galactitol-	2.9	0.0004	0.7	0.0005	1.2	0.2146	1.9	0.0231

	specific enzyme IIB component								
lmo2667	similar to PTS system galactitol-specific enzyme IIA component	2.3	0.0015	0.8	0.0316	1.1	0.7243	1.7	0.0817
lmo2668	similar to transcriptional antiterminator (BglG family)	2.2	0.0014	0.7	0.0058	1.2	0.2125	2	0.0071
lmo2670	conserved hypothetical protein	.	.	2.3	0.0000	2.5	0.0005	2.9	0.0001
lmo2671	unknown	1.9	0.0244	2.2	0.0000	1.2	0.7651	1.4	0.5402
lmo2672	weakly similar to transcription regulator	2	0.0027	2.7	0.0000	1.5	0.4151	1.8	0.2444
lmo2673	conserved hypothetical protein	6.9	0.0009	9.7	0.0000	26	0.0000	37.2	0.0000
lmo2674	ribose 5-phosphate isomerase B	9.8	0.0000	3.8	0.0000	10.1	0.0000	13.7	0.0000
lmo2695	similar to dihydroxyacetone kinase	1.6	0.1042	1.6	0.0003	1.8	0.0011	2.3	0.0019
lmo2696	similar to hypothetical dihydroxyacetone kinase	1.5	0.1100	1.8	0.0001	1.8	0.0010	3.6	0.0000
lmo2697	dihydroxyacetone kinase, phosphotransfer subunit	1.9	0.0947	1.5	0.0012	2	0.0046	3	0.0031
lmo2705	unknown	2.2	0.0424	2.2	0.0000	0.8	0.6344	0.9	0.9105
lmo2713	secreted protein with 1 GW repeat	1.4	0.1721	1.1	0.3510	1.2	0.5172	1.6	0.0337
lmo2720	acetyl-coenzyme A synthetase	1	0.9951	1.5	0.0022	0.7	0.2728	0.7	0.1591
lmo2724	similar to unknown proteins	2.8	0.0030	2.9	0.0000	4.7	0.0000	4.9	0.0000
lmo2733	PTS system, fructose-specific, IIABC component	1	0.9974	1.6	0.0016	0.9	0.8926	1	0.9573
lmo2738	conserved hypothetical protein	1.5	0.2984	1.4	0.0184	1.7	0.0157	1.3	0.4141
lmo2739	similar to hypothetical hemolysin transcriptional regulator, Sir2 family	1.7	0.0538	1.5	0.0038	1.5	0.0212	1.6	0.0327
lmo2740	unknown	1.1	0.9014	1.4	0.0004	1.7	0.0155	1.7	0.0187
lmo2741	major facilitator family transporter	1.7	0.0698	1.5	0.0010	1.6	0.0378	1.7	0.0179
lmo2742	unknown	0.7	0.1447	1.5	0.0015	0.9	0.6798	1.1	0.5745

lmo2748	general stress protein 26	11.6	0.0000	10	0.0000	15.3	0.0000	5.1	0.0063
lmo2758	similar to inosine-monophosphate dehydrogenase	1.7	0.0239	1.1	0.5999	0.9	0.5046	0.6	0.0061
lmo2763	similar to PTS cellobiose-specific enzyme IIC	2.3	0.2757	1.1	0.6010	2.3	0.0118	1.7	0.1991
lmo2765	similar to PTS cellobiose-specific enzyme IIA	1.9	0.5039	1.2	0.2317	3	0.0004	3.7	0.0306
lmo2791	Partition protein, ParA homolog	1.9	0.0260	1	0.7694	1.1	0.5956	1	0.9480
lmo2803	unknown	1	0.9951	1.2	0.4203	1.9	0.0437	1.3	0.4849
lmo2806	hypothetical secreted protein	1.2	0.7978	3	0.0000	1.8	0.2093	1.7	0.4461
lmo2832	similar to unknown proteins	1.2	0.8437	1.5	0.0083	1.1	0.8386	1.1	0.7141
lmof2365_0282	internalin D	3	0.0299	1.6	0.0245	2.8	0.2001	0.6	0.5607
lmof2365_0374	internalin	5.5	0.0001	0.9	0.7050	5.4	0.0046	3.4	0.2836
lmof2365_0459	hypothetical protein	.	.	0.9	0.8598	1.9	0.0465	1.5	0.7469
lmof2365_0470	conserved hypothetical protein	1.3	0.8989	1.3	0.1002	2.3	0.0038	2.9	0.0001
lmof2365_0703	hypothetical protein	2.2	0.0277	3.3	0.0000	2.2	0.0038	2.4	0.0186
lmof2365_0885	conserved domain protein	0.9	0.9870	1	0.9571	1.9	0.0408	2.3	0.1655
lmof2365_1122	hypothetical protein	.	.	2.2	0.4929	1.7	0.0439	1.3	0.4445
lmof2365_1273	MutT/nudix family protein	0.9	0.9004	1.1	0.6851	1.2	0.4237	1.7	0.0091
lmof2365_1394	hypothetical protein	1.9	0.1498	1.5	0.0068	1	0.9178	0.9	0.4528

lmof2365_1681	PTS system, sucrose-specific, IIBC component	3.8	0.0928	0.8	0.7639	1.8	0.5273	3.3	0.0202
lmof2365_2638	cell wall surface anchor family protein	4.8	0.0001	1	0.9467	0.9	0.8638	1.3	0.6331
lmof6854_0073.1	conserved hypothetical protein	0.9	0.8926	1.1	0.5644	1.3	0.1991	1.6	0.0052
lmof6854_2470	hypothetical protein	.	.	.	.	1.6	0.3807	4.7	0.0019
lmof6854_2470.3	gp27	.	.	1.7	0.0000	.	.	10.1	0.0510
lmof6854_2659.10	conserved hypothetical protein	0.6	0.4706	1.1	0.7376	1.6	0.0408	2.1	0.4946
lmoh7858_0080.5	hypothetical protein	3.2	0.0002	1	0.9174	0.9	0.8874	1	0.9347
lmoh7858_0487.7	transcriptional regulator, putative	0.9	0.9563	1.2	0.6487	2	0.0439	2	0.0438
lmoh7858_1167	response regulator, putative	0.9	0.8971	1.1	0.8741	2	0.0312	2	0.0351

[a]Probe name based on *L. monocytogenes* EGDe gene

[b]Common name based on EGDe annotation

[c]Fold changes represent transcript levels in the parent strain compared to the ΔsigB strain; significant genes ≥1.5-fold and adjusted p-value <0.05

[d] Adjusted p-value <0.05 considered significant

[e]E-value calculated by HMMER; E-value <0.1 considered valid hit

Table A2 [S3.1] RNA-Seq average GEI and TaqMan qRT-PCR absolute copy number of select genes.

Gene	Strain	RNA-Seq Average GEI[a]	TaqMan[b]	Study
ctc	10403S	2.70	6.99	Raengpradub et al. 2008, unpublished data[c]
ctc	Δ*sigB*	2.61	6.92	Raengpradub et al. 2008, unpublished data[c]
flaA	10403S	3.14	7.93	Unpublished data[d]
gadA	10403S	1.08	5.97	Unpublished data[cd]
gadA	Δ*sigB*	-0.94	2.33	Unpublished data[c]
gap	10403S	1.96	6.54	Raengpradub et al. 2008, unpublished data[c]
gap	Δ*sigB*	1.64	6.84	Raengpradub et al. 2008, unpublished data[c]
inlA	10403S	0.79	6.39	Unpublished data[d]
opucA	10403S	1.28	5.59	Raengpradub et al. 2008, unpublished data[c]
opucA	Δ*sigB*	0.73	4.80	Raengpradub et al. 2008, unpublished data[c]
plcA	10403S	0.65	5.49	Unpublished data[d]
rpoB	10403S	2.11	6.70	Raengpradub et al. 2008, unpublished data[cd]
rpoB	Δ*sigB*	2.13	6.82	Raengpradub et al. 2008, unpublished data[c]
sigB	10403S	1.70	6.94	Unpublished data[d]

[a]Log-transformed average GEI
[b]Log-transformed average absolute copy number calculated by TaqMan qRT-PCR
[c]Unpublished data from M. E. Palmer, Cornell University
[d]Unpublished data from R. A. Ivy, Cornell University

Table A2 [S3.2] ncRNAs identified in this study or previously identified

Description	Coordinates in 10403S	Length	10403S Average GEI[a]	Δ*sigB* Average GEI[b]	Fold change range (10403S/Δ*sigB*)[c]
Putative ncRNAs newly identified in this study					
putative ncRNA	161945..162111	167	32.70	34.10	0.82-1.11
putative ncRNA	222952..223741	790	1.99	2.17	0.67-1.38
putative ncRNA	409956..410100	145	43.80	82.82	0.4-0.7
putative ncRNA	419482..419602	121	269.20	306.99	0.3-1.46
putative ncRNA	477023..477185	163	7.46	5.46	0.36-3
putative ncRNA	479838..479991	154	56.02	67.52	0.19-1.6
putative ncRNA	836741..836942	202	15.52	11.34	0.78-1.96
putative ncRNA	938236..938563	328	14.47	29.94	0.44-0.53
putative ncRNA	1257547..1257724	178	20.56	23.43	0.24-2.07
putative ncRNA	1393256..1393496	241	52.11	65.68	0.51-1.49
putative ncRNA	1884385..1884664	280	25.20	45.98	0.34-0.89
putative ncRNA	2020305..2020575	271	189.49	224.23	0.52-1.18
putative ncRNA, *sbrE*	2072821..2073334	514	2359.89	20.95	30.82-470.9
putative ncRNA	2305436..2305610	175	20.62	49.18	0.33-0.53
putative ncRNA	2370319..2370547	229	45.73	17.84	0.8-6.31
ncRNAs in the Rfam database					
putative L10 leader	159701..159845	145	81.31	99.33	0.58-1.22
putative SAM riboswitch (S-box leader)	204783..204972	190	18.14	61.65	0.14-1.17
putative TPP riboswitch (THI element)	240868..241057	190	14.06	10.34	0.23-3.43
putative purine riboswitch	490215..490347	133	650.65	629.28	0.34-2.85
putative SAM riboswitch (S-box leader)	516988..517156	169	1.89	3.79	0.25-1
putative glucosamine-6-phosphate activated ribozyme	637782..638097	316	51.78	77.18	0.18-3.48
putative lysine riboswitch	707866..708136	271	57.98	59.25	0.56-2.13
putative SAM riboswitch (S-box leader)	762904..763066	163	103.57	189.42	0.33-1.12
putative PreQ1-I riboswitch	788075..788122	48	5.44	7.35	0.2-2.5
putative yybP-ykoY leader	902340..902520	181	11.41	16.88	0.65-0.7

putative cobalamin riboswitch	1037938..1038128	191	3.94	1.99	1.75-2
putative cobalamin riboswitch	1074606..1074806	201	2.96	3.97	0.67-0.86
putative glycine riboswitch	1230912..1231051	140	49.98	29.38	0.61-15.67
putative TPP riboswitch (THI element)	1319193..1319376	184	20.31	58.86	0.26-0.44
putative T-box leader	1352163..1352373	211	304.66	296.84	0.95-1.1
putative T-box leader	1412037..1412289	253	96.34	162.32	0.28-2.59
putative L21 leader	1435176..1435232	57	252.99	217.91	0.63-2.04
putative T-box leader	1447016..1447277	262	63.12	69.73	0.27-3.26
putative T-box leader	1455337..1455592	256	153.60	175.46	0.32-3.57
putative T-box leader	1500724..1500985	262	118.16	150.49	0.53-1.24
putative T-box leader	1534507..1534760	254	14.44	35.41	0.23-1.38
putative T-box leader	1534782..1535053	272	1086.72	1128.90	0.21-4.09
putative T-box leader	1569120..1569357	238	37.22	37.52	0.53-1.47
putative SAM riboswitch (S-box leader)	1574285..1574471	187	41.49	86.25	0.3-0.72
putative SAM riboswitch (S-box leader)	1597226..1597439	214	183.17	264.67	0.26-3.06
putative T-box leader	1660200..1660479	280	223.69	318.23	0.23-2.69
putative L19 leader	1707678..1707737	60	87.43	70.50	0.62-2.11
putative PyrR binding site element	1762762..1762865	104	1.67	1.54	0.5-3
putative PyrR binding site	1763445..1763553	109	5.12	3.11	0.8-3.5
putative purine riboswitch	1804062..1804230	169	18.56	33.91	0.33-0.93
putative FMN riboswitch (RFN element)	1865665..1865923	259	72.68	204.4	0.13-1.06
putative T-box leader	2134253..2134523	271	77.11	84.01	0.8-1.05
putative SAM riboswitch (S-box leader)	2327827..2328052	226	37.34	34.81	0.61-2.08
putative T-box leader	2505646..2505928	283	7.38	5.73	0.61-2.21
putative L13 leader	2524976..2525028	53	26.09	44.14	0.38-0.89
putative ykoK leader (M-box)	2605679..2605991	313	59.33	70.78	0.65-1.13
putative bacterial signal recognition particle (SRP)	2623466..2623799	334	1462.20	1295.75	0.62-2.71
putative T-box leader	2662484..2662739	256	93.57	70.33	1.24-1.43
ncRNAs previously described in *L. monocytogenes*[a]					
rliA (Mandin *et al.*, 2007)	392435..392658	224	0.00	0.00	1-1
rliB (Mandin *et al.*, 2007)	423176..423572	397	3.06	2.24	0.36-3
rliH (Mandin *et al.*, 2007)	1038953..1039381	429	8.79	11.08	0.48-1.16

sbrB (Nielsen *et al*., 2008) [e]	1133688..1133853	166	6.26	14.33	0.29-0.71
rliD (Mandin *et al*., 2007)	1217293..1217620	328	10.40	6.85	0.81-2.47
sbrA (Nielsen *et al*., 2008)	1257454..1257523	70	2.09	0.51	1-2
6S / *ssrS* (Mandin *et al*., 2007)	1404418..1404628	211	7921.44	6428.62	0.41-2.12
sbrC (Nielsen *et al*., 2008) [e]	1406547..1406619	73	0.00	0.00	1-1
rilE (Mandin *et al*., 2007)	1442719..1442941	223	0.00	0.00	1-1
putative L20 leader/*lhrB* (Christiansen *et al*., 2006) [f]	1705228..1705366	139	59.59	70.55	0.37-1.92
putative bacterial RNAse P class B; *rnpB* (Mandin *et al*., 2007)	1806961..1807510	550	1243.83	1422.30	0.65-1.12
rliF (Mandin *et al*., 2007)	1951412..1951631	220	0.00	0.00	1-1
Listeria Hfq-binding RNA, *lhrA* (Christiansen *et al*., 2006)	2193210..2193426	217	4532.33	4914.67	0.58-1.4
rliG (Mandin *et al*., 2007)	2232543..2232820	278	0.39	0.42	1-1
tmRNA (transfer messenger RNA / 10Sa RNA); *ssrA* (Mandin *et al*., 2007)	2346617..2347046	430	8566.23	7110.59	0.65-1.87
sbrD (Nielsen *et al*., 2008) [e]	2449849..2449916	68	2.28	2.04	0.33-2
rliI (Mandin *et al*., 2007)	2681235..2681460	226	29.23	38.99	0.53-1.18

[a] Average normalized number of reads matching each of the σ^B-dependent genes in the two 10403S datasets divided by the length of the genes times 100 bp;

[b] Average normalized number of reads matching each of the σ^B-dependent genes in the two Δ*sigB* datasets divided by the length of the genes times 100 bp;

[c] Average fold changes from the 10403S (two datasets) and Δ*sigB* (two datasets). ncRNAs with no matching reads in Δ*sigB* had their coverage manually set to 1 to allow for calculation of the fold change;

[d] ncRNAs previously described in *L. monocytogenes* have the article reference;

[e] ncRNA not experimentally validated;

[f] *lhrB* and the putative L20 leader ncRNA overlap and seem to be same ncRNA.

Table A2 [S3.3] Genes up-regulated by σ^B

Locus	EGD-e locus	Description	Avg. fold change (WT/Δ*sigB*)[a]	10403S Average GEI[b]	Δ*sigB* Average GEI[c]
LMRG_02371	lmo0122	similar to phage proteins	3.90	2.37	0.60
LMRG_02382	lmo0133	similar to *E. coli* YjdI protein	52.50	22.32	0.26
LMRG_02383	lmo0134	similar to *E. coli* YjdJ protein	62.00	22.12	0.00
LMRG_02414	lmo0169	similar to a glucose uptake protein	17.81	9.96	0.58
LMRG_02415	lmo0170	unknown	3.89	20.38	5.68
LMRG_02646	lmo0263 (*inlC2*)	internalin C2	140.67	12.80	0.13
LMRG_02851	lmo0263 (*inlD*)	internalin D	15.38	1.21	0.06
LMRG_02611	lmo0265	similar to succinyldiaminopimelate desuccinylase	204.50	17.95	0.00
LMRG_02602	lmo0274	unknown	3.17	2.89	0.91
LMRG_00013	lmo0321	similar to unknown proteins	11.40	11.50	1.85
LMRG_00064	lmo0372	similar to beta-glucosidase	4.26	2.40	0.66
LMRG_00098	lmo0405	similar to phosphate transport protein	6.52	4.54	0.84
LMRG_00126	lmo0433 (*inlA*)	Internalin A	5.86	6.19	1.06
LMRG_00127	lmo0434 (*inlB*)	Internalin B	6.00	2.71	0.47
LMRG_00131	lmo0439	weakly similar to a module of peptide synthetase	45.00	3.57	0.09
LMRG_00137	lmo0445	similar to transcription regulator	51.56	5.44	0.15
LMRG_00196	lmo0515	conserved hypothetical protein	17.73	15.42	0.90
LMRG_00221	lmo0539	similar to tagatose-1,6-diphosphate	14.54	132.74	9.30

		aldolase			
LMRG_00236	lmo0554	similar to NADH-dependent butanol dehydrogenase	120.00	13.54	0.08
LMRG_00237	lmo0555	similar to di-tripeptide transporter	8.38	18.76	2.27
LMRG_00275	lmo0593	similar to transport proteins (formate?)	5.75	9.01	1.88
LMRG_00278	lmo0596	similar to unknown proteins	170.50	32.33	0.09
LMRG_00285	lmo0602	weakly similar to transcription regulator	115.50	23.08	0.00
LMRG_00293	lmo0610	similar to internalin proteins, putative peptidoglycan bound protein (LPXTG motif)	7.87	3.33	0.57
LMRG_00311	lmo0628	unknown	24.00	8.64	0.00
LMRG_00312	lmo0629	unknown	3.21	7.60	2.39
LMRG_00341	lmo0654	unknown	7.10	56.61	7.94
LMRG_00342	lmo0655	similar to phosphoprotein phosphatases	4.38	7.41	1.69
LMRG_00357	lmo0669	similar to oxidoreductase	75.93	64.60	0.89
LMRG_00358	lmo0670	unknown	105.50	59.60	0.58
LMRG_00411	lmo0722	similar to pyruvate oxidase	19.82	12.69	0.76
LMRG_00469	lmo0781	similar to mannose-specific phosphotransferase system (PTS) component IID	59.58	29.59	0.65
LMRG_00470	lmo0782	similar to mannose-specific phosphotransferase system (PTS) component IIC	18.99	29.59	1.58
LMRG_02869	lmo0783	similar to mannose-specific phosphotransferase system (PTS) component IIB	90.75	24.90	0.35
LMRG_00472	lmo0784	similar to mannose-specific phosphotransferase system (PTS) component IIA	88.50	25.25	0.21

LMRG_00482	lmo0794	similar to B. subtilis YwnB protein	67.02	32.50	0.72
LMRG_00484	lmo0796	conserved hypothetical protein	4.21	43.88	10.61
LMRG_02244	lmo0819	unknown	3.01	18.35	6.09
LMRG_02304	lmo0880	similar to wall associated protein precursor (LPXTG motif)	269.00	19.38	0.00
LMRG_02011	lmo0911	unknown	4.04	33.90	8.58
LMRG_02013	lmo0913	similar to succinate semialdehyde dehydrogenase	330.38	30.05	0.11
LMRG_02036	lmo0937	unknown	54.38	44.68	0.91
LMRG_02052	lmo0953	unknown	167.00	73.18	0.48
LMRG_02094	lmo0994	unknown	81.50	22.06	0.16
LMRG_00583	lmo1140	unknown	11.93	47.84	4.28
LMRG_00687	lmo1241	unknown	21.43	7.42	0.45
LMRG_00745	lmo1295(*hfq*)	similar to host factor-1 protein	4.83	49.77	11.19
LMRG_00826	lmo1375	similar to aminotripeptidase	14.60	10.50	0.72
LMRG_00873	lmo1421	similar to glycine betaine/carnitine/choline ABC transporter (ATP-binding protein)	28.44	5.27	0.67
LMRG_00877	lmo1425 (*opuCD*)	similar to betaine/carnitine/choline ABC transporter (membrane p)	3.56	22.59	6.51
LMRG_00878	lmo1426 (*opuCC*)	similar to glycine betaine/carnitine/choline ABC transporter (osmoprotectant-binding protein)	3.77	19.78	5.41
LMRG_00885	lmo1433	similar to glutathione reductase	27.00	5.37	0.21
LMRG_01444	lmo1526	similar to unknown proteins	9.94	15.30	1.66
LMRG_01365	lmo1602	similar to unknown proteins	5.47	157.02	30.08
LMRG_01360	lmo1606	similar to DNA translocase	7.88	29.50	3.97
LMRG_02768	lmo1694	similar to CDP-abequose synthase	155.31	27.51	0.20
LMRG_02772	lmo1698	similar to ribosomal-protein-alanine N-	4.10	42.94	10.92

		acetyltransferase			
LMRG_00977	lmo1830	similar to conserved hypotheticl proteins	31.69	14.02	0.47
LMRG_01013	lmo1866	similar to conserved hypothetical proteins	2.63	4.87	1.79
LMRG_01030	lmo1883	similar to chitinases	155.50	14.66	0.05
LMRG_01151	lmo2003	similar to transcription regulator GntR family	14.67	3.15	0.32
LMRG_01217	lmo2067	similar to conjugated bile acid hydrolase	65.92	11.51	0.36
LMRG_01236	lmo2085	putative peptidoglycan bound protein (LPXTG motif)	104.38	24.72	0.23
LMRG_01284	lmo2130	similar to unknown protein	4.47	3.97	1.01
LMRG_02808	lmo2132	unknown	11.81	4.43	0.39
LMRG_01675	lmo2157 (*sepA*)	SepA	49.45	24.25	0.56
LMRG_01674	lmo2158	similar to *B. subtilis* YwmG protein	479.39	509.23	22.80
LMRG_01619	lmo2213	similar to unknown protein	94.50	18.81	0.00
LMRG_01602	lmo2230	similar to arsenate reductase	411.00	96.43	0.00
LMRG_01601	lmo2231	similar to unknown proteins	21.50	2.47	0.07
LMRG_01561	lmo2269	unknown	20.25	7.59	0.48
LMRG_02732	lmo2387	conserved hypothetical protein	29.33	3.49	0.12
LMRG_02736	lmo2391	conserved hypothetical protein similar to *B. subtilis* YhfK protein	11.76	39.48	4.54
LMRG_01850	lmo2398 (*ltrC*)	low temperature requirement C protein, also similar to *B. subtilis* YutG protein	2.80	50.03	18.94
LMRG_01814	lmo2434	highly similar to glutamate decarboxylases	126.00	11.95	0.12
LMRG_01794	lmo2454	unknown	84.50	50.24	0.76
LMRG_01785	lmo2463	similar to transport protein	9.61	4.07	0.42
LMRG_01764	lmo2484	similar to *B. subtilis* YvlD protein	4.31	21.34	5.05
LMRG_01763	lmo2485	similar to *B. subtilis* yvlC protein	3.93	32.87	8.47

LMRG_01754	lmo2494	similar to negative regulator of phosphate regulon	8.46	4.39	0.50
LMRG_02698	lmo2570	unknown	11.42	12.75	1.28
LMRG_02697	lmo2571	similar to nicotinamidase	9.84	25.15	2.99
LMRG_02696	lmo2572	similar to Chain A, Dihydrofolate Reductase	8.05	29.05	3.59
LMRG_02695	lmo2573	similar to zinc-binding dehydrogenase	7.52	25.91	3.83
LMRG_02146	lmo2602	conserved hypothetical protein	32.63	6.54	0.18
LMRG_02147	lmo2603	unknown	44.25	9.92	0.25
LMRG_02215	lmo2670	conserved hypothetical protein	3.00	25.23	8.58
LMRG_02216	lmo2671	unknown	3.13	27.29	8.82
LMRG_02217	lmo2672	weakly similar to transcription regulator	3.52	14.44	4.15
LMRG_02218	lmo2673	conserved hypothetical protein	150.50	31.92	0.11
LMRG_02219	lmo2674	similar to ribose 5-phosphate epimerase	5.42	52.93	9.94
LMRG_01972	lmo2724	similar to unknown proteins	6.95	20.50	3.72
LMRG_01963	lmo2733	similar to PTS system, fructose-specific IIABC component	7.95	1.35	0.32
LMRG_01948	lmo2748	similar to *B. subtilis* stress protein YdaG	207.50	49.37	0.00
LMRG_02448	lmo0019	unknown	29.63	4.79	0.17
LMRG_02472	lmo0043	similar to arginine deiminase	75.00	8.14	0.13
Noncoding	ND	putative ncRNA, *sbrE*	186.09	2359.89	20.95

[a]Average fold changes from the 10403S (two runs) and Δ*sigB* (two runs). Genes with no matching reads in Δ*sigB* had their coverage manually set to 1 to allow for calculation of the fold change;

[b]Average normalized number of reads matching each of the σ^B-dependent genes in the two 10403S datasets relative to the length of the genes times 100 bp;

[c]Average normalized number of reads matching each of the σ^B-dependent genes in the two Δ*sigB* datasets relative to the length of the genes times 100 bp.

Table A2 [S3.4] Comparison of genes found to be σ^B-dependent by microarray analysis and not by RNA-Seq

Transcript[a]	EGD-e locus[b]	Description[c]	σ^{Bd}	Fold [e]	Adj. p-value[f]	Fold change[g]	Adj. p-value[h]	Fold change[i]	Fold change[j]	Fold change[k]	Fold change[l]	q-value[m]	q-value[n]	q-value[o]	q-value[p]
LMRG_02632	lmo0210	similar to L-lactate dehydrogenase	+	5.2	0.0002	3.1	<0.0001	0.6	4.6*	0.4	2.9*	1.0000	<0.0001	1.0000	<0.0001
LMRG_02633	lmo0211	similar to *B. subtilis* general stress protein	+	2.1	0.0045	2.1	0.0003	1.1	1.9	0.8	1.4	<0.0001	<0.0001	1.0000	<0.0001
LMRG_02580	lmo0291	conserved hypothetical protein similar to B. subtilis YycJ protein	-	4.9	0.0016	2.1	0.0006	1.4	2.0*	1.8	2.5*	0.0004	<0.0001	<0.0001	<0.0001
LMRG_02320	lmo0896	indirect negative regulation of sigma B dependant gene expression	+	2.3	<0.0001	2.2	0.0003	2.9*	1.8	5.3*	3.3*	<0.0001	<0.0001	<0.0001	<0.0001
LMRG_02055	lmo0956	similar to N-acetylglucosamine-6P-phosphate deacetylase	+	2.9	0.0062	2.3	<0.0001	1.2	1.9	1.4	2.3*	0.0271	<0.0001	<0.0001	<0.0001
LMRG_02095	lmo0995	similar to *B. subtilis* YkrP protein	+	3.9	0.0000	2.8	0.0005	2.2	0.9	15.2*	6.3*	0.4254	1.0000	<0.0001	<0.0001
LMRG_00880	lmo1428	similar to glycine betaine/carnitine/choline ABC transporter	+	2.3	0.0033	2.9	<0.0001	1.9	2.5*	4.3*	5.7*	0.0001	<0.0001	<0.0001	<0.0001
LMRG_00884	lmo1432	unknown	-	4.5	0.0011	2.1	0.0002	1.3	1.3	2.5*	2.4*	0.2125	0.2927	<0.0001	<0.0001
LMRG_01366	lmo1601	similar to general stress protein	-	12	0.0001	4.2	<0.0001	3.4*	1.9	6.7*	3.7*	<0.0001	<0.0001	<0.0001	<0.0001
LMRG_01080	lmo1933	similar to GTP cyclohydrolase I	+	2.1	0.0159	2.1	0.0003	1.5	3.4*	2.8*	6.4*	0.5102	0.0116	0.0004	<0.0001
LMRG_01641	lmo2191	similar to unknown proteins	+	3.2	0.0001	2.7	0.0001	1.8	3.2*	1.0	1.9	<0.0001	<0.0001	0.4855	<0.0001
LMRG_01627	lmo2205	similar to phosphoglyceromutase 1	+	2.7	0.0022	2.2	0.0020	2.4*	2.9*	1.9*	2.4*	<0.0001	<0.0001	<0.0001	<0.0001

*Indicates significant RNA-Seq binomial comparison (fold change >2.0; q-value <0.05)

[a]*L. monocytogenes* 10403S locus description
[b]*L. monocytogenes* EGD-e locus number, Glaser et al., 2001
[c]Based on *L. monocytogenes* EGD-e, Glaser et al., 2001
[d]σ^B promoter present based on HMM developed in this study
[e]Fold change *L. monocytogenes* 10403S PrfA*/10403S PrfA*Δ*sigB,* Ollinger et al., 2009
[f]Adjusted p-value based on *L. monocytogenes* 10403S PrfA*/10403S PrfA*ΔsigB, Ollinger et al., 2009
[g]Fold change *L. monocytogenes* 10403S/Δ*sigB*, Raengpradub et al., 2008
[h]Adjusted p-value based on *L. monocytogenes* 10403S/Δ*sigB*, Raengpradub et al., 2008

[i]10403S replicate 1 / Δ*sigB* replicate 1

[j]10403S replicate 2 / Δ*sigB* replicate 1

[k]10403S replicate 2 / Δ*sigB* replicate 1

[l]10403S replicate 2 / Δ*sigB* replicate 2

[m]10403S replicate 1 > Δ*sigB* replicate 1

[n]10403S replicate 2 > Δ*sigB* replicate 1

[o]10403S replicate 2 > Δ*sigB* replicate 1

[p]10403S replicate 2 > Δ*sigB* replicate 2

Table A3 [S3.5] σ^B-dependent promoters used for HMM search

Gene	Motif
Initial set of experimentally validated motifs[a]	
prfA	TTGTTACTGCCTAATGTTTTTAGGGTATTTTAAA
lmo2230	ATGTTTCTAGTAATTTAAAAAGGGTAGATATTA
gadA	CGGTTTGTCTCTGTGGTTTAATGGGTATTGGTGA
lmo1433	TCGTTTGAAAGTGAAATCAGACGGGAAAACAAGC
bsh	ATGTTTTACTCCAAACTCCGAGGGTACTGGTAT
inlA	TAGTGTTATTTTGAACATAAAGGGTAGAGGATA
opuC	AAGTTTAAATCTATACTAGTTAGGGAAATTAGTT
lmo1421	AGGAATATTTAGGGATGATTTAGGGTAATTGGAT
lmo0669	ACGTTTTAGCGTAAAACAGGAGGGAAGACATAA
lmo2695	ACGTTTTGACTTTCTAGTAAAGGGAAATTGAGG
rsbV	ATGTTTTAATTTTATTTGTTAGGGTAAAATCGA
ltrC	ATGTTTAGAAATCCTGTAAACGTCTATCATACA
fri	ATGTTTAAGAAATTTTATCAGTGGTAAATACTTT
lmo0796	AGGTTTAATTTCTTAAGATTTAGGCTAGATTATA
lmo1830	CCGTTTTTTGTTTGTAATTTTAGGGTAGATGTGT
lmo2391	TGGTTTTATTTTTTACTCACCGGGAAAAGTTCT
Iteratively added motifs	
LMRG_00977T0	CCGTTTTTTCTTTCTAATTTTAGGGTAGATGTGT
LMRG_02146T0	TTGTTTTGGTTTAATGCCAAAGGGAATATATTA
LMRG_02011T0	TTGTTTTAACTTGCCCTCAGGCGGGTATTTATTA
LMRG_01674T0	ATGTTTTAGCTTTCTATATTGTGGAAAACACTA
LMRG_01794T0	CTGTTTTAAAAATAACGAGAGGGGTAATGATTT
LMRG_01236T0	CTGTTTTCTTTTGCTGTTTTATGGGTATTTAATG
LMRG_01365T0	AAGTTTTAGAGGGGAATACTCAGGGTATAGAAAA
LMRG_01972T0	TAGTTTAAGGTAAAACGAATTGGGTATTTTCTA
LMRG_02052T0	TTGTTTTACTTCTACTTTTTTTAGGGAATAAAATA
LMRG_00341T0	AGGATTACATTTCTATTTATTGGGGAAAAGTAGA
LMRG_00236T0	AGGTTTAAATTTTCTAAAAAAAGTGTATTATTAA
LMRG_02768T0	TGGTTTTAATACTACTAAAAAGGGAATAAACTA
LMRG_02448T0	TCTTTTTATTTTTCCAAAATAGGGTATACATAA
LMRG_01030T0	TAGTTTTATTTTCACTATGTTGGGTATTTTCTA
LMRG_00098T0	CTTTTTATATTTGTATAAAAGGGGTATAGACAA
LMRG_02646T0	TTGTTAATTTGGTCTAAAAAAGGGTATCTATTA
LMRG_01444T0	TCGTTTTTAATAGGACAGAAACGGGTACAGAATA
LMRG_02036T0	ATGTTTAAAGACTGATCTCACGGGAATATATAA
LMRG_00131T0	TTGTTTCACCGCACTGCTTTCAGGGAAACTATTA
LMRG_02013T0	CTGATTAAATTTTTCGATTTGTGGAAAACACTA
LMRG_02382T0	ACGTTTTCTTTTGGTTGATGAGTGGAATAGATGG
LMRG_02218T0	ATGCTTCTTTCTTTTATTTATGGGTATTAAGTA
LMRG_00098T0	TCTTTTTATATTTGTATAAAAGGGGTATAGACAA
LMRG_01619T0	CTGTTTCAATTATGAAAAACGTGGAAAATAAAG

[a]Taken from Raengpradub et al., 2008