LISTERIA MONOCYTOGENES PLACENTAL COLONZATION AND CONSEQUENCES FOR PREGNANCY OUTCOME

By

Kayla Nicole Conner

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Microbiology and Molecular Genetics – Doctor of Philosophy

2022

ABSTRACT

LISTERIA MONOCYTOGENES PLACENTAL COLONZATION AND CONSEQUENCES FOR PREGNANCY OUTCOME

By

Kayla Nicole Conner

Listeria monocytogenes (*Lm*) is a Gram-positive bacterium that causes the severe food-borne disease listeriosis. Listeriosis is particularly problematic in pregnant women as *Lm* colonizes the placenta, resulting in adverse pregnancy outcomes including stillbirth, miscarriage, and preterm labor. Despite numerous studies of placental listeriosis (PL) in various animal models, the mechanisms driving adverse outcomes following PL are largely uncharacterized. This dissertation addresses some of the field's knowledge gaps by analyzing the changes in placental gene expression and metabolism following infection with *Lm* and by characterizing a key *Lm* virulence factor, Internalin P (InlP), which plays a significant role in *Lm* placental colonization. Chapter 1 gives pertinent background information on the placenta, *Lm*, and PL and broadly addresses the knowledge gaps to be addressed by the rest of the dissertation.

Chapter 2 describes an *in vivo* study of PL in mice. Infected and control placentas were analyzed for differences in gene expression profiles between the two groups. We identified an enrichment of genes associated with eicosanoid biosynthesis, suggesting perturbations in eicosanoid metabolism in infected tissues. By quantifying placental eicosanoid concentrations through mass spectrometry, we found a significant increase in the concentrations of several eicosanoids with known roles in inflammation and/or labor. This study provides a likely explanation for temporal disruptions of labor following placental infection.

Chapters 3 and 4 discuss two studies of the *Lm* virulence factor InlP, a key player in placental colonization. InlP contributes to *Lm*'s placental pathogenesis likely by conferring the

ability of *Lm* to transcytose through placental layers. Prior studies reported that no homologs of InlP exist in *Listeria* species other than *Lm*. Chapter 3 describes our discovery that at least two other *Listeria* species, *L. ivanovii* and *L. seeligeri*, encode InlP homologs. We characterized the domain architectures and genomic neighborhoods of these homologs and speculated on their implications for *Listeria* evolution.

In chapter 4, I continue discussion of InlP and describe our identification and preliminary characterization of naturally occurring InlP variants. In this study, we used a bioinformatics approach to analyze *Lm* whole genome sequences (WGS) and identify InlP variants. We uncovered two InlP variants of interest in the *Lm* population. The first results from a start codon point mutation in the *inlP* gene, likely resulting in a truncated and potentially nonfunctional InlP protein product. The second is an InlP variant with a PRO to SER substitution in the InlP calcium binding loop, which is hypothesized to play a role in InlP activation or stabilization. These results provide two avenues for further investigation of InlP regulation and function and suggest the potential for InlP-dependent variation in placental colonization potential across *Lm* isolates.

In chapter 5, I summarize this dissertation. This chapter reflects on the results, implications, and challenges of each study outlined in the prior chapters. I discuss the unique challenges faced due to the ongoing COVID-19 pandemic and its effects on my graduate training. Finally, I share concluding remarks and propose future directions for this project and the field of PL. Together, the chapters of this dissertation describe novel findings that contribute to the field by assessing genetic and metabolic changes to the placenta due to listeriosis and further characterizing a known key placental virulence factor.

'You'll never do a whole lot unless you're brave enough to try.'
-Dolly Parton

This dissertation is dedicated to:
My parents, Jeremy and Michelle Conner. This victory is not just mine – it's ours.
My nana, Joan Byrge. I know you're smiling down on me. And I am certain you'd call me "Dr. Knee Hugger" if you were here.
My husband, Alan Halim. You are my rock, my joy, and my favorite person to troubleshoot experiments with.

ACKNOWLEDGEMENTS

In the dedication section of this dissertation, I quoted the great Tennessee queen, Dolly Parton – “You’ll never do a whole lot unless you’re brave enough to try.” It took a lot of bravery for a first-generation college student from a small east Tennessee town to pursue a Ph.D., but it should be known that this bravery was instilled in me by a long list of encouragers and mentors.

I want to start by thanking my graduate mentor, Dr. Jonathan Hardy. From day one of my lab rotation in the Hardy lab, he has been my biggest academic supporter. The early days of the Hardy lab were challenging – starting in a new lab comes with a long list of firsts and to-dos, and it can feel daunting. Jonathan quickly established an atmosphere of collaboration and comradery, often rolling up his sleeves and working alongside us at the bench to get the ball rolling. Hardy Lab group meetings ultimately became the place where I tackled my imposter syndrome; I learned to ask my questions and ask them loudly. I learned that my thoughts and ideas are valuable and that they should be spoken. Our exceptional lab environment is also heavily attributed to fellow Hardy Lab graduate students Jon Kaletka, Justin Lee, and Mike Witte. Thank you all for providing feedback, ideas, laughs, memes, and songs for the Party Hardy playlist. I consider myself incredibly lucky to have spent almost five years working alongside some of my favorite people.

This accomplishment would have been impossible without the added help of my graduate committee (Dr. Rob Abramovitch, Dr. Ripla Arora, Dr. Christopher Contag, and Dr. Andrew Olive) and our MMG graduate coordinator, Roseann Bills. I would like to thank my committee for offering thoughtful feedback on experiments, guidance for my career, and encouragement when things became frustrating. I would like to thank Roseann for answering all (approximately ten thousand) of my questions, putting out small fires, and generally working magic in many

situations. Keeping a whole department of Ph.D. students organized is far from easy (kind of like herding cats?), but Roseann somehow manages it beautifully and keeps us all on track.

Undoubtedly, I would not have pursued this path without the encouragement of past educators and mentors. I would like to acknowledge my high school chemistry teacher, Brenda Hardin, for being one of the first teachers who made me feel like becoming a scientist was attainable and for instilling a love of chemistry in me that ultimately helped me to choose my college major. I would like to also acknowledge my high school English teacher, Sherrie Collins, whose instruction ensured I have always been able to communicate my science effectively. Finally, I would like to acknowledge my undergraduate research mentor, Dr. Erin McClelland. As an undergraduate at Middle Tennessee State University, I studied novel small molecule inhibitors of the fungal pathogen *Cryptococcus neoformans* under her direction. It was in Dr. McClelland's lab that I realized my love for research and decided to pursue a career in microbiology. I am forever grateful to her for allowing me to join her lab and learn the ropes of basic scientific research.

Completing a Ph.D. is an extremely long and difficult process, but it becomes much more enjoyable when surrounded by loving, caring, and encouraging friends. To Emily, Lauren, Cody, Jon, Victoria, Justin, Kristin, Hunter, Zach, and Kate – thank you for pushing through this with me. Thank you for the times you've been willing to proofread things, listen to practice presentations, offer experimental and troubleshooting advice, and listen to complaints when they needed to come out. Thank you for the game nights and the belly laughs and the roasts. To Haley, Andrew, Lane, and Anneke – thank you for loving me so well, even if the Ph.D. journey didn't always make sense from the outside looking in. You have all laughed with me on my best days and cried with me on my worst; you've celebrated my victories as fiercely as you would your own, and my life is brighter with each of you in it.

As I've neared the end of my Ph.D. journey, I've learned that one of the best parts of this accomplishment is sharing it with the people closest to me, my family. Mom and Dad, you have emphasized to me from the beginning that I can achieve anything I want. I'm not the world's typical idea of a doctoral-level microbiologist – I was a first-generation college student, I'm the first in our family to hold an advanced degree, and I didn't have particularly outstanding GPAs in high school or college. There weren't a ton of resources available to me – no fancy science classrooms at school, no money for tutors or science camps – but there was a lot of love, and it was evident in all that you did for me, and it continues to be evident in all that you do for me now. Of all the science I'll ever do, none will ever top looking through the telescope with Dad or listening to Mom read about the Chernobyl disaster to me before helping me build a model of the reactor (I was a weird kid, I know… it's fine). To my brother, Noah, thank you for always keeping me laughing, for being someone I can tell my deepest secrets to, and for trusting me as your resource for all your evolution and vaccine questions. You've taught me that I should always channel the confidence of a teenage boy arguing about science with his older sister – a literal scientist.

The most difficult part of my Ph.D. was continuing after the loss of my grandmother, or as I called her, Nana. Nana was everything that I hope to be – fiery, outspoken, smart, sassy, funny, comforting, loving, a great cook (all hope has been lost for me there), and a master gardener. Nana brightened the lives of everyone around her. She was selfless and infinitely kind and thoughtful. She kept pictures of me in her purse and took every opportunity to tell anyone who would listen that her granddaughter was going to be a microbiologist. Undoubtedly, her pride would be insufferable if she was here for this.

Finally, I would like to acknowledge my husband, Alan Halim. You are the best thing that graduate school brought me, and all of this means so much more with you. You are understanding, patient, and kind beyond words. You've talked me through imposter syndrome, failed experiments, long days, and every "I want to quit" moment. You're the only person I ever want to quarantine with, even if it means I have to listen to you practice the ukulele. You're the best cat dad to Louis and Charlie, and your dad jokes make for the best belly laughs (also eye rolls). Thank you for keeping me grounded on the hard days and for celebrating with me on the good ones – there's a bottle of Lillet Blanc waiting for us at the end of this.

This accomplishment took a lot of bravery and willingness to try and fail. It also took a village of teachers, mentors, friends, and family to get here. To all who have shared this journey with me, thank you – your love, support, and presence in my life are valued more than I can possibly put into words.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

AA	Arachidonic Acid	**LRR**	Leucine Rich Repeat
ActA	Actin Assembly Inducing Protein	***Ls***	*Listeria seeligeri*
BHI	Brain Heart Infusion medium	**LTB_4**	Leukotriene B_4
BLAST	Basic Local Alignment Search Tool	**LXA_4**	Lipoxin A_4
CAP	Chorioallantoic Placenta	**OD_{580}**	Optical Density at 580nm
CFU	Colony Forming Units	**PBS**	Phosphate Buffered Saline
Cm	Chloramphenicol	**PGH_2**	Prostaglandin H_2
COX	Cyclooxygenase	**PCR**	Polymerase Chain Reaction
CVP	Choriovitelline Placenta	**PGA_2**	Prostaglandin A_2
Il-1	Interleukin 1	**PGD_2**	Prostaglandin D_2
Il-8	Interleukin 8	**PGE_2**	Prostaglandin E_2
Inl	Internalin	**PGF**	Prostaglandin F
IVIS	*In Vivo* Imaging System	**$PGF_{2\alpha}$**	Prostaglandin $F_{2\alpha}$
LB	Luria Burtani Medium	**PGH_2**	Prostaglandin H_2
Lc	*Listeria costaricensis*	**PLA_2**	Phospholipase A_2
Lin	*Listeria innocua*	**RNAseq**	RNA Sequencing
Liv	*Listeria ivanovii*	**Strep**	Streptomycin
LLO	Listeriolysin O	**TNF-α**	Tumor Necrosis Factor α
Lm	*Listeria monocytogenes*	**WGS**	Whole Genome Sequencing
LOX	Lipoxygenase	**WT**	Wild Type

CHAPTER 1

PLACENTAL LISTERIOSIS: HOW *LISTERIA MONOCYTOGENES* BREACHES THE PLACENTAL BARRIER, AND CONSEQUENCES FOR PREGNANCY OUTCOME

INTRODUCTION

Placental Structure and Cell Types

The placenta is a transient organ in mammals that serves as the interface between the mother and fetus and is expelled during delivery (**Fig. 1.1A**). This organ is the site of gas exchange, metabolic and endocrine functions, and vascular rearrangement; processes that are critical for the maintenance of the pregnancy and fetal development [1]. The major cell type of the placenta is the fetally-derived trophoblast [2]. Trophoblasts begin forming in the first stages of development, differentiating from the blastocyst [2]. This gives rise to cytotrophoblasts, which can develop further into the fused, multinucleated syncytiotrophoblast layer or invasive extravillous trophoblasts [2,3].

The placenta is anchored into the decidua by protruding villi which are composed of extravillous trophoblast cells (**Fig. 1.1B**). In humans, these cells invade the decidua during early pregnancy, forming villous trees that maintain the placenta's close contact with maternal circulation throughout pregnancy [2]. The syncytiotrophoblasts form the outermost cell layer of the placenta and play important endocrine functions throughout pregnancy [3]. This cell layer is fused, lacking cell-cell junctions, making it largely resistant to pathogens [4–7]. The syncytiotrophoblast is in direct contact with maternal blood and is a major site for physiological exchange between the mother and fetus, necessitating its resistance to outside invaders [3].

While the trophoblasts are inarguably critical for placental physiology, other placental cell types also contribute to its function. Mesenchymal-derived Hofbauer cells are placenta-resident macrophages and are the primary antigen-presenting cells of the placenta [3]. Hofbauer cells are also responsible for producing several growth factors and cytokines that ultimately promote

trophoblast differentiation [3]. The placenta also houses fetal vascular cells, such as pericytes and fetal endothelial cells, which are critical for regulating placental blood flow [3].

Placental Metabolism and Labor

The placenta is an active regulatory participant throughout the duration of pregnancy, including labor. Through its production of estrogen, progesterone, oxytocin, relaxin, and prostaglandins, the placenta helps to regulate temporal and physical aspects of parturition (**Fig. 1.2**) [8]. Progesterone is maintained at high concentrations throughout pregnancy, inhibiting contractility of the myometrium until the onset of labor [9,10]. Relaxin has various functions throughout pregnancy and induces the production of enzymes that contribute to cervical ripening in late pregnancy [11]. The production of estrogen at term promotes uterine contractility and increases oxytocin receptor expression, allowing oxytocin to induce myometrial contractions and prostaglandin production [12–14]. Prostaglandins have been studied comprehensively for their role in pregnancy and the regulation of labor by promoting inflammation [15,16]. The onset of labor requires a switch from an anti-inflammatory state to a pro-inflammatory state at the maternal-fetal interface, mimicking inflammation at other bodily sites caused by infection and other disease.

Prostaglandins

Prostaglandins are 20-carbon fatty acids derived from arachidonic acid (AA). Prostaglandin biosynthesis begins with the liberation of AA from phospholipids of the cell membrane (**Fig. 1.3**) [17]. This process is primarily mediated by Phospholipase A_2 (PLA_2), which hydrolyzes the phospholipid backbone at the sn-2 position to yield free AA [18]. Once released from the phospholipid bilayer, free AA can diffuse to other cells, become reincorporated into the

phospholipid bilayer, or undergo metabolization by cyclooxygenase (COX) or lipoxygenase (LOX) to yield prostanoids and lipoxins, respectively [17]. Two subclasses of prostanoids, thromboxanes and prostaglandins, are produced through further enzymatic and non-enzymatic processing by various isomerases and dehydration reactions [17]. Conversion of AA into prostaglandins depends heavily on COX [17]. Two structurally similar but functionally unique isoforms of COX, COX-1 and COX-2, exist throughout the body to carry out this function [19]. COX-1 is a constitutively active isoform of COX while COX-2 is inducible and typically responsible for inflammatory responses. Both form prostaglandin H_2 (PGH_2), which may then be metabolized further into other prostaglandins [19].

Late in pregnancy, Prostaglandin E_2 (PGE_2) and Prostaglandin $F_2\alpha$ ($PGF_2\alpha$) are prominent mediators of cervical ripening and promote the induction of labor (**Fig. 1.2**) [16,20,21]. Both have been shown to induce elastin, collagenase, and matrix metalloproteinases that break down fetal membranes and promote cervical and myometrial contractility. Synthetic versions of these prostaglandins, like Misoprostol and Dinoprostone, have been used in the clinic to artificially induce labor since the 1970s [16,20,22].

Placental Dysfunction and Preterm Labor

The timing of labor requires coordination of numerous mediators, including prostaglandins, originating from the placenta (**Fig. 1.2**) [15]. Placental perturbations can disrupt this process, sometimes resulting in preterm labor [21]. Preterm labor, defined as the onset of labor before 37 weeks of gestation in humans, occurs in an estimated 5-18% of all pregnancies and is the leading cause of neonatal death worldwide [23]. Preterm labor remains a problem throughout childhood as it can result in developmental abnormalities and neurodevelopmental disorders [24,25]. Preterm

labor is a syndrome with multiple etiologies including maternal stress, preeclampsia, vascular disorders, and infection [26,27].

Even though relatively very few pathogens can invade and colonize the placenta, placental infections are a significant public health concern. An estimated 30% of all preterm births can be attributed to underlying infection [27]. Pathogens in the reproductive tract and at the maternal-fetal interface can be detected by toll-like receptors which ultimately drive the production of pro-inflammatory chemokines (IL-1β), cytokines (IL-6, TNF-α and IL-8), and prostaglandins [28–30]. This can result in an overall pro-inflammatory state at the maternal-fetal interface and induction of preterm labor.

Models for Study of the Placenta

The placenta is incredibly diverse in structure, shape, and vasculature across placenta-harboring species. Thus, it is important to be mindful when choosing experimental models for placental studies. There are four primary placental classifications to consider in choosing a model: 1) placenta type, 2) chorioallantoic placenta shape, 3) histological structure, and 4) vasculature. Additionally, placenta researchers must consider the available genetic tools and cost associated with each possible model (**Table 1.1**).

The first classification differentiates between the two types of placentas - choriovitelline (CVP) and chorioallantoic (CAP) [31]. CVPs are considered a more primitive form of this organ and are typically found in marsupials. These primitive placentas are sometimes referred to as the "yolk sac placenta," may also serve as the primary placenta during early pregnancy before the CAP takes over as the pregnancy progresses in some species. The CAP is the larger, more complex

placenta that functions throughout the duration of pregnancy in most non-marsupial species [31].

Second, the placenta can be classified in terms of shape of the chorioallantoic placenta. Four possible shapes exist: diffuse (large and covers most of the fetus), multicotyledonary (small spots of tissue scattered across the fetus), zonary (forms a band that wraps around the center of the fetus), and discoid (small disc on one side of the fetus) [31].

The third level of classification focuses on the histological structure of the placenta and is primarily defined by the level of invasion of trophoblast cells [31]. Epitheliochorial placentas are considered the most superficial type with no destruction or invasion of maternal tissues and loose association between trophoblasts and endometrial epithelium. Endotheliochorial placentas exhibit degradation of the epithelium and connective tissue in the uterus, which leads to direct contact between the trophoblasts and endometrium. Finally, hemochorial placentas are considered the most invasive type, with all maternal tissue layers being degraded, leading to direct contact between the fetal chorion and maternal blood. The hemochorial placenta can be further characterized based on the number of trophoblast layers found in the placenta (referred to as 'hemomonochorial' for one layer, 'hemodichorial' for two layers, etc.) [31].

The fourth and final classification of placentas is based on the structure of the fetal-maternal interface in hemochorial placentas and consists of two types: villous and labyrinthine [31]. Villous placentas are found in primates (including humans) and exhibit branched villi that are bathed in maternal blood. These villi are branched by extravillous trophoblast cells that invade deeply into the decidua, the thick layer of the maternal uterus that lies closest to the fetus. Labyrinthine placentas, found in rodents, are characterized by anchoring trophoblast giant cells that associate more loosely with the decidua without invading [31].

Human placentas are categorized as chorioallantoic, discoid, hemomonochorial, and villous. Mouse placentas are characterized as chorioallantoic, discoid, hemotrichorial, and labyrinthine. Thus, despite being metabolically similar, murine and human placentas differ in histological structure and vasculature. The most common models used for the study of placental infection are choriocarcinoma cell lines, mice, gerbils, guinea pigs, nonhuman primates, and human placental explants [32]. Each model has its own benefits and drawbacks, considering factors such as cost, genetic malleability, and similarity to the human placenta (**Table 1.1**) [32]. Guinea pigs and nonhuman primates have historically been used as the most accurate models of human placental dysfunction, as their placenta structures most closely match those of humans. However, mouse models offer many benefits such as cost minimization, short gestation period, well-characterized genetics, and similar molecular/biochemical changes associated with labor to those that occur in humans [32,33]. Further, results obtained in *in vivo* mouse models and mouse-derived cell lines can be validated in human cell lines and explants [32,33].

Listeria and the Placenta – Background and Epidemiology

Listeria monocytogenes is a bacterial pathogen known to colonize the human placenta and drive a detrimental pro-inflammatory response [34]. Due to this ability, paired with its well-characterized physiology and genetic malleability, *L. monocytogenes* has been used in numerous studies of pregnancy-associated infection and its outcomes for decades.

The *Listeria* genus encompasses 21 identified species including the human pathogen *Listeria monocytogenes* (*Lm*), the causative agent of listeriosis [34]. *Lm* is a Gram-positive, facultative anaerobic bacterium with a saprophytic lifestyle [34]. Many of *Lm*'s characteristics which allow it to survive in the soil environment (including resistance to high salt concentrations

and ability to divide at low temperatures) also render it resistant to many common antimicrobial practices used in food production [34]. While the average adult will encounter multiple *Lm* exposures each year with no adverse outcomes, immunocompromised individuals remain vulnerable to listeriosis [35]. Alarmingly, listeriosis comes with a ~20% fatality rate among those infected [35].

Listeriosis ranks third in deaths due to foodborne illness worldwide, exceeding those caused by *Salmonella* and *Clostridium* [35]. In human hosts, *Lm* can cause gastrointestinal, central nervous system, and perinatal disease [34]. Pregnant individuals are approximately 10 times more likely to contract listeriosis compared to the non-pregnant population, and an estimated 17% of all annual listeriosis cases are pregnancy-associated [36]. The actual number of annual listeriosis cases is likely underestimated due to the subclinical nature of the disease, which is predicted to result in many pregnancy-associated listeriosis cases being unaccounted for. Pregnancy-associated listeriosis can result in multiple adverse pregnancy outcomes for both the mother and fetus, including spontaneous abortion, miscarriage, preterm labor, abnormal fetal development, and deadly neonatal disease [37].

To date, there have been 14 identified *Lm* serovars, with only three (1/2a, 1/2b, and 4b) responsible for 95% of all listeriosis cases [34,35,38–40]. While 1/2b isolates account for most food contaminants, serovar 4b is responsible for >50% of all human cases of listeriosis [34]. Additionally, serovar 4b strains are found more frequently in pregnancy-associated cases of listeriosis than in non-pregnancy-associated cases [34]. Together, these findings suggest that serovar-specific adaptations may dictate the fitness of *Lm* in specific environments or host niches.

Virulence Factors and Infectious Cycle

Lm encodes an arsenal of virulence factors that are critical for its intracellular lifecycle, many of which are regulated by the thermosensitive master virulence regulator PrfA (**Fig. 1.4**) [34]. Once ingested with contaminated food, *Lm* can invade host gut epithelial cells by binding host cell E-cadherin via the surface protein Internalin A (InlA) [41]. From there, Lm can disseminate to the liver where it binds C-Met on hepatocytes via InlB [42]. Once internalized by the host cell, *Lm* secretes the beta hemolysin Listeriolysin O (LLO) which releases it from the phagocytic lysosome and into the cytosol [43]. There, the actin polymerization protein ActA allows *Lm* to undergo cell-to-cell spread by polymerizing host actin, creating "actin rockets" that propel it into neighboring cells [44–46].

Internalins

While many *Lm* virulence factors have been characterized, the internalins are continually being studied for both their role in different *Listeria* species and for their basic biochemical properties. The internalins are a large family of proteins found across the *Listeria* genus [47]. In *Lm* alone, over 20 internalins have been identified to date [47]. The internalins can be divided into two main categories – secreted and anchored [47]. The secreted internalins share C-terminal signal peptides, while the anchored internalins share anchoring domains that lock them onto the *Lm* surface (**Fig. 1.5**) [47]. All internalins share characteristic leucine rich repeat (LRR) domains at their core, which can vary in length and number between these proteins [47]. LRRs are found in proteins across the domains of life, including other bacterial pathogens, and have been shown in many cases to be the sites of protein-protein interactions [48–50]. The continued study of

internalins will undoubtedly offer answers for questions in not only *Listeria* biology, but also basic biochemistry.

While InlA and InlB remain the most well-characterized internalins, others have been shown to contribute to virulence in the human host. InlC has been shown to enhance *Lm* cell-to-cell spread in the liver through its interactions with the host protein Tuba [51,52]. InlF interacts with host vimentin and enhances *Lm*'s ability to cross the blood-brain-barrier and colonize the brain [53]. Finally, InlP has been shown to interact with the host protein afadin and, through an unclear mechanism, enable *Lm* to transcytose through the layers of the placenta [54,55].

Breaching the Placental Barrier

The placenta is inherently resistant to most pathogens, and evidence suggests that the pathogens able to colonize this organ likely use multiple, diverse strategies to breach its protective barriers [7]. The syncytiotrophoblast has been experimentally shown to resist invasion by bacterial pathogens including *Lm*, but damage to this cell layer can allow invaders to cross [4,6]. The underlying cytotrophoblasts and anchoring extravillous trophoblasts are much more susceptible to infection [7]. Previous studies suggest that these cell types are the primary routes of entry for *Lm* into the placenta [56]. Further, there is evidence that *Lm* primarily remains intracellular in the placenta and may use a trojan horse strategy to spread from infected maternal cells to extravillous trophoblasts, allowing it to bypass the protective syncytiotrophoblast layer and traverse the placental cell layers [56].

Concluding Remarks and Dissertation Overview

The molecular mechanisms driving temporal and physical regulation of parturition are relatively well-understood, but many questions remain regarding preterm labor. Preterm labor has numerous causes, including placental infection by specialized pathogens that can breach this organ's protective barriers. The maternal-fetal interface reaches a pro-inflammatory state at the onset of labor, and this state can be achieved before term with the presence of placental pathogens driving the production of pro-inflammatory cytokines, chemokines, and prostaglandins.

Lm is one pathogen known to invade the placenta and has been used as a model in studies of placental infection. This bacterial pathogen is known to induce preterm labor, but the mechanisms underlying this remain unclear. *Lm*'s well-understood lifecycle and repertoire of virulence factors offer clues as to how it colonizes the placenta and interacts with the host to drive the pro-inflammatory, pro-parturition state. Future studies should address the genetic and metabolic perturbations of the placenta resulting from *Lm* invasion which could ultimately drive this state and its outcomes. Additionally, future studies of *Lm* virulence factors (like InlP) could inform on how *Lm* and other placental pathogens have evolved specialized mechanisms for placental invasion. Ultimately, the study of preterm labor and its causes remains important for public health worldwide. Further elucidation of this syndrome and the mechanisms driving it could inform on potential therapeutic strategies to reduce the occurrence of preterm labor and its devastating outcomes for both mother and fetus.

This dissertation serves to address some of the gaps in knowledge regarding 1) perturbations in placental gene expression and metabolism resulting from placental infection and 2) the role of InlP in placental invasion by *Lm*. In Chapter 2, we use a mouse model of placental

listeriosis and differential gene expression analysis to identify over- and underexpressed genes in the placenta following *Lm* colonization. We also identify disturbances to the eicosanoid pathway, resulting in altered placental prostaglandin concentrations which could contribute to infection-induced preterm labor. In Chapter 3, we use a computational comparative genomics approach to identify InlP homologs in non-*monocytogenes* species of *Listeria*. Finally, in Chapter 4, we identify naturally occurring InlP variants in the *Lm* population using computational analysis of publicly available whole genome sequences. We pair this InlP data with available metadata for each isolate to establish serovar-specificity of some variants of interest. Finally, in Chapter 5, we conclude this dissertation with an overview of our major findings and discussion of future directions.

APPENDIX

FIGURES AND TABLES

Figure 1.1 The human placenta. The human placenta is attached to the uterine wall via extravillous trophoblasts in the placenta (A). The outermost layer of the placenta is a multinucleated layer of syncytiotrophoblasts. Beneath this layer are the cytotrophoblasts which give rise to the syncytiotrophoblasts and extravillous trophoblasts (B). Figure created with biorender.com.

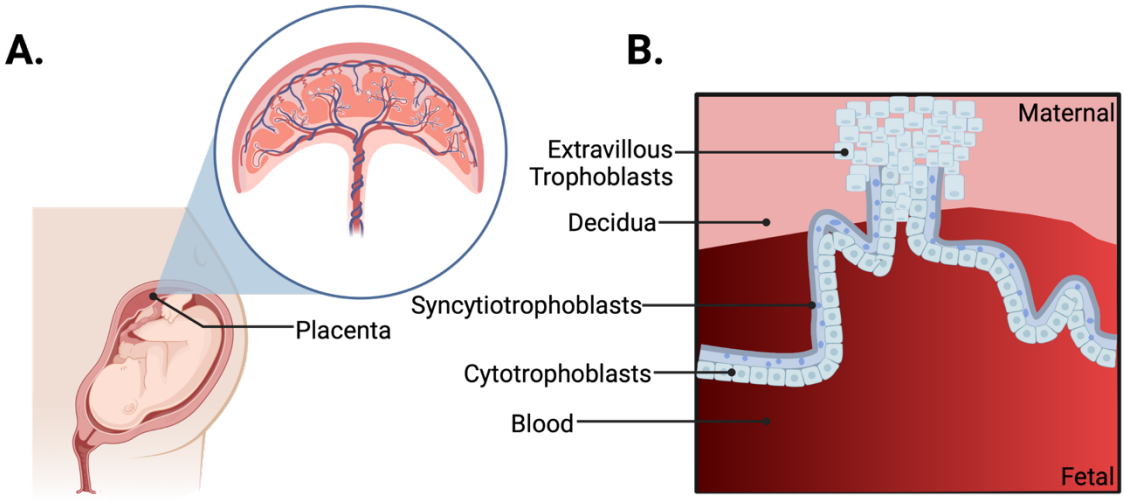


Figure 1.2 Parturition in humans. At term, fetal stress drives the release of cortisol from the fetal adrenal gland. This cortisol acts on the placenta to decrease progesterone production which, in turn, increases the expression of oxytocin receptors and production of prostaglandins. Prostaglandins stimulate cervical softening and uterine contractions. Continued oxytocin production drives this process by further inducing placental prostaglandin production and uterine contractions. Figure created with biorender.com.

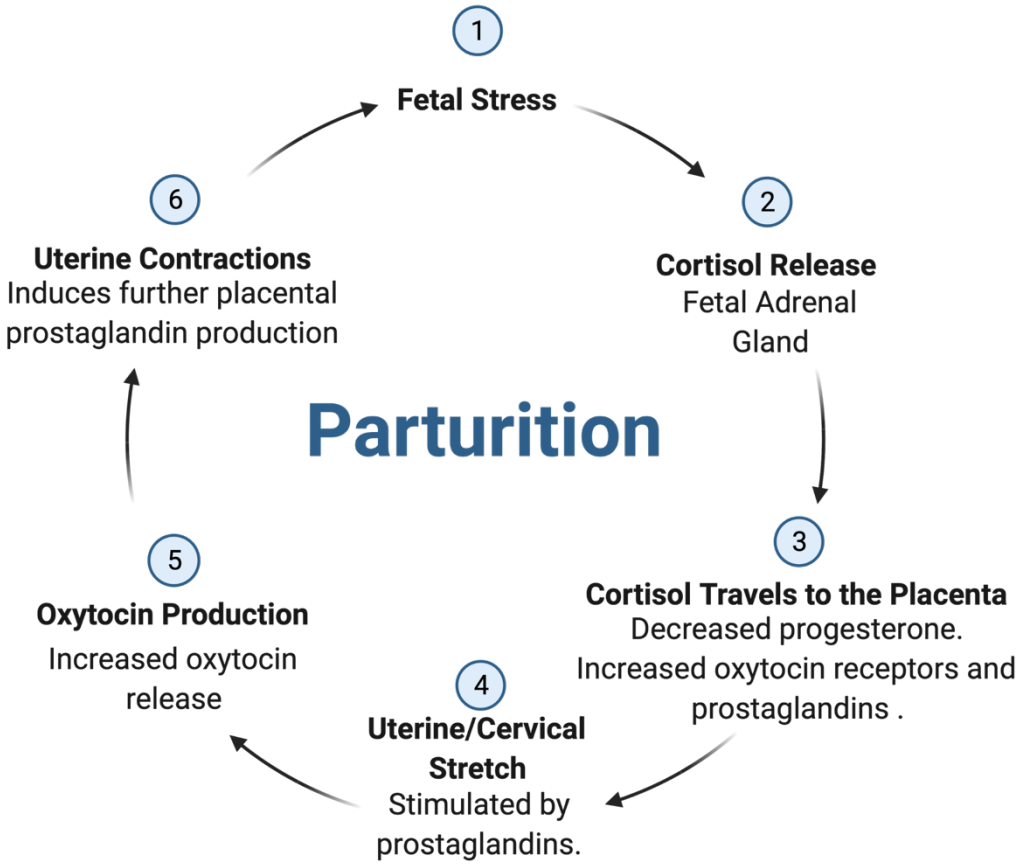


Figure 1.3 The eicosanoid pathway. The eicosanoid pathway begins with the liberation of free arachidonic acid from phospholipid bilayers by phospholipase A_2. Free arachidonic acid can be shunted to three major pathways – the cytochrome P450/epoxygenase pathway, the lipoxygenase pathway, or the cyclooxygenase pathway. Enzymatic processing of arachidonic acid by cyclooxygenase enzymes yields prostaglandin H_2 which can be further enzymatically processed into other prostaglandins. Figure created with biorender.com.

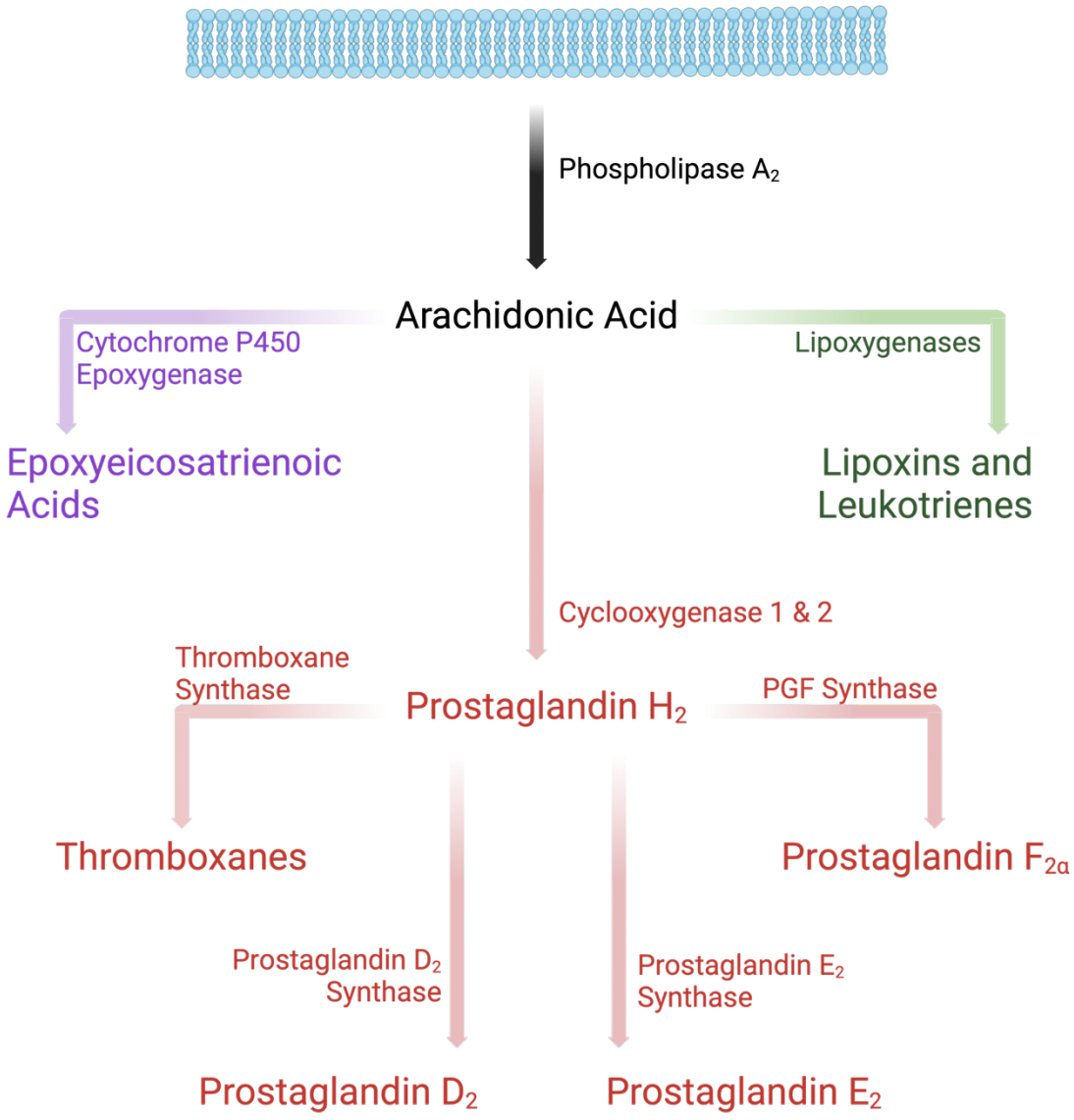


Figure 1.4 *Listeria monocytogenes* infectious cycle. The bacterial pathogen *Listeria monocytogenes* enters the host cell via phagocytosis or internalin-mediated entry (1). Once in the phagocytic lysosome (2), *L. monocytogenes* expresses the beta hemolysin listeriolysin O to break down the lysosome membrane, releasing the bacterium into the cytosol (3). In the cytosol, *L. monocytogenes* can multiply and can use the virulence factor ActA to polymerize host actin (4), creating actin rockets that propel it through the cell membrane and into neighboring cells (5) where the cycle can begin again.

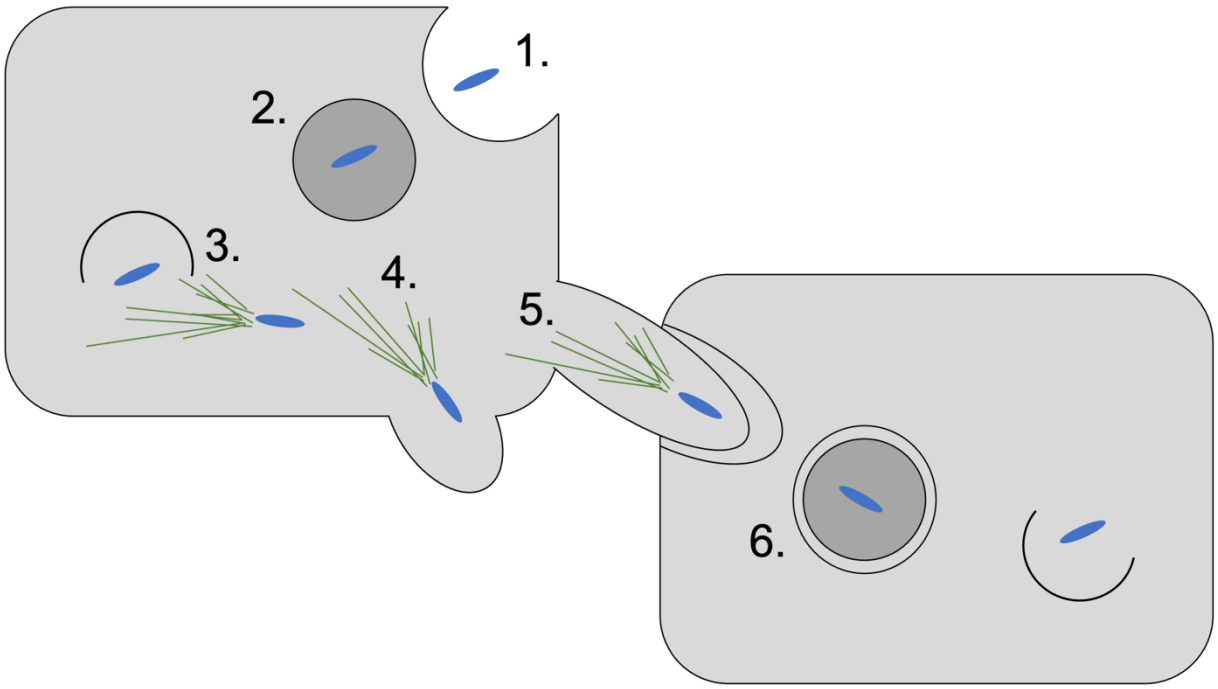


Figure 1.5 General internalin structure. The internalins are a large family of proteins in the *Listeria* genus. The internalins can be divided into two main categories: anchored (A) and secreted (B). The anchored internalins harbor C-terminal LPTXG or GW domains that anchor them into the *Listeria* cell membrane. Alternatively, secreted internalins contain an N-terminal signal peptide and tend to be much smaller than the anchored internalins. All internalins share a characteristic core of leucine rich repeats which can vary in size and number. Figure created with biorender.com.

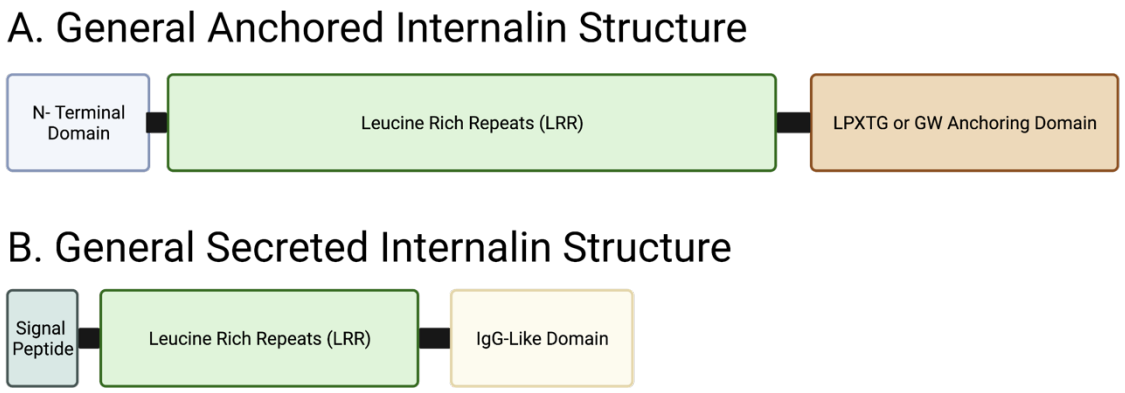


Table 1.1 Comparison of models used in placental studies

Species/Model	CAP Shape	Histological Structure	Fetal-Maternal Interface	Genetic Malleability	Cost
Human (Explants)	Discoid	Hemomonochorial	Villous	N/A	$$$
Mouse	Discoid	Hemotrichorial	Labyrinthine	++++	$$
Non-Human Primate	Discoid	Hemomonochorial	Villous	++	$$$$$
Guinea Pig	Discoid	Hemomonochorial	Labyrinthine	++	$$$

REFERENCES

REFERENCES

[1] G.J. Burton, A.L. Fowden, The placenta: a multifaceted, transient organ, Phil. Trans. R. Soc. B. 370 (2015) 20140066. https://doi.org/10.1098/rstb.2014.0066.

[2] M.Y. Turco, A. Moffett, Development of the human placenta, Development. 146 (2019) dev163428. https://doi.org/10.1242/dev.163428.

[3] Y. Wang, S. Zhao, Vascular Biology of the Placenta, Morgan & Claypool Life Sciences, San Rafael (CA), 2010. http://www.ncbi.nlm.nih.gov/books/NBK53247/ (accessed May 21, 2022).

[4] J.R. Robbins, K.M. Skrzypczynska, V.B. Zeldovich, M. Kapidzic, A.I. Bakardjiev, Placental syncytiotrophoblast constitutes a major barrier to vertical transmission of *Listeria monocytogenes*, PLoS Pathog. 6 (2010) e1000732. https://doi.org/10.1371/journal.ppat.1000732.

[5] J.A. Guttman, B.B. Finlay, Tight junctions as targets of infectious agents, Biochim Biophys Acta. 1788 (2009) 832–841. https://doi.org/10.1016/j.bbamem.2008.10.028.

[6] V.B. Zeldovich, C.H. Clausen, E. Bradford, D.A. Fletcher, E. Maltepe, J.R. Robbins, A.I. Bakardjiev, Placental syncytium forms a biophysical barrier against pathogen invasion, PLoS Pathog. 9 (2013) e1003821. https://doi.org/10.1371/journal.ppat.1003821.

[7] J.R. Robbins, A.I. Bakardjiev, Pathogens and the placental fortress, Curr. Opin. Microbiol. 15 (2012) 36–43. https://doi.org/10.1016/j.mib.2011.11.006.

[8] S. Vannuccini, C. Bocchi, F.M. Severi, J.R. Challis, F. Petraglia, Endocrinology of human parturition, Annales d'Endocrinologie. 77 (2016) 105–113. https://doi.org/10.1016/j.ando.2016.04.025.

[9] S. Mesiano, E.-C. Chan, J.T. Fitter, K. Kwek, G. Yeo, R. Smith, Progesterone Withdrawal and Estrogen Activation in Human Parturition Are Coordinated by Progesterone Receptor A Expression in the Myometrium, The Journal of Clinical Endocrinology & Metabolism. 87 (2002) 2924–2930. https://doi.org/10.1210/jcem.87.6.8609.

[10] T. Zakar, F. Hertelendy, Progesterone withdrawal: key to parturition, American Journal of Obstetrics and Gynecology. 196 (2007) 289–296. https://doi.org/10.1016/j.ajog.2006.09.005.

[11] L.T. Goldsmith, G. Weiss, Relaxin in Human Pregnancy, Annals of the New York Academy of Sciences. 1160 (2009) 130–135. https://doi.org/10.1111/j.1749-6632.2008.03800.x.

[12] G. Weiss, Endocrinology of Parturition, The Journal of Clinical Endocrinology & Metabolism. 85 (2000) 4421–4425. https://doi.org/10.1210/jcem.85.12.7074.

[13] R.M. Kamel, The onset of human parturition, Arch Gynecol Obstet. 281 (2010) 975–982. https://doi.org/10.1007/s00404-010-1365-9.

[14] J.R.G. Challis, S.G. Matthews, W. Gibb, S.J. Lye, Endocrine and Paracrine Regulation of Birth at Term and Preterm, Endocrine Reviews. 21 (2000) 514–550. https://doi.org/10.1210/edrv.21.5.0407.

[15] J.R.G. Challis, S.J. Lye, W. Gibb, Prostaglandins and Parturition, Ann NY Acad Sci. 828 (1997) 254–267. https://doi.org/10.1111/j.1749-6632.1997.tb48546.x.

[16] R. Bakker, S. Pierce, D. Myers, The role of prostaglandins E1 and E2, dinoprostone, and misoprostol in cervical ripening and the induction of labor: a mechanistic approach, Arch Gynecol Obstet. 296 (2017) 167–179. https://doi.org/10.1007/s00404-017-4418-5.

[17] E. Ricciotti, G.A. FitzGerald, Prostaglandins and Inflammation, Arterioscler Thromb Vasc Biol. 31 (2011) 986–1000. https://doi.org/10.1161/ATVBAHA.110.207449.

[18] A.M. Vasquez, V.D. Mouchlis, E.A. Dennis, Review of four major distinct types of human phospholipase A2, Advances in Biological Regulation. 67 (2018) 212–218. https://doi.org/10.1016/j.jbior.2017.10.009.

[19] W.L. Smith, D.L. DeWitt, R.M. Garavito, Cyclooxygenases: Structural, Cellular, and Molecular Biology, Annu. Rev. Biochem. (2000) 145–182.

[20] D.M. Olson, C. Amman, Role of the prostaglandins in labour and prostaglandin receptor inhibitors in the prevention of preterm labour, Frontiers in Bioscience. 12 (2007) 1329–1343.

[21] J.R.G. Challis, D.M. Sloboda, N. Alfaidy, S.J. Lye, W. Gibb, F.A. Patel, W.L. Whittle, J.P. Newham, Prostaglandins and mechanisms of preterm birth, Reproduction. 124 (2002) 1470–1626.

[22] L.D. Levine, Cervical ripening: Why we do what we do, Seminars in Perinatology. (2019) 151216. https://doi.org/10.1016/j.semperi.2019.151216.

[23] L. Liu, H.L. Johnson, S. Cousens, J. Perin, S. Scott, J.E. Lawn, I. Rudan, H. Campbell, R. Cibulskis, M. Li, C. Mathers, R.E. Black, Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000, The Lancet. 379 (2012) 2151–2161. https://doi.org/10.1016/S0140-6736(12)60560-1.

[24] M.K. Mwaniki, M. Atieno, J.E. Lawn, C.R. Newton, Long-term neurodevelopmental outcomes after intrauterine and neonatal insults: a systematic review, The Lancet. 379 (2012) 445–452. https://doi.org/10.1016/S0140-6736(11)61577-8.

[25] C.J.L. Murray *et al.* Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, The Lancet. 380 (2012) 2197–2223. https://doi.org/10.1016/S0140-6736(12)61689-4.

[26] R. Romero, S.K. Dey, S.J. Fisher, Preterm labor: One syndrome, many causes, Science. 345 (2014) 760–765. https://doi.org/10.1126/science.1251816.

[27] R.L. Goldenberg, J.F. Culhane, J.D. Iams, R. Romero, Epidemiology and causes of preterm birth, Lancet. 371 (2008) 75–84. https://doi.org/10.1016/S0140-6736(08)60074-4.

[28] R. Romero, R. Gomez, T. Chaiworapongsa, G. Conoscenti, J. Cheol Kim, Y. Mee Kim, The role of infection in preterm labour and delivery, Paediatr Perinat Epidemiol. 15 (2001) 41–56. https://doi.org/10.1046/j.1365-3016.2001.00007.x.

[29] M.A. Elovitz, Z. Wang, E.K. Chien, D.F. Rychlik, M. Phillippe, A New Model for Inflammation-Induced Preterm Birth, The American Journal of Pathology. 163 (2003) 2103–2111. https://doi.org/10.1016/S0002-9440(10)63567-5.

[30] V. Agrawal, E. Hirsch, Intrauterine infection and preterm labor, Seminars in Fetal and Neonatal Medicine. 17 (2012) 12–19. https://doi.org/10.1016/j.siny.2011.09.001.

[31] S. Furukawa, Y. Kuroda, A. Sugiyama, A Comparison of the Histological Structure of the Placenta in Experimental Animals, J Toxicol Pathol. 27 (2014) 11–18. https://doi.org/10.1293/tox.2013-0060.

[32] D.E. Lowe, J.R. Robbins, A.I. Bakardjiev, Animal and Human Tissue Models of Vertical *Listeria monocytogenes* Transmission and Implications for Other Pregnancy-Associated Infections, Infect. Immun. 86 (2018). https://doi.org/10.1128/IAI.00801-17.

[33] S.L. Adamson, Y. Lu, K.J. Whiteley, D. Holmyard, M. Hemberger, C. Pfarrer, J.C. Cross, Interactions between trophoblast cells and the maternal and fetal circulation in the mouse placenta, Dev. Biol. 250 (2002) 358–373.

[34] J.A. Vazquez-Boland, M. Kuhn, P. Berche, T. Chakraborty, G. Dominguez-Bernal, W. Goebel, B. Gonzalez-Zorn, J. Wehland, J. Kreft, *Listeria* Pathogenesis and Molecular Virulence Determinants, Clinical Microbiology Reviews. 14 (2001) 584–640. https://doi.org/10.1128/CMR.14.3.584-640.2001.

[35] *Listeria* (Listeriosis), Centers for Disease Control and Prevention, n.d. https://www.cdc.gov/listeria/index.html (accessed March 6, 2022).

[36] *Listeria* During Pregnancy, American Pregnancy Association, n.d. https://americanpregnancy.org/healthy-pregnancy/pregnancy-concerns/listeria-during-pregnancy/.

[37] Z. Wang, X. Tao, S. Liu, Y. Zhao, X. Yang, An Update Review on *Listeria* Infection in Pregnancy, IDR. Volume 14 (2021) 1967–1978. https://doi.org/10.2147/IDR.S313675.

[38] M. Doumith, C. Buchrieser, P. Glaser, C. Jacquet, P. Martin, Differentiation of the Major *Listeria monocytogenes* Serovars by Multiplex PCR, J Clin Microbiol. 42 (2004) 3819–3822. https://doi.org/10.1128/JCM.42.8.3819-3822.2004.

[39] M. Doumith, C. Cazalet, N. Simoes, L. Frangeul, C. Jacquet, F. Kunst, P. Martin, P. Cossart, P. Glaser, C. Buchrieser, New Aspects Regarding Evolution and Virulence of *Listeria monocytogenes* Revealed by Comparative Genomics and DNA Arrays, Infect Immun. 72 (2004) 1072–1083. https://doi.org/10.1128/IAI.72.2.1072-1083.2004.

[40] J. McLauchlin, R.T. Mitchell, W.J. Smerdon, K. Jewell, *Listeria monocytogenes* and listeriosis: a review of hazard characterisation for use in microbiological risk assessment of foods, International Journal of Food Microbiology. 92 (2004) 15–33. https://doi.org/10.1016/S0168-1605(03)00326-X.

[41] J. Mengaud, H. Ohayon, P. Gounon, R.-M. Mège, P. Cossart, E-Cadherin Is the Receptor for Internalin, a Surface Protein Required for Entry of *L. monocytogenes* into Epithelial Cells, Cell. 84 (1996) 923–932. https://doi.org/10.1016/S0092-8674(00)81070-3.

[42] Y. Shen, M. Naujokas, M. Park, K. Ireton, InlB-Dependent Internalization of *Listeria* Is Mediated by the Met Receptor Tyrosine Kinase, Cell. 103 (2000) 501–510. https://doi.org/10.1016/S0092-8674(00)00141-0.

[43] C. Geoffroy, J.L. Gaillard, J.E. Alouf, P. Berche, Purification, characterization, and toxicity of the sulfhydryl-activated hemolysin listeriolysin O from *Listeria monocytogenes*, Infect Immun. 55 (1987) 1641–1646. https://doi.org/10.1128/iai.55.7.1641-1646.1987.

[44] E. Domann, J. Wehland, M. Rohde, S. Pistor, M. Hartl, W. Goebel, M. Leimeister-Wächter, M. Wuenscher, T. Chakraborty, A novel bacterial virulence gene in *Listeria monocytogenes* required for host cell microfilament interaction with homology to the proline-rich region of vinculin., The EMBO Journal. 11 (1992) 1981–1990. https://doi.org/10.1002/j.1460-2075.1992.tb05252.x.

[45] C. Kocks, E. Gouin, M. Tabouret, P. Berche, H. Ohayon, P. Cossart, *L. monocytogenes*-induced actin assembly requires the actA gene product, a surface protein, Cell. 68 (1992) 521–531. https://doi.org/10.1016/0092-8674(92)90188-I.

[46] L.G. Tilney, D.A. Portnoy, Actin filaments and the growth, movement, and spread of the intracellular bacterial parasite, *Listeria monocytogenes*., Journal of Cell Biology. 109 (1989) 1597–1608. https://doi.org/10.1083/jcb.109.4.1597.

[47] K. Ireton, R. Mortuza, G.C. Gyanwali, A. Gianfelice, M. Hussain, Role of internalin proteins in the pathogenesis of *Listeria monocytogenes*, Mol Microbiol. (2021) mmi.14836. https://doi.org/10.1111/mmi.14836.

[48] B. Kobe, The leucine-rich repeat as a protein recognition motif, Current Opinion in Structural Biology. 11 (2001) 725–732. https://doi.org/10.1016/S0959-440X(01)00266-4.

[49] B. Kobe, J. Deisenhofer, The leucine-rich repeat: a versatile binding motif, Trends in Biochemical Sciences. 19 (1994) 415–421. https://doi.org/10.1016/0968-0004(94)90090-6.

[50] M. Lecuit, H. Ohayon, L. Braun, J. Mengaud, P. Cossart, Internalin of *Listeria monocytogenes* with an intact leucine-rich repeat region is sufficient to promote internalization, Infect Immun. 65 (1997) 5309–5319. https://doi.org/10.1128/iai.65.12.5309-5319.1997.

[51] F. Engelbrecht, S.-K. Chun, C. Ochs, J. Hess, F. Lottspeich, W. Goebel, Z. Sokolovic, A new PrfA-regulated gene of *Listeria monocytogenes* encoding a small, secreted protein which belongs to the family of internalins, Molecular Microbiology. 21 (1996) 823–837. https://doi.org/10.1046/j.1365-2958.1996.541414.x.

[52] L. Polle, L.A. Rigano, R. Julian, K. Ireton, W.-D. Schubert, Structural Details of Human Tuba Recruitment by InlC of *Listeria monocytogenes* Elucidate Bacterial Cell-Cell Spreading, Structure. 22 (2014) 304–314. https://doi.org/10.1016/j.str.2013.10.017.

[53] P. Ghosh, E.M. Halvorsen, D.A. Ammendolia, N. Mor-Vaknin, M.X.D. O'Riordan, J.H. Brumell, D.M. Markovitz, D.E. Higgins, Invasion of the Brain by *Listeria monocytogenes* Is Mediated by InlF and Host Cell Vimentin, MBio. 9 (2018) e00160-18. https://doi.org/10.1128/mBio.00160-18.

[54] C. Faralla, G.A. Rizzuto, D.E. Lowe, B. Kim, C. Cooke, L.R. Shiow, A.I. Bakardjiev, InlP, a New Virulence Factor with Strong Placental Tropism, Infect. Immun. 84 (2016) 3584–3596. https://doi.org/10.1128/IAI.00625-16.

[55] C. Faralla, E.E. Bastounis, F.E. Ortega, S.H. Light, G. Rizzuto, L. Gao, D.K. Marciano, S. Nocadello, W.F. Anderson, J.R. Robbins, J.A. Theriot, A.I. Bakardjiev, *Listeria monocytogenes* InlP interacts with afadin and facilitates basement membrane crossing, PLoS Pathog. 14 (2018) e1007094. https://doi.org/10.1371/journal.ppat.1007094.

[56] A. Le Monnier, N. Autret, O.F. Join-Lambert, F. Jaubert, A. Charbit, P. Berche, S. Kayal, ActA Is Required for Crossing of the Fetoplacental Barrier by *Listeria monocytogenes*, Infection and Immunity. 75 (2007) 950–957. https://doi.org/10.1128/IAI.01570-06.

CHAPTER 2

INFECTION WITH *LISTERIA MONOCYTOGENES* ALTERS THE PLACENTAL TRANSCRIPTOME AND EICOSANOME

PUBLICATION NOTICE

The following dissertation chapter describes perturbations in placental gene expression and eicosanoid concentrations following infection with *L. monocytogenes*. The author i) assisted RNAseq data analysis, ii) assisted with animal infection and sample collection for semi-targeted lipidomics, iii) completed data analysis for semi-targeted lipidomics with the assistance of Dr. Todd Lydic, iv) generated all data visializations, v) wrote the manuscript. The RNA sequencing (infection, sample collection, and raw data collection) was conducted by Dr. Jonathan Hardy with assistance from Dr. Derek Holman at Stanford University prior to starting the Hardy laboratory at Michigan State University.

The following chapter has been submitted for publication to *Placenta* and is currently in peer review. It is available as a preprint through bioRxiv: "*K.N. Conner, D. Holman, T. Lydic, J.W. Hardy, Infection with Listeria monocytogenes alters the placental transcriptome and eicosanome, bioRxiv, https://doi.org/10.1101/2022.04.14.48831*."

ABSTRACT

Placental infection and inflammation are risk factors for adverse pregnancy outcomes, including preterm labor. However, the mechanisms underlying these outcomes are poorly understood. To study this response, we have employed a pregnant mouse model of placental infection caused by the bacterial pathogen *Listeria monocytogenes*, which infects the human placenta. Through *in vivo* bioluminescence imaging, we confirm the presence of placental infection and quantify relative infection levels. Infected and control placentas were collected on embryonic day 18 for RNA sequencing to evaluate gene expression signatures associated with infection by *Listeria*. We identified an enrichment of genes associated with eicosanoid biosynthesis, suggesting an increase in eicosanoid production in infected tissues. Because of the known importance of eicosanoids in inflammation and timing of labor, we quantified eicosanoid levels in infected and uninfected placentas using semi-targeted mass spectrometry. We found a significant increase in the concentrations of several key eicosanoids: leukotriene B4, lipoxin A4, prostaglandin A2, prostaglandin D2, and eicosatrienoic acid. Our study provides a likely explanation for dysregulation of the timing of labor following placental infection. Additionally, our results suggest potential biomarkers of placental pathology and targets for clinical intervention.

INTRODUCTION

To ensure the development of the allogeneic fetus, placental immune responses must be precisely balanced between protective immunity and deleterious inflammation [1,2]. Bacterial infection of the placenta can affect this balance, leading to adverse pregnancy outcomes even in the absence of severe disease [1,2]. One such infection is prenatal listeriosis caused by the Gram-positive bacterium *Listeria monocytogenes* (*Lm*). *Lm* is an opportunistic foodborne pathogen that primarily affects the immunocompromised, especially pregnant individuals, who are typically exposed to *Lm* through contaminated meat and dairy products [3]. Following ingestion, *Lm* invades the gut epithelium and traffics in maternal monocytes to the female reproductive organs where it uses cell to cell spread to invade the placenta [3]. Invasion of the placenta can result in a myriad of adverse pregnancy outcomes including preterm labor and downstream abnormal development of the offspring [4–6]. Despite great strides that have been made in the understanding of *Lm* invasion of the placenta, little information is available on the molecular mechanisms underlying listeriosis-associated preterm labor.

Labor and parturition are complicated processes controlled by many genetic, metabolic, and physical factors within the female reproductive tract. Eicosanoids, a family of hormone-like fatty acids, dictate the timing of labor by signaling cervical ripening, breaking down fetal membranes, and promoting myometrial contractility [7–9]. These lipids are produced enzymatically by all cells in the body beginning with the liberation of arachidonic acid from cell membrane phospholipids [10]. Downstream processing by cyclooxygenase (COX) and lipoxygenase (LOX) enzymes yields the two eicosanoid classes: prostaglandins and lipoxins, respectively [10]. Eicosanoids are key players in the delicate balance between protective immunity

and deleterious inflammation throughout the body, including the placenta [10]. While associations have been made between eicosanoid pathway perturbations and placental pathology, little information exists regarding infection-induced perturbations to the eicosanoid pathway and downstream consequences in the placenta.

Due to its well characterized lifecycle and genetic malleability, *Lm* has been used as a model for placental infection for decades [11]. In this study, we use a pregnant CD1 mouse model of bioluminescent *Lm* placental infection to begin exploring infection-induced eicosanoid pathway perturbations. We demonstrate through RNA sequencing that mouse placentas colonized with *Lm* have gene expression profiles associated with placental dysfunction and preterm labor. We verify, using semi-targeted mass spectrometry, that these aberrant gene expression profiles result in significant changes to placental eicosanoid concentrations, which we refer to as the placental eicosanome. Together, our data identify a likely mechanism for the induction of preterm labor associated with placental listeriosis infection.

MATERIALS AND METHODS

Strains/Bacterial Culture

The bacterial strain used in this study is the bioluminescent *Listeria monocytogenes* strain Xen32 (Perkin Elmer, Inc.). Cultures were grown overnight, shaking at 37°C in brain heart infusion (BHI) broth supplemented with kanamycin for selection. On the day of mouse infection, overnight cultures were subcultured in fresh BHI supplemented with kanamycin for selection and grown to an OD_{600} of 0.5. The subculture was then diluted in sterile phosphate buffered saline (PBS) to yield 10^6 colony forming units (CFU) per mL.

Animals and *In Vivo* Imaging

We All mouse experiments were approved by the Institutional Animal Care and Use Committees at Michigan State University and Stanford University. Mice were housed at the Stanford University Research Animal Facility and the Michigan State University Clinical Center animal facility under the care of Campus Animal Resources. The BSL-2 animal procedures were approved under Stanford University Protocol 12342 (formerly 8158) and Michigan State University Animal Use Protocol 201800030. Timed gestation day 11 (E11) pregnant CD-1 mice were delivered on that day from Charles River Laboratories. On E14.5, mice were infected via tail vein injection with 2 x 10^5 CFU of *Listeria monocytogenes* Xen32 in 200mL phosphate buffered saline prepared as described above *(see "Strains/Bacterial Culture*"). Uninfected control mice were not injected. On E18.5, mice were imaged using the PerkinElmer In Vivo Imaging System (IVIS) to confirm placental infection, then humanely sacrificed under anesthesia according to approved guidelines. Uterine horns were immediately excised and imaged separately using the

IVIS to identify infected placentas. Placentas were excised and snap frozen on dry ice then frozen at -80°C for downstream analyses. All animals were imaged using the IVIS for 5 minutes prior to euthanasia, and uterine horns were imaged for 1 minute following excision. Image analysis was performed using the Living Image software by Caliper Life Sciences, and average radiance (light intensity) is expressed as photons per second per centimeter squared per steradian (photons/s/cm^2/str).

RNA Sequencing

Twenty infected and four uninfected mouse placentas were excised for downstream RNA sequencing (RNAseq). Tissues were snap frozen on dry ice and stored at -80°C until homogenization. Tissues were homogenized by suspending them in Qiagen Buffer RLT and passing them each subsequently through 16G, 18G, 20G, and 22G needles. Total RNA was extracted from each placenta using the Qiagen RNeasy Midi kit and DNase treated with DNase I (Qiagen) according to the manufacturer's instructions. Isolated RNA was analyzed for RNA integrity (RIN) values by the Stanford PAN Facility prior to submission for RNAseq analysis by SeqMatic Inc., Mountain View, CA. Single-read sequencing on libraries was performed using the Illumina Genome Analyzer IIx. Data was analyzed on the Galaxy webserver [12]. Raw read files from RNAseq analysis were assessed for quality using FastQC [13], and adapters were removed using Trimmomatic sliding window trimming [14]. To align reads to the mouse reference genome (GRCm39), we used Bowtie2 [15], and resulting alignment files were analyzed for read counts with FeatureCounts [16]. Finally, differential expression analysis was carried out using DESeq2 [17]. Gene ontology and pathway analyses were performed by submitting respective lists for significantly up- and down-regulated genes to g:Profiler with default options

(https://biit.cs.ut.ee/gprofiler/gost) [18]. Gene ontology networks were generated using GOnet with custom GO terms related to the eicosanoid pathway (https://tools.dice-database.org/GOnet/) [19].

Lipidomics

Semi-targeted mass spectrometry (MS) analysis was performed on six infected and six uninfected mouse placentas that had been snap frozen and kept at -80°C. Placentas were homogenized in methanol acidified with formic acid. Samples were then incubated overnight at -20°C for protein precipitation, then centrifuged. Supernatants were subjected to solid phase extraction using Phenomenex Strata-X 33-micron SPE columns as previously described to concentrate eicosanoids and remove biological matrix components. Eluates were reconstituted in methanol containing 0.01% butylated hydroxytoluene, then centrifuged immediately prior to analysis. Fatty acids and their oxygenated derivatives were analyzed by high resolution/accurate mass (HRAM)-LC-MS. Data-dependent product ion spectra were collected on the four most abundant ions at 30,000x resolution using the FT analyzer. Lipidomics data was analyzed using the Metaboanlyst software according to statistical methods previously published by Xi *et al* [20,21]. Lipid concentrations were normalized to placenta mass, log transformed, and subjected to Pareto scaling prior to statistical analyses.

Availability of Data and Materials

Raw sequencing files, normalized count tables, and DESeq2 outputs are provided as supplementary material and can be accessed through the NCBI Gene Expression Omnibus under Accession GSE201038.

RESULTS

Placental infection by *Lm* alters placental gene expression

At the dose of $2x10^5$ colony forming units (CFU) of bioluminescent *Lm* in pregnant CD1 mice, a range of infection levels is observed across placentas in a single uterine horn, which permits the analysis of many outcomes of prenatal listeriosis including stillbirth and fetal abnormality [22]. Using this model, *in vivo* bioluminescence imaging (BLI) was employed to identify and isolate infected placentas (**Fig 1**). RNA sequencing analysis of 20 infected placentas and 4 control placentas from uninfected animals revealed 498 significantly underexpressed and 862 significantly overexpressed ($Log_2FC \leq -1$ or ≥ 1; adjusted P value ≤ 0.05) in the infected placentas **(Fig. 2A, Supplementary Material)**. The top five overexpressed genes following infection included *Zbp1*, *GM12250*, *Igtp*, *Tap1*, and *Ido1* **(Fig. 2A, Supplementary Material).** These results were expected considering the various immunoregulatory roles these genes are known to play. We observed minimal variability in the four uninfected sample gene expression profiles, which formed their own distinct cluster (**Fig. 2B**). Conversely, the infected sample gene expression profiles displayed considerable variability, which is consistent with the range of infection levels in our model (**Fig. 1, Fig. 2B**).

To better understand the pathways associated with significantly dysregulated genes, we performed functional profiling using g:Profiler. This analysis revealed several pathways of interest in both up- and down-regulated gene data sets (**Supplementary Material**). Interestingly, pathways associated with underexpressed genes were largely related to ion transport across the membrane (**Supplementary Material**). As expected, most pathways associated with overexpressed genes were linked to pro-inflammatory processes typical of bacterial infection, consistent with the expected infiltration and activation of immune cells (**Supplementary Material**). Notably, GO terms related to prostanoid and prostaglandin biosynthesis were enriched in our upregulated gene data set (**Supplementary Material**). To visualize overexpressed gene networks associated with eicosanoid metabolism, we submitted our overexpressed genes and custom gene ontology ID list to GOnet to generate a visual custom gene ontology network (**Fig. 2C**).

Following gene ontology analysis, we became interested in the enrichment of eicosanoid metabolism genes due to the known roles of eicosanoids in pregnancy and listeriosis elsewhere in the body. In our RNAseq data, we observed a significant overexpression (approximately 2.3-fold increase, adjusted P value ≤ 0.05) in the *Ptgs2* gene encoding cyclooxygenase 2, a key enzyme in the eicosanoid pathway (**Fig. 2C, Supplementary Material**). While *Ptgs2* encoding cyclooxygenase 2 was significantly overexpressed, the *Ptgs1* gene encoding cyclooxygenase 1 (the constitutive housekeeping isoform of this enzyme) was not significantly dysregulated (**Supplementary Material**). In addition to *Ptgs2*, we observed overexpression of several other eicosanoid-associated genes (**Fig. 2C, Supplementary Material**). We hypothesized that, due to overexpression of several genes associated with eicosanoid production, the concentrations of these lipids would be increased in infected placentas. Specifically, we hypothesized that infected placentas would harbor increased concentrations of prostaglandins due to the upregulation of

several enzymes implicated in prostaglandin synthesis (**Fig. 3**). In addition, because eicosanoid pathway enzymes can be regulated by post-transcriptional mechanisms including allosteric induction [23], it was important to measure the pathway products themselves to fully characterize changes in this pathway.

Lm infection alters eicosanoid concentrations in the placenta

Because we wanted to know if eicosanoid levels were perturbed along with eicosanoid pathway gene expression, we carried out semi-targeted mass spectrometry to measure concentrations of various eicosanoids in infected and uninfected placentas (**Fig. 1**)**.** Our analysis revealed distinct profiles for infected versus uninfected placentas **(Fig. 4).** We observed 12 eicosanoids showing a ≥2-fold increase or decrease in concentration in the placenta following infection with *Lm* (**Fig. 4**). Strikingly, leukotriene B_4 (LTB_4) exhibited a ~25-fold increase following infection (**Fig. 4, Supplementary Material**). Also of note were prostaglandin A_2 (PGA2), prostaglandin E_2 (PGE_2), prostaglandin D_2 (PGD_2), and prostaglandin $F_{2\alpha}$ ($PGF_{2\alpha}$) which showed ~4.8-, ~2.4-, ~2.1, and ~2.3-fold increases following infection, respectively (**Fig. 4, Supplementary Material**). Of these dysregulated eicosanoids, nine reached statistical significance ($p \leq 0.05$) including LTB_4, LXA_4, PGA_2, PGD_2, and eicosatrienoic acid (**Fig. 5, Supplementary Material)**. Together, these data supported our hypothesis that altered gene expression in the placenta results in changes in placental eicosanoid profiles, which we refer to as the placental eicosonome.

DISCUSSION

Pregnancy complications including preterm birth are relatively common, and preterm birth is the leading cause of infant mortality worldwide [24,25]. While many factors can contribute to the occurrence of preterm birth, the outcome can be developmentally devastating for the infant. Infants born prematurely are more likely to exhibit breathing problems, sensory problems, and developmental delay [26]. Infection is a well-known cause of preterm birth, necessitating studies of prenatal responses to distinct pathogens [25]. Because of the crucial role of the placenta in immune responses during pregnancy, pathogens that infect this organ are especially important to understand. For example, it will be crucial to distinguish placental infection from other prenatal infections such as chorioamnionitis, which may elicit completely different responses and require different interventions. In addition, placental infection can induce inflammatory responses, which have been associated with preterm birth [27]. Therefore, animal models of placental infection are vital tools in understanding preterm birth.

Listeria monocytogenes is a known placental pathogen that can cause preterm labor as well as other perinatal pathologies [11,28]. Animal models of prenatal listeriosis have revealed details of placental infection, including the target cell type, bacterial virulence factors and molecular mechanisms of invasion [29]. However, host placental responses to this bacterium have not been previously defined and may reveal clues as to the function of the placenta in prenatal resistance to infection. Our data sheds light on the molecular and metabolic mechanisms underlying listeriosis-induced preterm labor. We have shown using a pregnant mouse model of placental listeriosis that infected placentas harbor distinct gene expression profiles compared to their uninfected counterparts. Unsurprisingly, we have identified an enrichment of genes associated with

inflammation and response to infection in infected placentas. We were particularly interested to observe an enrichment of genes associated with eicosanoid biosynthesis and metabolism following infection. Though this result is not entirely surprising due to the role of eicosanoids in inflammation, it was noteworthy considering that eicosanoids are known to play critical roles in the regulation of labor, as well as other aspects of pregnancy such as placental function [30]. This discovery warrants further investigation due to previous associations between placental eicosanoid dysregulation and pathological pregnancy outcomes in previous studies [31].

To determine if eicosanoid concentrations were perturbed along with gene expression profiles, we employed a semi-targeted mass spectrometry approach to quantify the eicosanoid concentrations in infected and uninfected mouse placentas. This analysis highlighted perturbations in eicosanoid concentrations in infected placentas. We noted significant increases in the concentrations of LTB_4, LXA_4, PGA_2, PGD_2, and eicosatrienoic acid. Previous studies strongly support the association between the eicosanoids we have identified as increased in placental infection and placental pathology, including LTB_4 [32], LXA_4 [33], and PGD_2 [34].

Broadening our understanding of molecular mechanisms underlying listeriosis-induced adverse pregnancy outcomes has the potential to propel the development of improved clinical interventions for pregnancy associated listeriosis and other placental infections. Our study offers insight into the genetic and metabolic changes that take place in the placenta following *Lm* infection. While our study begins to offer possible mechanisms of listeriosis-induced preterm labor, much remains to be investigated.

To our knowledge, this is the first study associating increased PGA_2 concentrations with placental infection or preterm labor. This is noteworthy as PGA_2 is a known degradation product resulting from the dehydration of PGE_2, which has been studied extensively for its role in the

timing and induction of parturition. Increased PGA_2 concentrations could imply an increase in upstream PGE_2 production and its subsequent degradation, which could be contribute to dysregulation of labor. Future studies should address the mechanistic role of this eicosanoid in the context of infection-induced preterm labor.

Our observations confirm that the known role of eicosanoids in infection and inflammation in other tissues also applies to the placenta, where the eicosanoids are also known to function in the timing of labor. It is noteworthy that many of the eicosanoids identified in our study have been implicated in pathological pregnancy outcomes and placental disease. In addition, the induction of specific prostaglandins and leukotrienes suggests the possibility of receptor-specific interventions. It is important to identify new detection and intervention methods that can be utilized to prevent adverse pregnancy outcome. We propose that future studies assess eicosanoid concentrations in maternal circulation to assess the usefulness of eicosanoids as clinical biomarkers of placental disease. Further, we suggest that eicosanoid synthesis and uptake be studied as a potential route of intervention in the prevention of infection-induced adverse pregnancy outcome.

DECLARATIONS

Competing interests

The authors declare that they have no competing interests.

Funding

KNC is supported by Michigan State University (MSU) Microbiology and Molecular Genetics departmental fellowships. This work was supported by MSU start-up funds granted to JWH.

Contributions

Conceptualization: JWH, KNC; Investigation: KNC, DH, TL, JWH; Analysis: TL, KNC; Writing: All Authors; Visualization: KNC; Project Administration: JWH; Resources and Funding Acquisition: JWH.

APPENDIX

FIGURES AND TABLES

Figure 2.1 *In vivo* bioluminescence imaging of *Lm* in the placenta. (A) Example of bioluminescence imaging of a pregnant mouse infected with *Lm* on E14.5 and imaged on E18.5. (B) Excised uterine horns from a similar animal showing the placentas used for RNAseq. RNA from infected placentas (arrows) was sequenced and compared to controls from uninfected mice. The false color scale is photons/second.

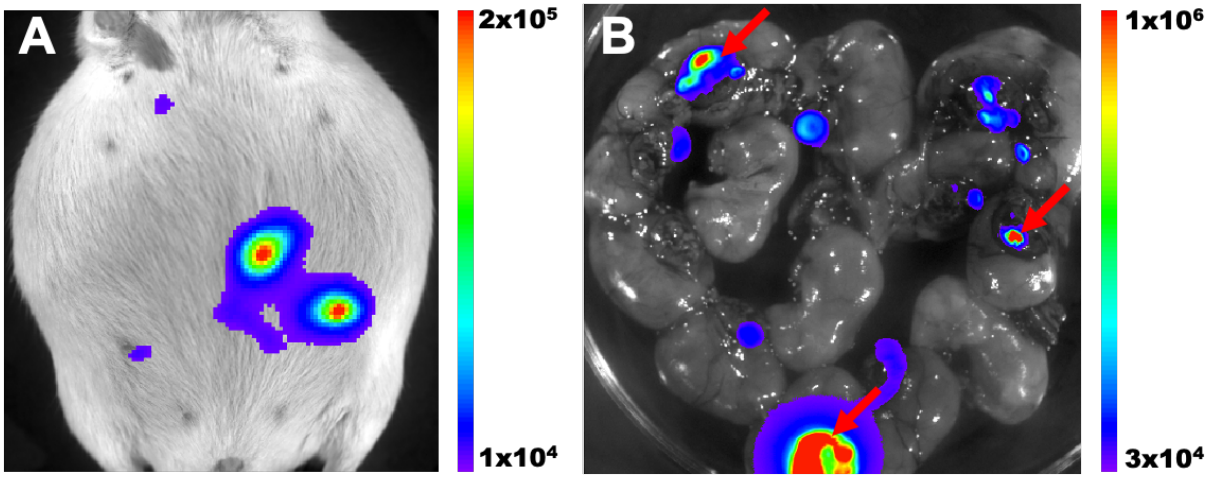


Figure 2.2 Gene expression profiles are altered in *Lm*-infected placentas. Differentially expressed genes in *Listeria*-infected placentas (compared to uninfected) were determined using DEseq2 and expressed as a volcano plot (A). Significantly overexpressed genes (fold change ≥ 2; Adjusted P value ≤ 0.05) are highlighted in red while significantly underexpressed genes (fold change ≤ -2; Adjusted P value ≤ 0.05) are highlighted in blue. Values are presented as Log_2 Fold Change and Log_{10} Adjusted P Value. The top 50 differentially expressed genes are expressed as a heatmap of normalized counts per sample (B). Heatmap was generated using Heatmapper [35], and sample clustering was computed with average linkage clustering and Euclidian distance measurement (represented by the sample dendogram). Gene ontology analysis was conducted using g:profiler, and network visualization was generated using GOnet (C). GO terms are in blue-green rectangles and gene names are in orange ovals.

A.

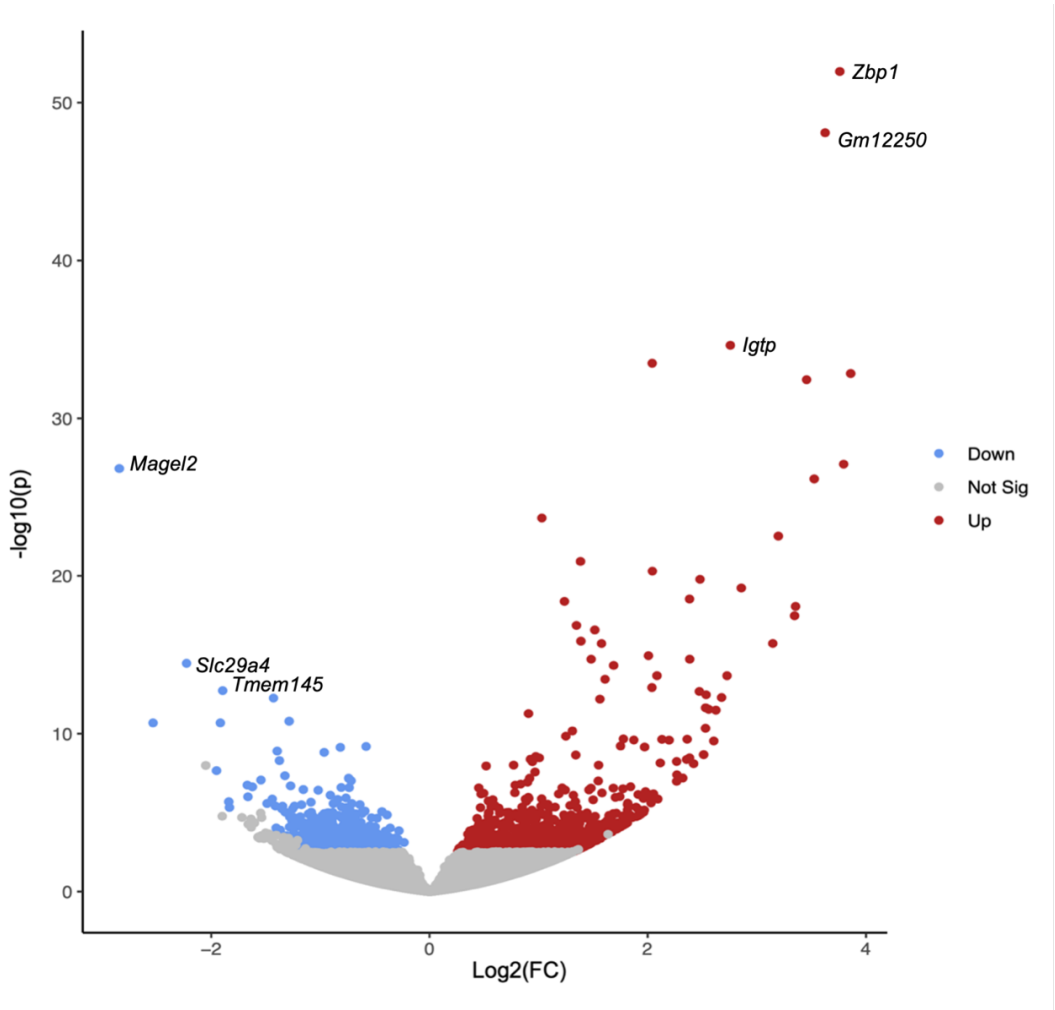


Figure 2.2 (Cont'd)

B.

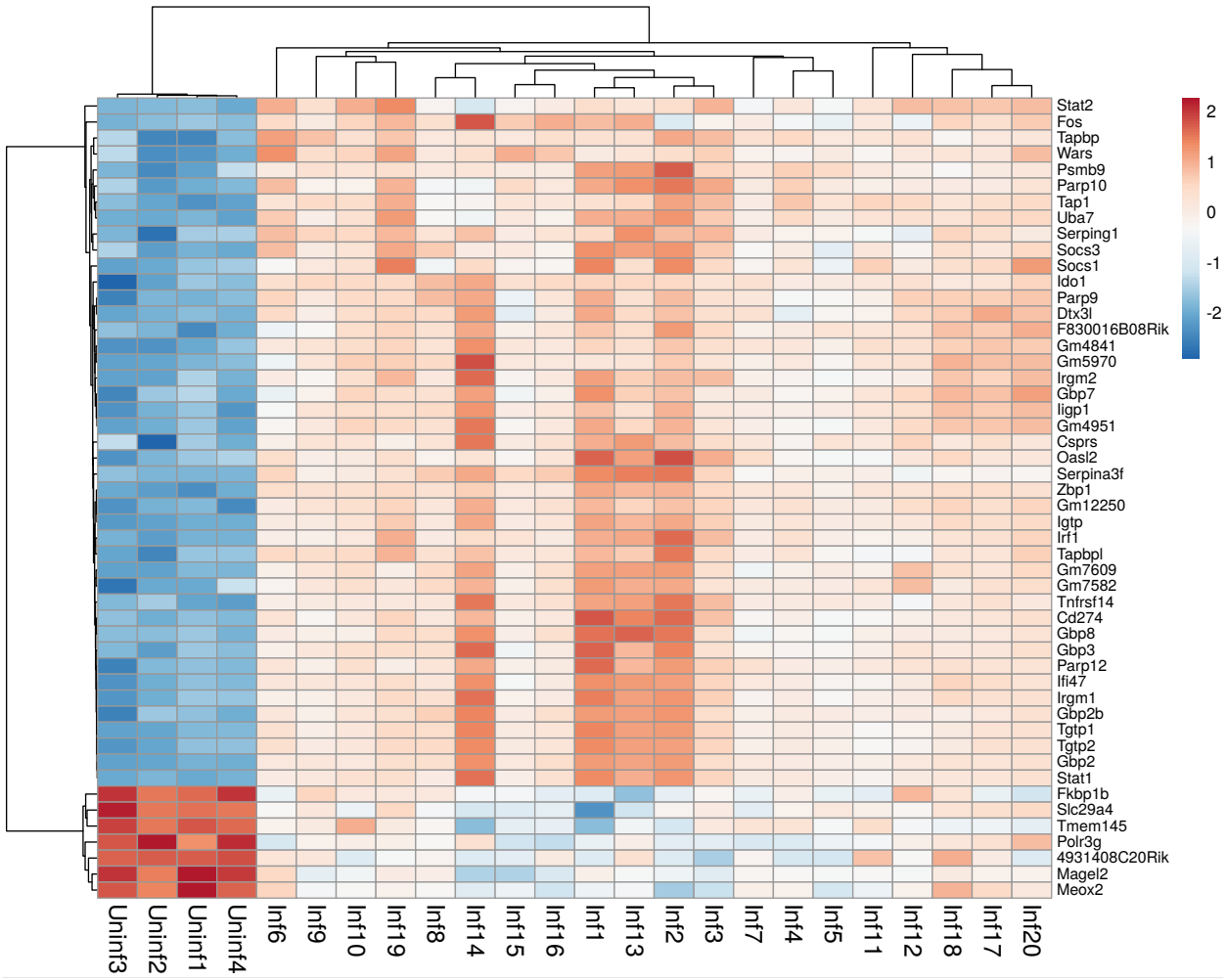


Figure 2.2 (Cont'd)

C.

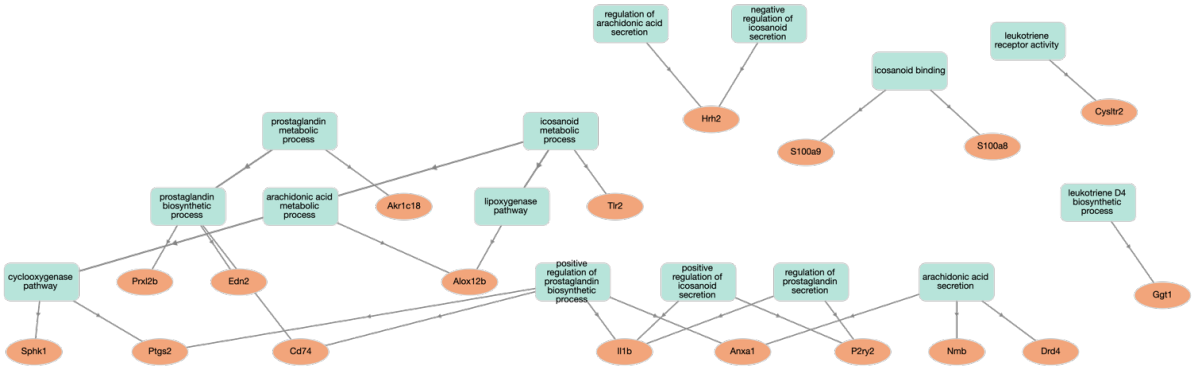


Figure 2.3 *Lm* infection results in upregulation of key eicosanoid pathway enzymes and increased concentrations of specific eicosanoids in the placenta. This adapted eicosanoid pathway figure illustrates the points at which this pathway is altered by listeriosis in the placenta. Genes for enzymes represented in blue text as well as eicosanoids represented in blue text are significantly overexpressed in our data sets.

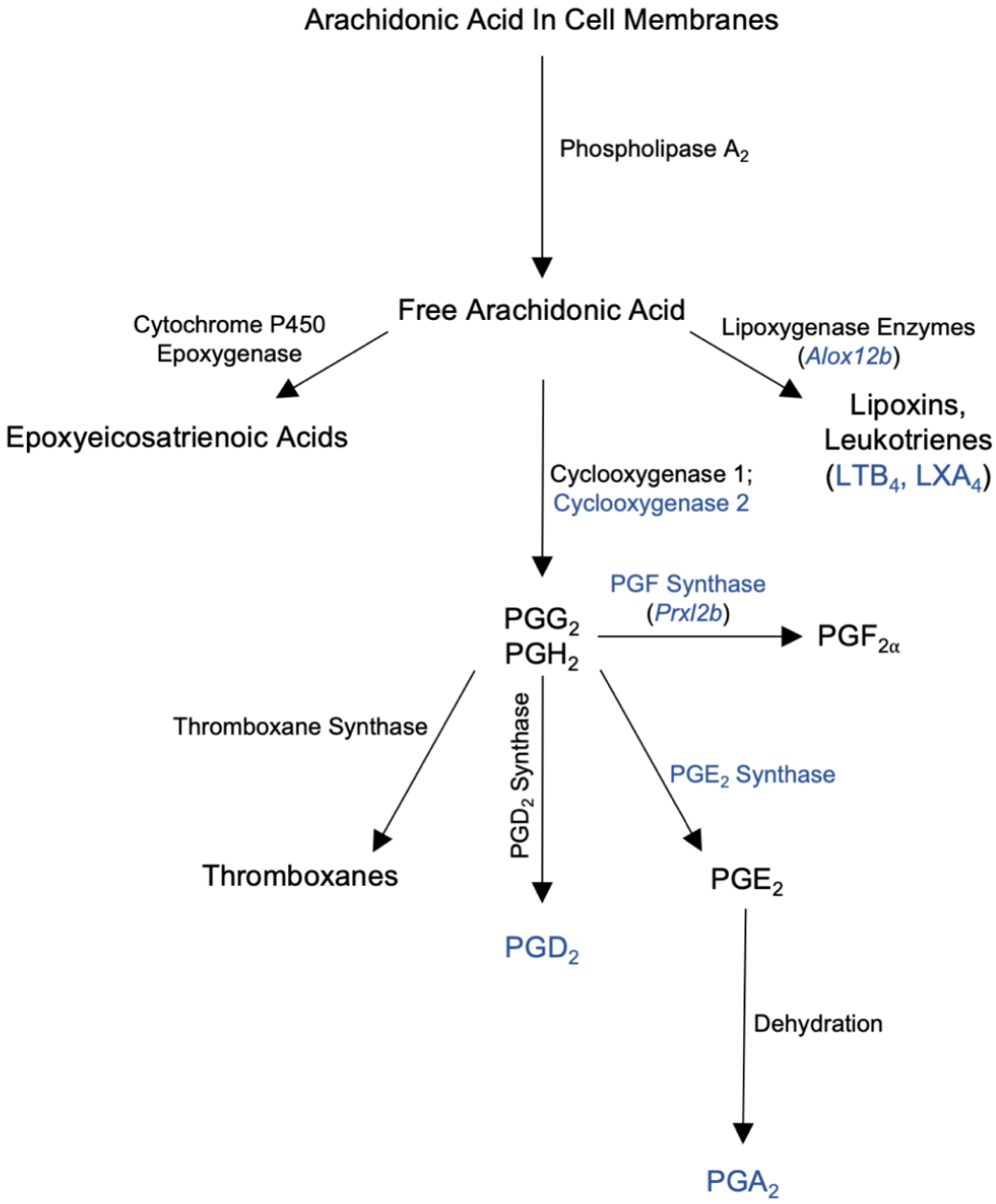


Figure 2.4 *Lm* infection alters the placental eicosanome. Eicosanoid profiles for infected and uninfected placentas were assessed using semi-targeted mass spectrometry. A heatmap was generated using Metaboanalyst to compare relative eicosanoid concentrations in infected versus uninfected placental samples (A). Fold change was analyzed using Metaboanalyst and is expressed as a dot plot with each dot representing the Log2 fold change (infected/uninfected) of each compound in our eicosanoid panel (B). Eicosanoids with >2-fold change are represented by pink dots, and significantly overexpressed eicosanoids are labeled.

A.

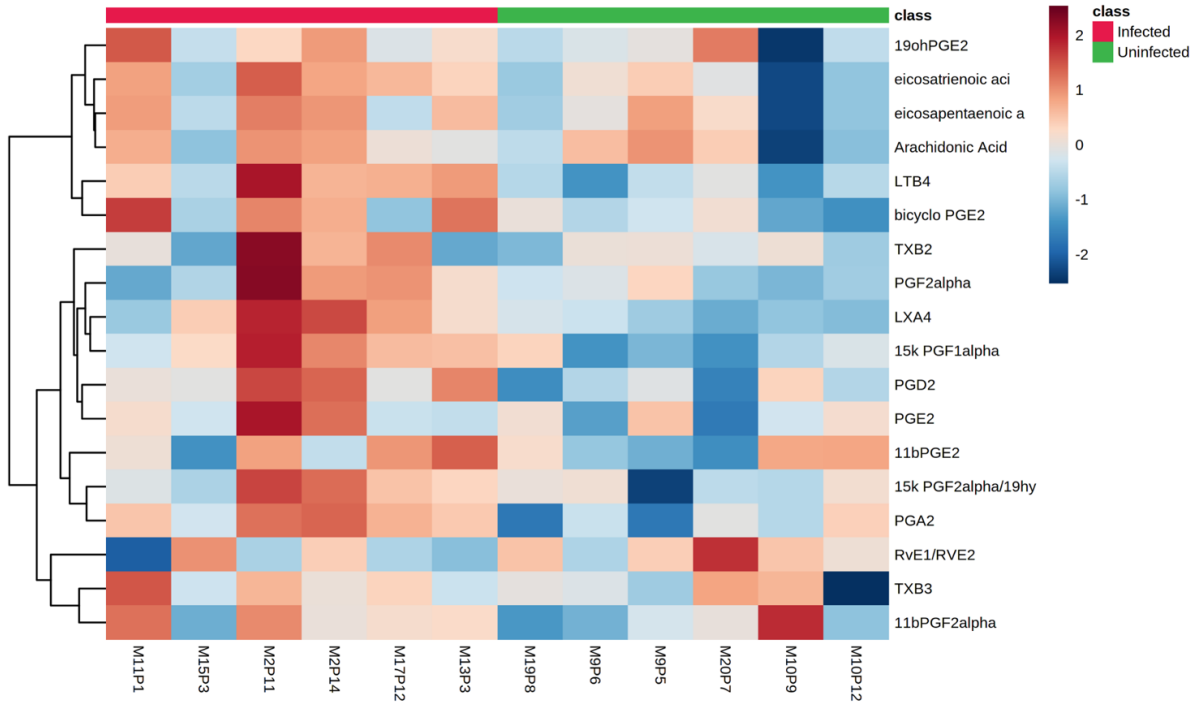


Figure 2.4 (Cont'd)

B.

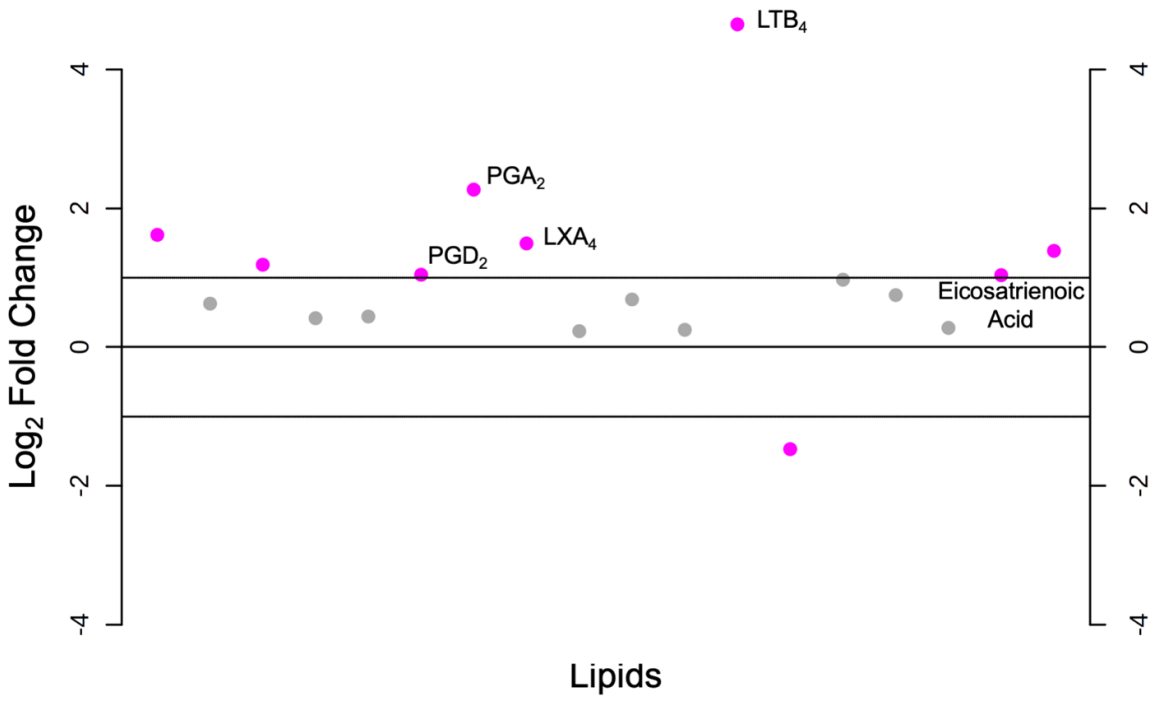
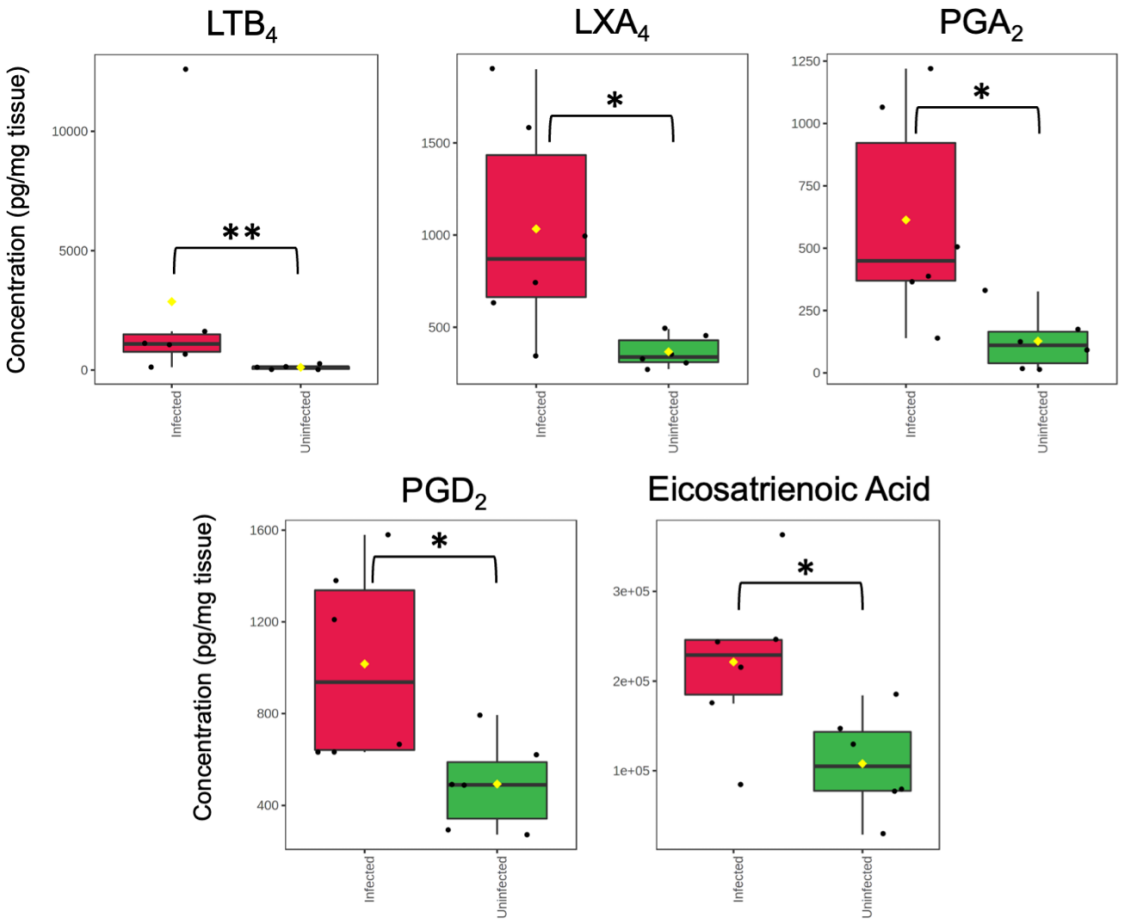


Figure 2.5 Several eicosanoids are significantly overexpressed following placental infection. Eicosanoid profiles for infected and uninfected placentas were assessed using semi-targeted mass spectrometry. Data was analyzed using Metaboanalyst, and eicosanoids which reached statistical significance (*p < 0.05, **p < 0.01) are represented as box plots below. Infected samples are in red while uninfected samples are in green. Each dot represents one sample. Concentrations are expressed as pg/mg of tissue.



REFERENCES

REFERENCES

[1] S.E. Ander, M.S. Diamond, C.B. Coyne, Immune responses at the maternal-fetal interface, Sci. Immunol. 4 (2019). https://doi.org/10.1126/sciimmunol.aat6114.

[2] V.B. Zeldovich, A.I. Bakardjiev, Host defense and tolerance: unique challenges in the placenta, PLoS Pathog. 8 (2012) e1002804. https://doi.org/10.1371/journal.ppat.1002804.

[3] J.A. Vazquez-Boland, M. Kuhn, P. Berche, T. Chakraborty, G. Dominguez-Bernal, W. Goebel, B. Gonzalez-Zorn, J. Wehland, J. Kreft, *Listeria* Pathogenesis and Molecular Virulence Determinants, Clinical Microbiology Reviews. 14 (2001) 584–640. https://doi.org/10.1128/CMR.14.3.584-640.2001.

[4] T. Mateus, J. Silva, R.L. Maia, P. Teixeira, Listeriosis during Pregnancy: A Public Health Concern, ISRN Obstet Gynecol. 2013 (2013) 851712. https://doi.org/10.1155/2013/851712.

[5] *Listeria* (Listeriosis), Centers for Disease Control and Prevention, n.d. https://www.cdc.gov/listeria/index.html (accessed March 6, 2022).

[6] *Listeria* During Pregnancy, American Pregnancy Association, n.d. https://americanpregnancy.org/healthy-pregnancy/pregnancy-concerns/listeria-during-pregnancy/.

[7] D.M. Olson, C. Amman, Role of the prostaglandins in labour and prostaglandin receptor inhibitors in the prevention of preterm labour, Frontiers in Bioscience. 12 (2007) 1329–1343.

[8] R. Bakker, S. Pierce, D. Myers, The role of prostaglandins E1 and E2, dinoprostone, and misoprostol in cervical ripening and the induction of labor: a mechanistic approach, Arch Gynecol Obstet. 296 (2017) 167–179. https://doi.org/10.1007/s00404-017-4418-5.

[9] J.R.G. Challis, D.M. Sloboda, N. Alfaidy, S.J. Lye, W. Gibb, F.A. Patel, W.L. Whittle, J.P. Newham, Prostaglandins and mechanisms of preterm birth, Reproduction. 124 (2002) 1470–1626.

[10] E. Ricciotti, G.A. FitzGerald, Prostaglandins and Inflammation, Arterioscler Thromb Vasc Biol. 31 (2011) 986–1000. https://doi.org/10.1161/ATVBAHA.110.207449.

[11] J.R. Robbins, A.I. Bakardjiev, Pathogens and the placental fortress, Curr. Opin. Microbiol. 15 (2012) 36–43. https://doi.org/10.1016/j.mib.2011.11.006.

[12] E. Afgan, D. Baker, B. Batut, M. van den Beek, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, B.A. Grüning, A. Guerler, J. Hillman-Jackson, S. Hiltemann, V. Jalili, H. Rasche, N. Soranzo, J. Goecks, J. Taylor, A. Nekrutenko, D. Blankenberg, The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update, Nucleic Acids Research. 46 (2018) W537–W544. https://doi.org/10.1093/nar/gky379.

[13] S. Andrews, FastQC: A Quality Control Tool for High Throughput Sequence Data, 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

[14] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics. 30 (2014) 2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

[15] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat Methods. 9 (2012) 357–359. https://doi.org/10.1038/nmeth.1923.

[16] Y. Liao, G.K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features, Bioinformatics. 30 (2014) 923–930. https://doi.org/10.1093/bioinformatics/btt656.

[17] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 550. https://doi.org/10.1186/s13059-014-0550-8.

[18] J. Reimand, M. Kull, H. Peterson, J. Hansen, J. Vilo, g:Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments, Nucleic Acids Research. 35 (2007) W193–W200. https://doi.org/10.1093/nar/gkm226.

[19] M. Pomaznoy, B. Ha, B. Peters, GOnet: a tool for interactive Gene Ontology analysis, BMC Bioinformatics. 19 (2018) 470. https://doi.org/10.1186/s12859-018-2533-3.

[20] B. Xi, H. Gu, H. Baniasadi, D. Raftery, Statistical Analysis and Modeling of Mass Spectrometry-Based Metabolomics Data, in: D. Raftery (Ed.), Mass Spectrometry in Metabolomics, Springer New York, New York, NY, 2014: pp. 333–353. https://doi.org/10.1007/978-1-4939-1258-2_22.

[21] J. Chong, D.S. Wishart, J. Xia, Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis, Current Protocols in Bioinformatics. 68 (2019). https://doi.org/10.1002/cpbi.86.

[22] J. Hardy, B. Kirkendoll, H. Zhao, L. Pisani, R. Luong, A. Switzer, M.V. McConnell, C.H. Contag, Infection of pregnant mice with *Listeria monocytogenes* induces fetal bradycardia, Pediatr Res. 71 (2012) 539–545. https://doi.org/10.1038/pr.2012.2.

[23] P. Subramanian, R.V. Stahelin, Z. Szulc, A. Bielawska, W. Cho, C.E. Chalfant, Ceramide 1-phosphate acts as a positive allosteric activator of group IVA cytosolic phospholipase A2 alpha and enhances the interaction of the enzyme with phosphatidylcholine, J Biol Chem. 280 (2005) 17601–17607. https://doi.org/10.1074/jbc.M414173200.

[24] S. Beck, D. Wojdyla, L. Say, A.P. Betran, M. Merialdi, J.H. Requejo, C. Rubens, R. Menon, P.F.A. Van Look, The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity, Bull. World Health Organ. 88 (2010) 31–38. https://doi.org/10.2471/BLT.08.062554.

[25] R.L. Goldenberg, J.F. Culhane, J.D. Iams, R. Romero, Epidemiology and causes of preterm birth, Lancet. 371 (2008) 75–84. https://doi.org/10.1016/S0140-6736(08)60074-4.

[26] J.P. Vogel, S. Chawanpaiboon, A.-B. Moller, K. Watananirun, M. Bonet, P. Lumbiganon, The global epidemiology of preterm birth, Best Practice & Research Clinical Obstetrics & Gynaecology. 52 (2018) 3–12. https://doi.org/10.1016/j.bpobgyn.2018.04.003.

[27] G.L. Mendz, N.O. Kaakoush, J.A. Quinlivan, Bacterial aetiological agents of intra-amniotic infections and preterm birth in pregnant women, Front. Cell. Infect. Microbiol. 3 (2013). https://doi.org/10.3389/fcimb.2013.00058.

[28] M.B. Vigliani, A.I. Bakardjiev, Intracellular Organisms as Placental Invaders, Fetal Matern Med Rev. 25 (2014) 332–338. https://doi.org/10.1017/S0965539515000066.

[29] D.E. Lowe, J.R. Robbins, A.I. Bakardjiev, Animal and Human Tissue Models of Vertical *Listeria monocytogenes* Transmission and Implications for Other Pregnancy-Associated Infections, Infect. Immun. 86 (2018). https://doi.org/10.1128/IAI.00801-17.

[30] X. Feng, Y. Zhang, Y. Zhang, X. Yang, D. Man, L. Lu, T. Xu, Y. Liu, C. Yang, H. Li, L. Qi, H. Su, X. Zhou, Z. Xu, Prostaglandin I2 mediates weak vasodilatation in human placental microvessels, Biol Reprod. 103 (2020) 1229–1237. https://doi.org/10.1093/biolre/ioaa156.

[31] M. Szczuko, J. Kikut, N. Komorniak, J. Bilicki, Z. Celewicz, M. Ziętek, The Role of Arachidonic and Linoleic Acid Derivatives in Pathological Pregnancies and the Human Reproduction Process, IJMS. 21 (2020) 9628. https://doi.org/10.3390/ijms21249628.

[32] A. López Bernal, D.J. Hansell, T.Y. Khong, J.W. Keeling, A.C. Turnbull, Placental leukotriene B4 release in early pregnancy and in term and preterm labour, Early Human Development. 23 (1990) 93–99. https://doi.org/10.1016/0378-3782(90)90132-3.

[33] L.O. Perucci, P.C. Santos, L.S. Ribeiro, D.G. Souza, K.B. Gomes, L.M.S. Dusse, L.P. Sousa, Lipoxin A4 Is Increased in the Plasma of Preeclamptic Women, AJHYPE. 29 (2016) 1179–1185. https://doi.org/10.1093/ajh/hpw053.

[34] S. Kumar, T. Palaia, C.E. Hall, L. Ragolia, Role of Lipocalin-type prostaglandin D2 synthase (L-PGDS) and its metabolite, prostaglandin D2, in preterm birth, Prostaglandins & Other Lipid Mediators. 118–119 (2015) 28–33. https://doi.org/10.1016/j.prostaglandins.2015.04.009.

[35] S. Babicki, D. Arndt, A. Marcu, Y. Liang, J.R. Grant, A. Maciejewski, D.S. Wishart, Heatmapper: web-enabled heat mapping for all, Nucleic Acids Res. 44 (2016) W147–W153. https://doi.org/10.1093/nar/gkw419.

CHAPTER 3

NOVEL INTERNALIN P HOMOLOGS IN *LISTERIA*

PUBLICATION NOTICE

The following chapter has been accepted for publication by *Microbial Genomics* and is currently available as a preprint through bioRxiv: "*K.N. Conner, J.T. Burke, J. Ravi, J.W. Hardy, Novel Internalin P homologs in Listeria, bioRxiv. https://doi.org/10.1101/2022.01.19.476994*."

ABSTRACT

Listeria monocytogenes (*Lm*) is a bacterial pathogen that causes listeriosis in immunocompromised individuals, particularly pregnant women. Several virulence factors support the intracellular lifecycle of *Lm* and facilitate cell-to-cell spread, allowing it to occupy multiple niches within the host and cross protective barriers, including the placenta. One family of virulence factors, internalins, contributes to *Lm* pathogenicity by inducing specific uptake and conferring tissue tropism. Over 25 internalins have been identified thus far, but only a few have been extensively studied. Internalins contain leucine-rich repeat (LRR) domains which enable protein-protein interactions, allowing *Lm* to bind host proteins. Notably, other *Listeria* species express internalins but cannot colonize human hosts, prompting questions regarding the evolution of internalins within the genus *Listeria*. Internalin P (InlP) promotes placental colonization through interaction with the host protein afadin. Though prior studies of InlP have begun to elucidate its role in *Lm* pathogenesis, there remains a lack of information regarding homologs in other *Listeria* species. Here, we have used a computational evolutionary approach to identify InlP homologs in additional *Listeria* species. We found that *L. ivanovii londoniensis* (*Liv*) and *L. seeligeri* (*Ls*) encode InlP homologs. We also found InlP-like homologs in *L. innocua* and the recently identified species *L. costaricensis.* All newly identified homologs lack the full-length LRR6 and LRR7 domains found in *Lm*'s InlP. These findings inform on the evolution of one key *Lm* virulence factor, InlP, and serve as a springboard for future evolutionary studies of *Lm* pathogenesis as well as mechanistic studies of *Listeria* internalins.

INTRODUCTION

Prenatal infection remains a major public health concern. Annually, nearly 13 million infants are born prematurely worldwide, and an estimated 30% of these preterm births can be attributed to prenatal infection, though the actual number may be higher due to the subclinical nature of many prenatal infections [1]. To better detect and treat these infections to prevent adverse pregnancy outcomes, we must better understand the pathogens that cause them. *Listeria monocytogenes* is widely used in prenatal infection research due to its well-characterized lifecycle and ease of use in laboratory experiments [2,3].

The *Listeria* genus comprises 17 species, including the human pathogens *L. ivanovii* (*Liv*) and *L. monocytogenes* (*Lm*) [4]. These Gram-positive facultative intracellular bacterial pathogens are the most typical causative agent of listeriosis in humans [4,5]. While relatively rare, listeriosis can result in severe morbidity and mortality in immunocompromised individuals [5,6]. Pregnant people are particularly at risk for listeriosis, as *Lm* can colonize the placenta and cause adverse pregnancy outcomes such as preterm birth, neonatal meningitis, miscarriage, and stillbirth [5]. They are approximately ten times more likely to contract listeriosis compared to their immunocompetent counterparts, comprising 17% of all annual cases of listeriosis [7,8]. *Lm* employs several virulence factors that aid in its invasion of various host niches and the breach of protective host barriers [5,6]. Previous studies have addressed the roles of various *Lm* virulence factors, such as ActA, Internalin A (InlA), and Internalin B (InlB) in the context of pregnancy [9–11]. Faralla *et al*. followed up with two studies focusing on Internalin P (InlP), a key virulence factor for the invasion of the placenta [12,13].

Typically, listeriosis begins with the consumption of contaminated food items. Once in the digestive system, *Lm* uses several virulence factors, including the internalins, to colonize gut epithelial cells and spread throughout the host [5,6,14]. Internalins contribute to this spread by conferring tissue tropism; for example, InlA binds E-cadherin on gut epithelial cells while InlB binds C-Met expressed by hepatocytes [14]. These interactions are enabled by Leucine-Rich Repeat (LRR) domains found in all internalins [14,15]. LRR domains are found in an array of functionally diverse proteins across the domains of life. LRRs are found in ribonuclease inhibitors in humans and pigs, connectin in *Drosophila*, adenylate cyclase in *Saccharomyces,* transmembrane kinase I in *A. thaliana*, and various virulence factors in pathogens such as *Y. pestis* and *L. monocytogenes* [15,16]. Internalins may have as few as four (InlG) or as many as fourteen (InlA) LRR domains [5,14]. Notably, internalins and other virulence factors are relatively well-conserved across the genus, including species that are considered non-pathogenic to humans, such as *L. seeligeri* and *L. innocua* [4,17,18]. While the details remain unclear, differences in pathogenicity have been attributed to minor genetic variations and differences in the expression of virulence factor genes [19]. The precise roles of the various internalin genes in *Listeria* and their evolutionary relationships remain critical open questions in *Listeria* biology.

Internalin P (InlP) is an *Lm* virulence factor known to enhance placental colonization in the pregnant host. This is likely accomplished by enabling *Lm* to transcytose through the basal membrane underlying the syncytiotrophoblast, the protective outer layer of placental cells that serves as a barrier between maternal and fetal blood [12]. Further characterization revealed that InlP encompasses nine LRR domains and binds the human protein afadin, which is a nectin-like protein found in cell-cell junctions and thought to play a significant role in cellular adhesion [13].

Initial InlP studies identified a structural homolog of InlP, Lmo2027, in *Lm*, but information regarding InlP homologs in other *Listeria* species has been incomplete [12]. In this study, we used comparative genomics and protein sequence-structure-function analyses to identify InlP homologs in the genomes of *L. seeligeri* (*Ls*), *L. ivanovii londoniensis* (*Liv*), *L. innocua* (*Lin*), and *L. costaricensis* (*Lc*). The bioinformatic analysis presented here serves as a springboard for future studies of *Listeria* evolution and pathogenesis pertaining to the internalin protein family, including its ability to colonize the human placenta.

MATERIALS AND METHODS

Identification of InlP Homologs

To identify $InlP_{Lm}$ homologs across evolutionary lineages, we submitted the $InlP_{Lm}$ amino acid sequence (accession: WP_014601135.1) to MolEvolvR (http://www.jravilab.org/molevolvr) [20]. The query returned hits for homologous proteins across bacterial phyla. While many species carried homologous proteins (*e.g.*, *Nostoc spp.* and *Beggiatoa leptomitoformis*), we chose to filter out hits with low similarity and divergent domain architectures and genomic contexts for our detailed study; we thus focused on the *Listeria* genus (including 45,530 *L. monocytogenes*, 740 *L. innocua*, 169 *L. seeligeri*, 44 *L. ivanovii*, and 1 *L. costaricensis* genomes). Within this dataset, we selected the hits with the highest percent similarity and unique domain architectures as representative homologs for further analysis. Accession numbers provided by MolEvolvR were used to query the NCBI RefSeq Protein Database for corresponding nucleotide sequences, locus tags, and isolate source (where available) for homologous genes [21]. BioCyc [22]

(https://biocyc.org) and NCBI RefSeq [21] protein databases were used to identify genomic contexts (neighboring genes).

Calculation of Percent Identity and Percent Similarity

Percent identity and percent similarity values for predicted homologs were provided by MolEvolvR [20] (**Table S1**; https://github.com/jravilab/inlp_listeria). Nucleotide sequences for *inlP* in *L. monocytogenes*, *L. ivanovii londoniensis*, *L. seeligeri,* and *L. costaricensis* were aligned using Clustal Omega [23] (https://www.ebi.ac.uk/Tools/msa/clustalo/), and resulting alignments were submitted in FASTA format to the Sequence Manipulation Suite [24] (SMS; http://www.bioinformatics.org/sms2) to calculate percent nucleotide identity. The homolog similarity and identity matrix was generated using MatGAT2.01 with the BLOSUM 62 matrix and default options [25].

Multiple Sequence Alignment, Phylogenetic Trees, and Protein Models

Multiple sequence alignments for homologous amino acid and nucleotide sequences were generated using Kalign [26] and visualized using JalView (Version 2.11.1.4) [27] with default parameters. Multiple sequence alignments in Figures 1 and 2 were generated using the msaplot function in the ggtree R package with default parameters [28]. Neighbor-joining trees were constructed using the ape package in R [29]. Domain architectures were determined using Interproscan with default parameters. Three-dimensional protein models were produced using SWISS-MODEL with the *L. monocytogenes* Internalin P crystal structure (PDB: 5hl3) as a template and visualized using ChimeraX [30–32]. All data, analyses, visualizations, and Table S1 for InlP *Listeria* homologs are available here: https://github.com/jravilab/inlp_listeria.

RESULTS

Listeria ivanovii londoniensis and *Listeria seeligeri* encode Internalin P Homologs

A To begin investigating evolutionary conservation of the *L. monocytogenes* Internalin P ($InlP_{Lm}$), we started with an extensive homology search and protein characterization of InlP-like proteins in diverse lineages across the tree of life using MolEvolvR [20] (http://jravilab.org/molevolvr). Most homologs were present only within the genus *Listeria*. To further ensure that all homologs are being identified, we picked other representative InlP homologs from *L. ivanovii* and *L. seeligeri* as new starting points for our homology search and characterization (using MolEvolvR [20]; see *Methods*). We found several hits in our multi-start search including proteins that contain transmembrane domains, resembling InlB rather than InlP (**Fig. 3.2**). Other hits carried neither the signature LRR (**Fig. 3.2**) or Internalin_N (**Fig. 3.6**) domains characteristic of internalins. Therefore, we restricted our full set of homologs to only InlP-like proteins resulting in 64 representative proteins with distinct domain architectures from each *Listeria* species including *L. monocytogenes*, *L. seeligeri*, *L. ivanovii*, *L. innocua*, and *L. costaricensis* (**Fig. 3.2, 3.6; Table S1**). Homologs from *L. seeligeri* (*Ls*) and *L. ivanovii* (*Liv*) showed >65% amino acid similarity compared to $InlP_{Lm}$, while homologs from *L. innocua* (*Lin*) and *L. costaricensis* (*Lc*) showed 52.6% and 53% similarity, respectively (**Fig. 3.3**). We found that several homologs lacked predicted signal peptide domains suggesting they are not secreted like InlP; this was corroborated by the presence of predicted transmembrane LPXTG motifs, which indicate that these homologs are more likely to be membrane-anchored InlB-like proteins rather than secreted InlP homologs (**Fig. 3.2, 3.6**). Investigation of the InlP-like proteins in *Lc and Lin* showed that it is unlikely that they are functional InlP homologs (discussed below). Further

investigation of domain architectures and genomic contexts of putative homologs ultimately revealed one InlP homolog encoded within the *L. ivanovii londoniensis* genome and three InlP paralogs encoded by *L. seeligeri* (discussed below). Here, we refer to these homologs as InlP$_{Lm}$, InlP$_{Li}$, InlP$_{Ls1}$, InlP$_{Ls2}$, and InlP$_{Ls3}$, respectively, to indicate species and gene order.

We determined the similarity of InlP homologs in *Liv* and *Ls* at the nucleotide and amino acid levels towards functional characterization. The *inlP* gene in *L. ivanovii londoniensis* (*inlP*$_{Liv}$), as well as the three *L. seeligeri* paralogs (*inlP*$_{Ls1}$, *inlP*$_{Ls2}$, and *inlP*$_{Ls3}$), shared ~70% identity with the *inlP*$_{Lm}$ gene and ~52–65% identity at the amino acid level when compared to InlP$_{Lm}$ (**Fig. 3.7**). However, the newly identified homologs in *L. ivanovii londoniensis* and *L. seeligeri* shared much higher percent amino acid similarity with InlP$_{Lm}$ — InlP$_{Liv,}$ InlP$_{Ls1,}$ and InlP$_{Ls3}$ showed ~70% similarity to InlP$_{Lm}$ (**Fig. 3.3**). Notably, the flanking *L. seeligeri* paralogs, InlP$_{Ls1}$ and InlP$_{Ls3}$, were more similar to each other than to the third paralog or to InlP$_{Lm}$ (**Fig. 3.3**). To further investigate these new *Listeria* InlP proteins, we next explored their genomic neighborhoods.

Because the *Listeria* genus maintains a high degree of synteny across species, we investigated the genomic contexts of identified homologs compared to the InlP$_{Lm}$ gene, which is flanked upstream by an amino acid permease gene and downstream by an NADPH dehydrogenase gene (**Fig. 3.4A**). We hypothesized that functional homologs of InlP$_{Lm}$ would be flanked by these same genes in other *Listeria* species. We, therefore, determined the genomic neighborhoods of *inlP* homologs in *L. ivanovii londoniensis, L. ivanovii ivanovii*, *L. seeligeri, L. costaricensis,* and *L. innocua* (*see Methods*). We found that the *inlP* genes in *L. ivanovii londoniensis* and *L. seeligeri* were flanked upstream by an amino acid permease gene and downstream by an NADPH dehydrogenase gene mirroring the *L. monocytogenes* genomic context, suggesting the identified homologs are likely true homologs of the *inlP* gene (**Fig. 3.4**). Interestingly, we found that the

gene encoding the InlP-like protein in *Lc* was flanked upstream by an amino acid permease and downstream by a gene encoding a LapB repeat-containing protein, inconsistent with genomic neighborhoods seen in other *Listeria* species (**Fig. 3.4E**). Also inconsistent with other genomic neighborhoods, the InlP-like protein in *Lin* was flanked upstream by a DUF5110-containing protein-encoding gene and downstream by the *ssrA* gene (**Fig. 3.4F**). To quantify the similarity of the flanking genes in *L. ivanovii* londoniensis, *L. seeligeri*, and *L. monocytogenes*, we calculated their pairwise similarity. At the amino acid level, the products of these flanking genes had >95% similarity to *L. monocytogenes*. Notably, while *L. ivanovii londoniensis* encoded an *inlP* homolog in this region, *L. ivanovii ivanovii* did not (**Fig. 3.4**). Additionally, while *L. ivanovii londoniensis* encoded only one copy of *inlP* (EL212_RS12905; $inlP_{Li}$), *L. seeligeri* encoded three copies (LSE_RS12040, LSE_RS12045, and LSE_RS12050; $inlP_{Ls1,}$ $inlP_{Ls2,}$ and $inlP_{Ls3,}$ respectively) (**Fig. 3.4**).

Internalin P Homologs in *L. ivanovii* and *L. seeligeri* lack the full-length LRR6 and LRR7 domains found in *L. monocytogenes* InlP

To delineate the evolution of InlP within *Listeria*, we generated a multiple sequence alignment (**Fig. 3.5**) and constructed a phylogenetic tree of the homologs (**Fig. 3.1**). While we observed several amino acid substitutions throughout the length of the proteins, the most striking difference between $InlP_{Lm}$ and its homologs was in the Leucine-Rich Repeat (LRR) regions — a partial lack of LRR6 and complete lack of LRR7 — in the *L. ivanovii* and *L. seeligeri* homologs (**Fig. 3.1**). Additionally, we noted a lack of conservation in a previously described calcium-binding loop present in $InlP_{Lm}$ (amino acid residues 132–135; **Fig. 3.1**) [10]. This observation was of particular interest since this calcium-binding loop might play a role in protein signaling or

stabilization of protein-protein interactions between InlP$_{Lm}$ and host afadin. To better visualize the structural differences in these homologs, we generated models of InlP$_{Li}$, InlP$_{Ls1}$, InlP$_{Ls2}$, and InlP$_{Ls3}$ based on the previously resolved crystal structure of InlP$_{Lm}$ (**Fig. 3.1**; see *Methods*). These models illustrate the similarity in the overall structure of the five homologous proteins, and the lack of LRR7 and full-length LRR6 are visible in *Li* and *Ls* homologs (**Fig. 3.1**; green/yellow regions). Additionally, the calcium-binding loop region is discernible in all five homologous proteins but appears structurally diverse in *L. ivanovii* and *L. seeligeri* homologs.

In summary, we have discovered novel Internalin P homologs in *Listeria*, traced their evolution, and uncovered potential functional implications pertaining to heterogeneity in key InlP domains. All InlP homolog data (along with characterizations in terms of domain architectures and modeling) are available at https://github.com/jravilab/inlp_listeria.

DISCUSSION

While previous studies have addressed many of the physical and mechanistic properties of InlP, the conservation of the *inlP* gene within or outside of the *Listeria* genus remains incompletely characterized. Here, we have provided insight into Internalin P in other *Listeria* species aside from *L. monocytogenes* that will drive future mechanistic studies of InlP as well as evolutionary studies of *Listeria* pathogenesis pertaining to the internalins.

First, we analyzed the InlP amino acid sequence with MolEvolvR [20] to retrieve homologous proteins across evolutionary lineages. MolEvolvR [20] is a powerful new bioinformatic web application to characterize protein families using molecular evolution and phylogeny (http://jravilab.org/molevolvr). The MolEvolvR [20] InlP search returned a list of potential homologs including those found in *Listeria* species. In this article, we focus on homologs

in *Listeria* since these species carried the classic Internalin and LRR domains. Using MolEvolvR [20], we identified InlP-like proteins in *L. innocua, L. seeligeri, L. ivanovii,* and *L. costaricensis*; only homologs in *L. seeligeri* and *L. ivanovii* londoniensis expressed an amino acid percent similarity value >65%. We found that the lower identity proteins are more likely to be Internalin B homologs based on their sequence, domain architecture, and structure.

To determine if the newly identified proteins were true homologs of $InlP_{Lm}$, we explored their domain architectures and genomic context. Consistent with the synteny observed in *Listeria* genomes, we found that the *inlP* domain architectures and genomic neighborhoods were highly conserved in *L. ivanovii londoniensis* and *L. seeligeri*, but not in *L. ivanovii ivanovii, L. innocua,* or *L. costaricensis.* While it is possible that these species could encode *inlP* homologs elsewhere in their genomes, it seems unlikely considering their domain architectures and lower conservation in sequence compared to other homologs. It is more likely that the homologous proteins identified in *Lc* and *Lin* are independent of InlP, but in the same class of small, secreted internalins that encompasses InlP, InlC, and InlH, among others [14].

Notably, we found that *L. ivanovii londoniensis* encoded a functional homolog for *inlP* while *L. ivanovii ivanovii* did not; *L. ivanovii ivanovii* encoded a pseudogene instead (**Fig. 3.8**). Historically, *L. ivanovii londoniensis* and *L. ivanovii ivanovii* have been distinguished biochemically [33]. Recently, Hupfeld *et al.,* noted that the two subspecies could also be distinguished based on bacteriophage susceptibility: *L. ivanovii ivanovii* strains are sensitive to bacteriophages, while *L. ivanovii londoniensis* strains encode a type II-A CRISPR-Cas system rendering them resistant to many phages [34]. Our finding that only *L. ivanovii londoniensis*, and not *L. ivanovii ivanovii*, encodes the *inlP* gene provides another avenue for distinguishing between these two subspecies and could be beneficial to public health laboratories seeking to differentiate

between them among clinical and food isolates. Additionally, because the evolution of virulence factors in *Listeria* remains mysterious, the specific presence of *inlP* in subspecies such as *ivanovii londoniensis* and its absence in *ivanovii ivanovii* may provide clues as to how *Listeria* evolves the ability to infect different cells and tissues.

Since *L. ivanovii* has been implicated in human and animal placental infection, it was not entirely surprising to find that it encoded the gene for InlP, an internalin known to enhance placental colonization. It was surprising, however, to find three copies of the *inlP* gene in *L. seeligeri* since it has not been significantly indicated in human or animal pathogenesis [35]. Our analyses suggested that InlP$_{Ls2}$ was the most similar to InlP$_{Lm}$ and InlP$_{Liv}$. It is possible that this paralog (InlP$_{Ls2}$) is the ancestral one, and InlP$_{Ls1}$ and InlP$_{Ls3}$ resulted from subsequent duplication events. The presence of multiple paralogs within *L. seeligeri* suggests that InlP could have alternative functions apart from enhancing placental colonization. *Listeria* species are frequently found in environmental isolates, as they readily reside in soil. It is possible that InlP provides a fitness advantage in this environment.

One of the main questions resulting from the discovery of InlP homologs centers on the evolutionary timeline of the *Listeria* genus: which InlP came first? Our discovery that InlP$_{Liv}$, InlP$_{Ls1}$, InlP$_{Ls2}$, and InlP$_{Ls3}$ do not contain the full-length LRR6 and LRR7 domains found in InlP$_{Lm}$ begins to offer potential answers to this question. It is plausible that *L. monocytogenes*, *L. seeligeri*, and *L. ivanovii londoniensis* shared a common ancestor that passed down the *inlP* gene, and a subsequent insertion event in *L. monocytogenes* led to the full-length InlP containing LRR6 and LRR7. Conversely, it is likely that *L. monocytogenes* carries the ancestral copy of InlP (InlP$_{Lm}$); the full-length *inlP* gene could have undergone a deletion resulting in the loss of LRR6 and LRR7 in InlP$_{Liv}$ and InlP$_{Ls}$, although it is less likely to observe several deletion events as against a single

insertion event. Future studies on the evolution of the *Listeria* genus and the larger family of internalin proteins will be required to answer this question more rigorously and determine their possible links to pathogenicity.

An additional structural difference noted between newly identified InlP homologs resides in the Ca^{2+}-binding loop of LRR3. Previously, this loop has been hypothesized to play a role in InlP signaling, activation, or stabilization in complex with its binding partner afadin [13]. Structural heterogeneity is visible in the Ca^{2+} regions of InlP homolog models; InlP homologs in *Liv* and *Ls* appear to have more open loops compared to $InlP_{Lm.}$ The ability of these loops to bind calcium, and their relative binding affinities will be an important avenue for future investigation, especially as more details regarding the function and regulation of $InlP_{Lm}$ come to light.

Recent studies made several fundamental discoveries regarding the physical and mechanistic properties of $InlP_{Lm}$ and its activity in the placenta, but many questions remain unanswered [12,13]. The discovery of InlP homologs in *L. ivanovii londoniensis* and *L. seeligeri*, two species that have not been substantially implicated in cases of placental infection, is compelling. Future studies will investigate the activity of these homologs to determine if they bind afadin and if they are able to enhance placental colonization of *L. monocytogenes* as well as endogenous InlP. Further, structural differences between these homologs suggest potential binding sites for the InlP-afadin interaction, which has not been resolved to date.

In summary, we report that *L. ivanovii londoniensis* and *L. seeligeri* encode homologs for the *L. monocytogenes* virulence factor InlP. Identified homologs in all three species are housed within similar genomic neighborhoods, flanked by the same housekeeping genes upstream and downstream; further, *L. seeligeri* encodes three copies of the *inlP* gene in this region. All four homologs are similar (>70%) to InlP in *L. monocytogenes*, the main structural difference resulting

from the lack of full-length LRR6 and LRR7 regions in InlP$_{Li}$, InlP$_{Ls1}$, InlP$_{Ls2}$, and InlP$_{Ls3}$. Our findings will serve as a springboard for future evolutionary studies of internalins in the *Listeria* genus and will bolster future *in vitro* and *in vivo* studies of InlP in the context of virulence and pathogenicity.

DECLARATIONS

Competing interests

The authors declare that they have no competing interests.

Funding

KNC is supported by Michigan State University (MSU) Microbiology and Molecular Genetics departmental fellowships. This work was supported by MSU start-up funds granted to JWH and JR.

Contributions

Conceptualization: KNC, JTB, JR, JWH. *Methodology:* KNC, JTB, JR. *Software:* JTB, JR. *Validation:* KNC, JTB, JR, JWH. *Formal Analysis:* KNC, JTB, JR. *Investigation:* KNC, JTB, JR. *Resources:* JR, JWH. *Data Curation:* KNC, JTB, JR. *Writing:* KNC, JR, JWH. *Visualization:* KNC, JTB, JR. *Supervision:* JR, JWH. *Project Administration:* JR, JWH. *Funding:* JR, JWH.

APPENDIX

Figure 3.1 Phylogeny and structure models of InlP and representative homologs. A phylogenetic tree was generated using the amino acid sequences of InlP homologs in *L. monocytogenes* (blue), *L. ivanovii londoniensis* (Purple), and *L. seeligeri* (green). Three-dimensional models were generated using SWISS-MODEL with the crystal structure for $InlP_{Lm}$ (PDB: 5hl3) as a template, then visualized using ChimeraX. Multiple sequence alignments were generated using MolEvolvR and illustrate the complete LRR6 and LRR7 insertion present in $InlP_{Lm}$ (inserted motif is highlighted in yellow in the 3D model in a backdrop of a blue protein structure model; also indicated with the arrow). The legend shows the colors of the amino acid residues indicated in the multiple sequence alignment. The height of the MSA (for each of the 5 sequences) has been increased to show the colors more distinctly, and to highlight the missing motif indicated by the arrow.

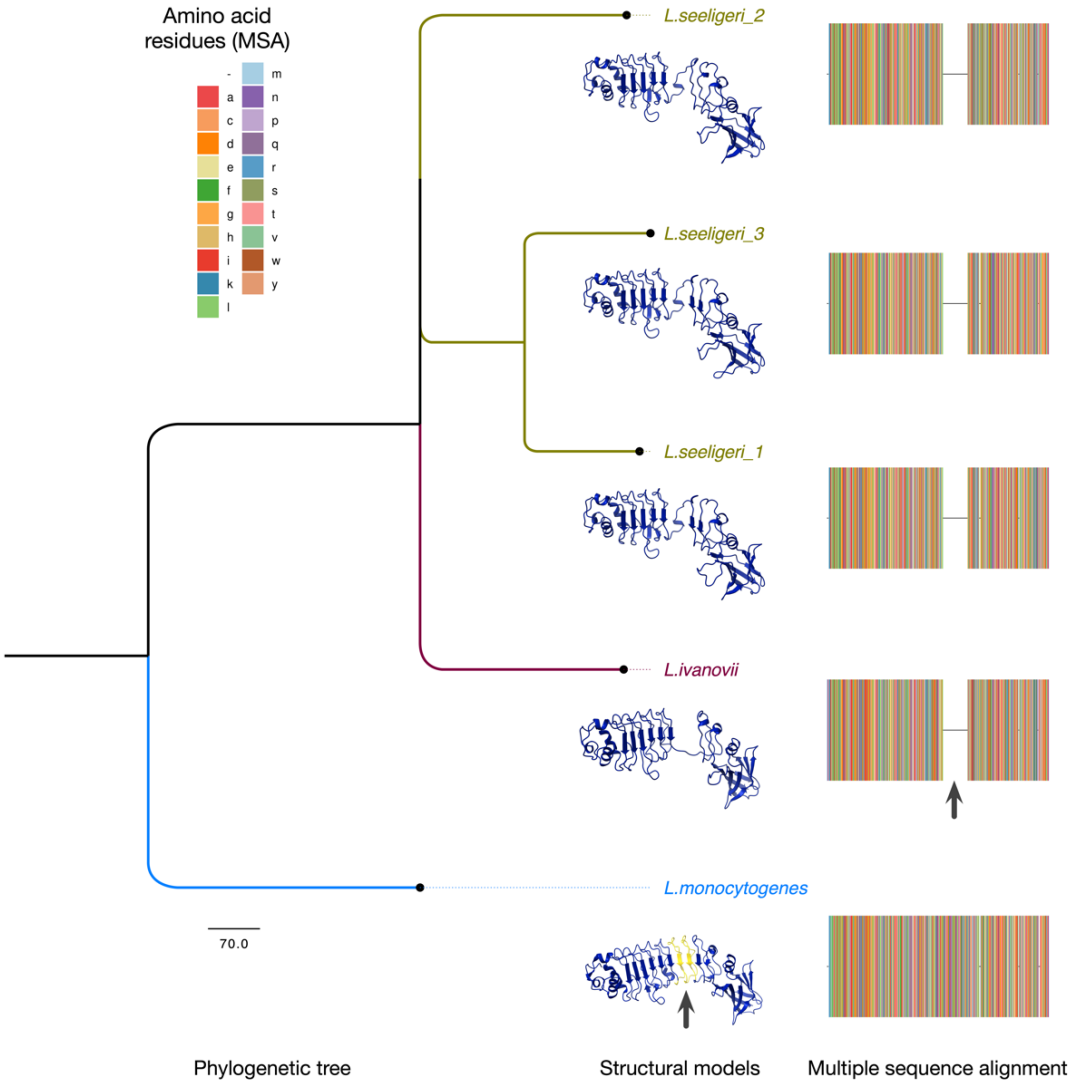


Figure 3.2 Phylogeny and domain architectures of putative Internalin P homologs. A multiple sequence alignment and phylogenetic tree were generated using the amino acid sequences of putative InlP homologs identified using MolEvolvR and five *Listeria* InlP starting points (See *Methods*). The phylogeny of InlP-like proteins (with branch lengths marked using dN/dS ratio with the ggtree package) has been overlaid with their domain architectures (generated using MolEvolvR, showing only Pfam domain architectures). The two legends show the colors of the amino acids indicated in the multiple sequence alignment (corresponding to the MSA, right panel) and the Pfam domain annotations (corresponding to the domain architecture, left panel). The arrow lengths and the overall colored MSA segments correspond to the relative lengths of each of the InlP-like proteins.

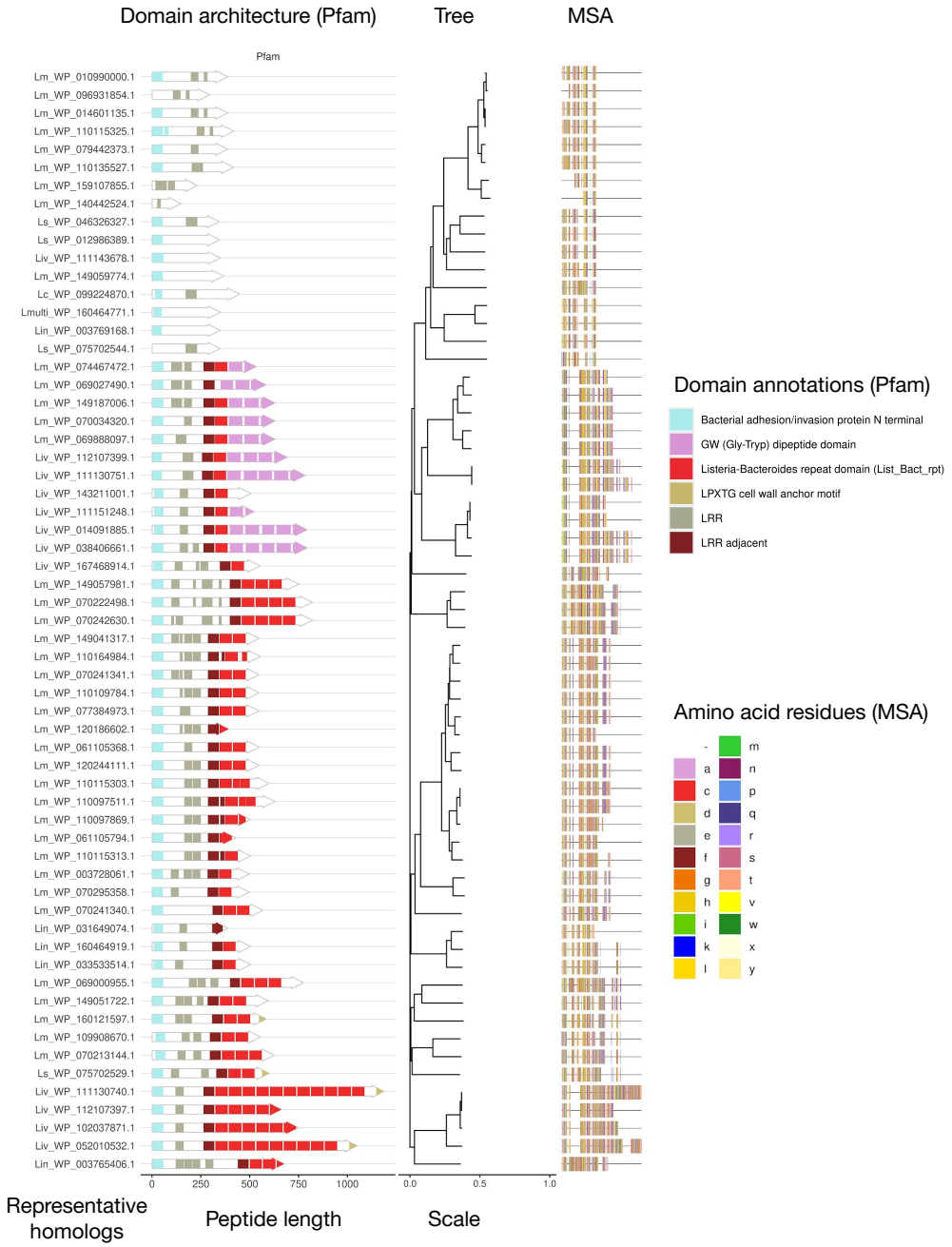


Figure 3.3 Percent similarity and identity of Internalin P homologs in representative *Listeria* species. Percent similarity and percent identity were calculated for the amino acid sequences for representative Internalin P homologs from five *Listeria* species, *L. monocytogenes, L. ivanovii, L. seeligeri, L. costaricensis,* and *L. innocua*. Matrix showing similarity and identity values was generated using MatGAT2.01 with the BLOSUM 62 matrix and default options selected.

Representative InlP homologs	*Liv*	*Lm*	*Ls1*	*Ls2*	*Ls3*	*Lin*	*Lc*
BFirmic_Livanovii_WP_126300113.1		58.1	74.6	78.6	75.2	37.8	32.1
BFirmic_Lmonocytogenes_WP_014601135.1	70.9		57.9	57.5	57.2	32.4	38.2
BFirmic_Lseeligeri_WP_012986389.1	85.7	72.7		77.3	90.5	36.9	33.1
BFirmic_Lseeligeri_WP_012986390.1	89.7	70.9	87.0		76.7	36.5	30.8
BFirmic_Lseeligeri_WP_012986391.1	84.9	72.2	94.8	85.3		36.6	32.5
BFirmic_Linnocua_WP_003769168.1	61.7	52.6	58.9	60.3	58.6		28.1
BFirmic_Lcostaricensis_WP_099224870.1	45.7	53.0	46.1	43.4	45.9	43.0	

Percentage identity

Percentage similarity

Figure 3.4 Genomic context of newly identified Internalin P gene homologs. Genes homologous to the *L. monocytogenes inlP* (A) were identified in various other *Listeria* species. Gene order was maintained in *L. monocytogenes*, *L. seeligeri*, and *L. ivanovii londoniensis*. Notably, *L. seeligeri* encodes three copies of the *inlP* gene. All homologs categorized as "true" homologs were flanked upstream by an amino acid permease gene (blue) and downstream by an NADPH dehydrogenase gene (orange). *L. ivanovii ivanovii* contains a pseudogene (purple) and an uncharacterized gene encoding a hypothetical protein (red) in this region. Putative homologous genes in *L. costaricensis* and *L. innocua* did not mirror genomic neighborhoods seen in the other *Listeria* species. All *inlP* homologs are represented in green. Genomic context was determined using RefSeq genomic records and the BioCyc genome browsers for each species (see *Methods*).

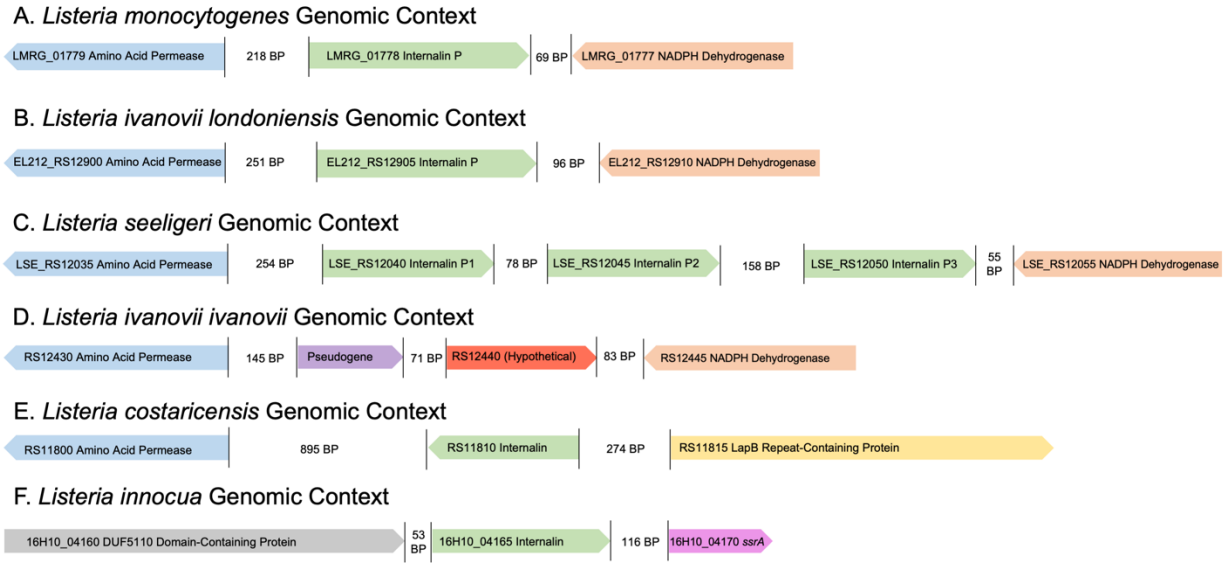


Figure 3.5 Multiple sequence alignment of Internalin P homologs. Amino acid sequences for identified Internalin P homologs in *L. monocytogenes*, *L. ivanovii londoniensis*, and *L. seeligeri* were aligned using Kalign and visualized using Jalview. Below the alignment is the consensus sequence for the four homologs. The red box indicates the insertion present only in the InlP$_{Lm}$ homolog. The *L. monocytogenes* homolog is used as the reference InlP protein (for residue numbering).

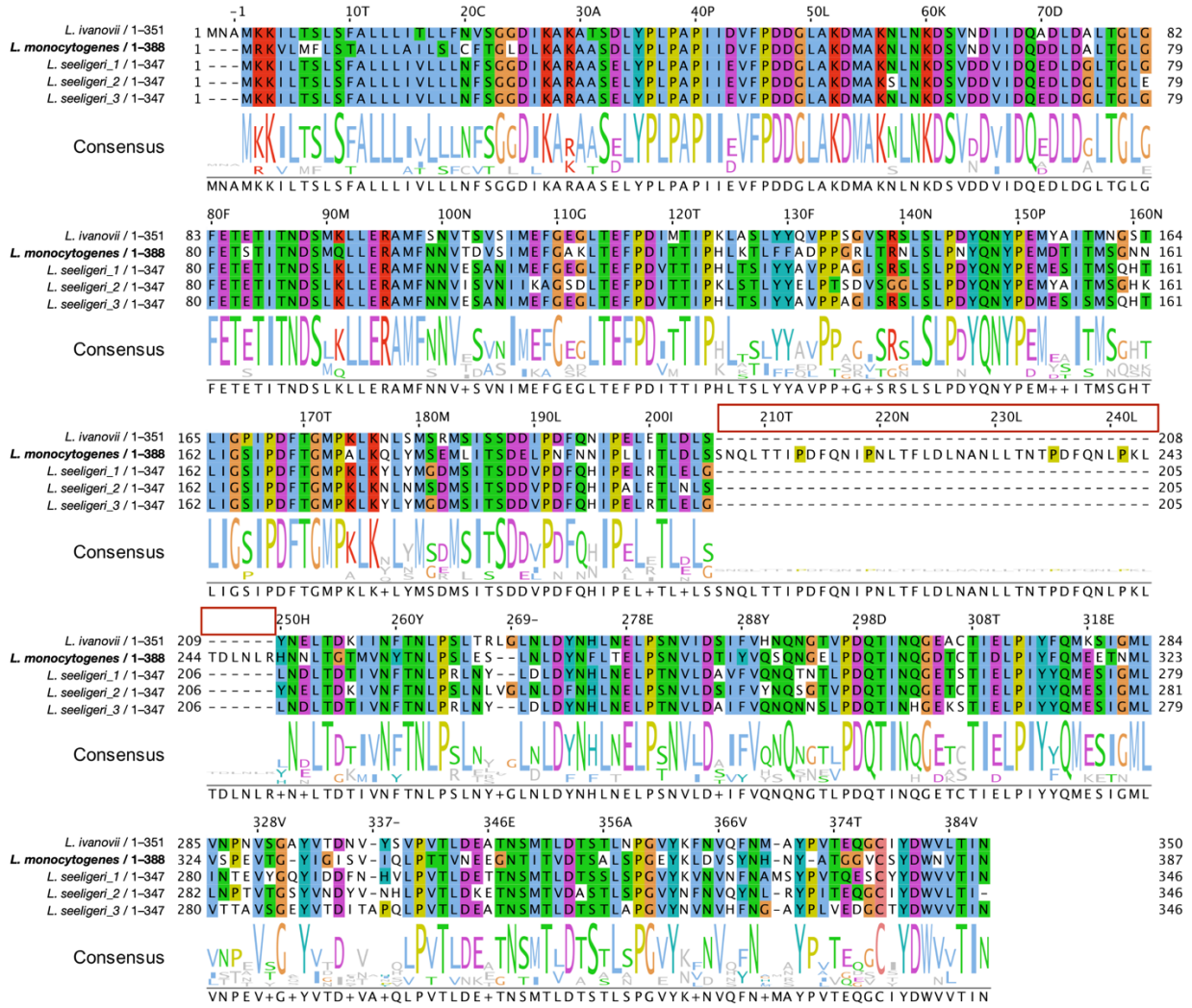


Figure 3.6 Phobius and Gene3D domain architectures of identified Internalin P homologs. Cellular localizations (Phobius) and domain architectures (Gene3D) were determined and visualized using MolEvolvR. Figure legends correspond to different domain and localization predictions.

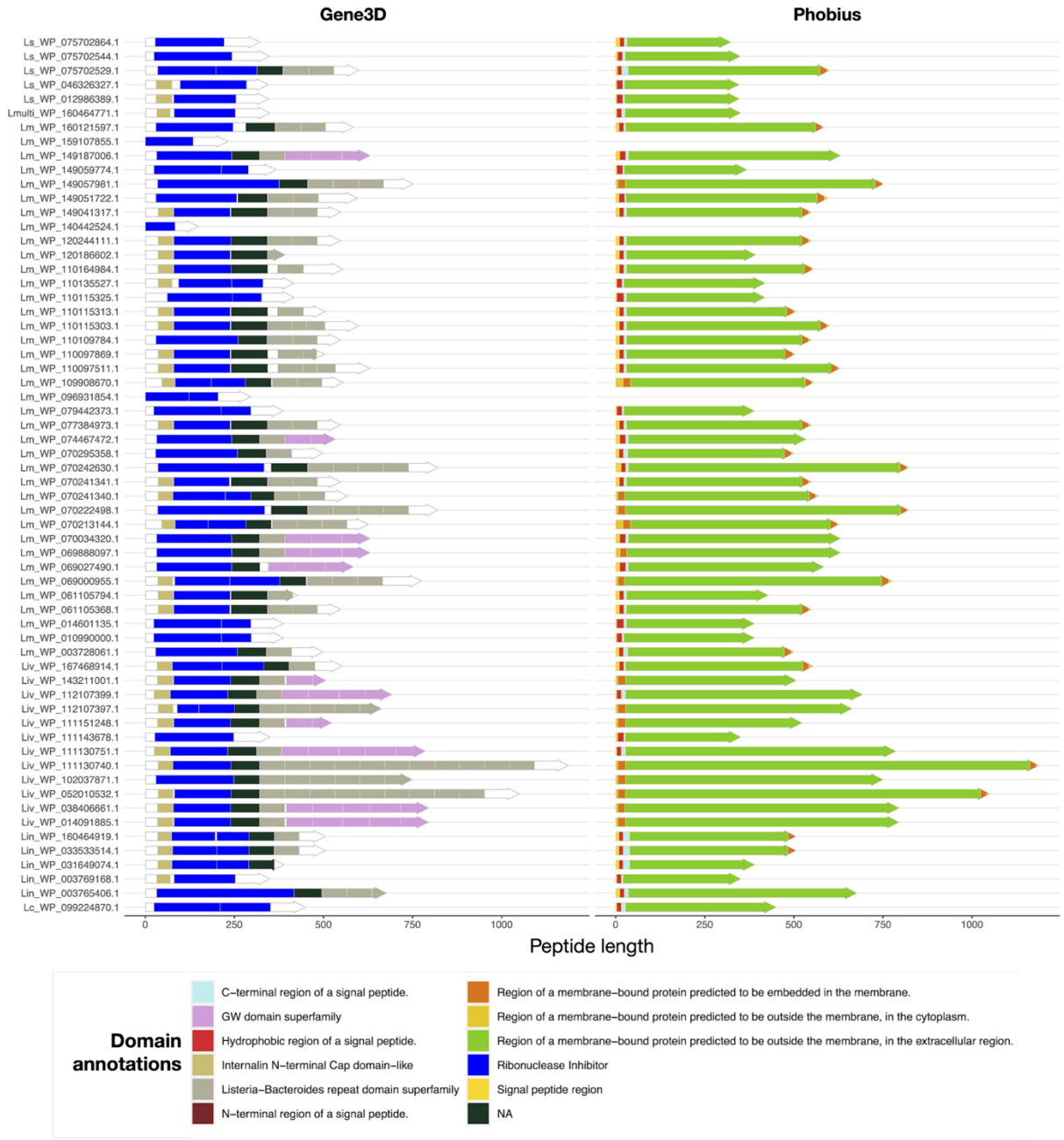


Figure 3.7 Percent nucleotide identity of *inlp* genes in *L. monocytogenes, L. ivanovii londoniensis*, and *L. seeligeri*. Heatmap representing pairwise percent identities of the nucleotide sequences of Internalin P homologs in *L. monocytogenes*, *L. ivanovii londoniensis*, and *L. seeligeri*. Darker and lighter shades of blue represent higher and lower percent identities, respectively.

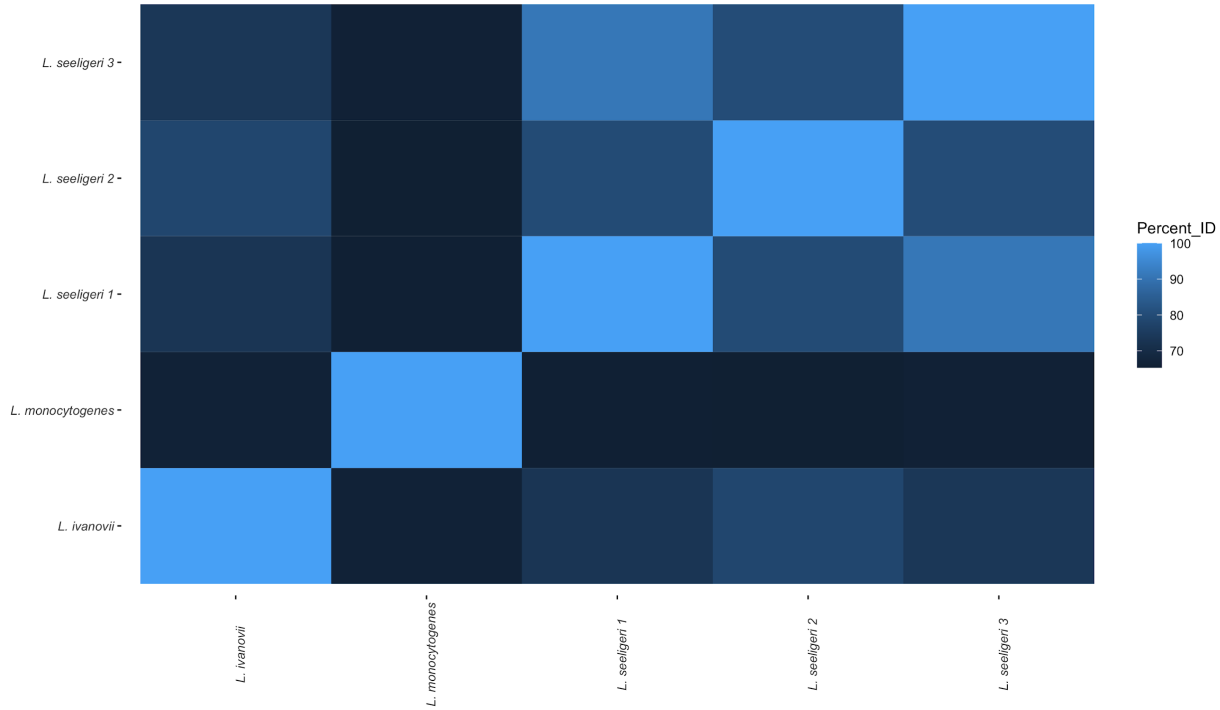
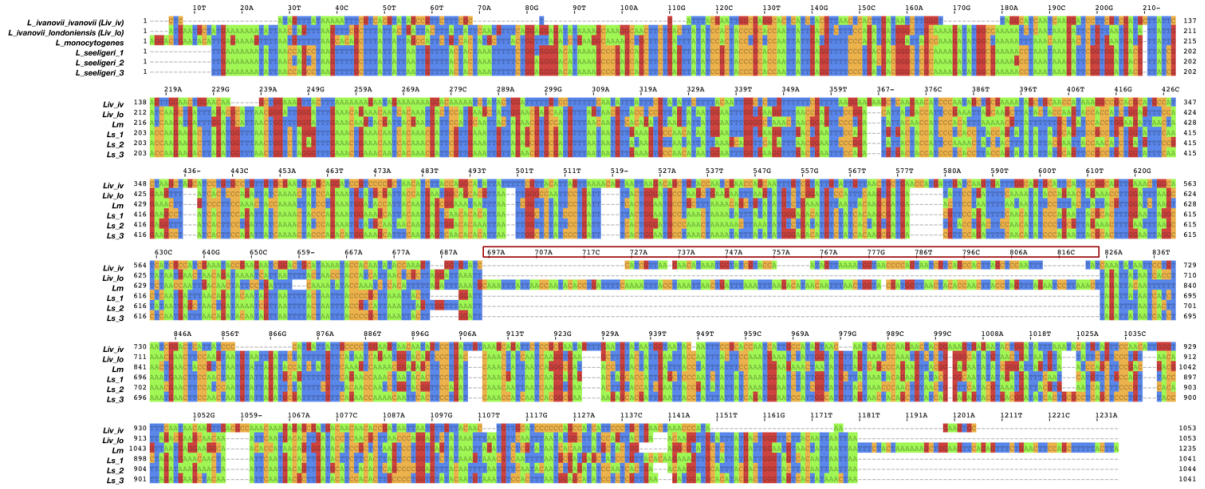


Figure 3.8 Nucleotide sequence alignment of *inlP* genes in *L. monocytogenes*, *L. ivanovii* subsp. *ivanovii*, *L. ivanovii* subsp. *londoniensis*, and *L. seeligeri*. Nucleotide sequences for *inlP* sequences in *L. monocytogenes*, *L. ivanovii* subsp. *londoniensis,* and *L. seeligeri* were aligned with the pseudogene region corresponding to *inlP* in *L. ivanovii* subsp. *ivanovii* using Clustal Omega and visualized using Jalview. The *L. monocytogenes inlP* gene is used as the reference for nucleotide numbering.



REFERENCES

REFERENCES

[1] R.L. Goldenberg, J.F. Culhane, J.D. Iams, R. Romero, Epidemiology and causes of preterm birth, Lancet. 371 (2008) 75–84. https://doi.org/10.1016/S0140-6736(08)60074-4.

[2] D.E. Lowe, J.R. Robbins, A.I. Bakardjiev, Animal and Human Tissue Models of Vertical *Listeria monocytogenes* Transmission and Implications for Other Pregnancy-Associated Infections, Infect. Immun. 86 (2018). https://doi.org/10.1128/IAI.00801-17.

[3] H.A. Morrison, D. Lowe, J.R. Robbins, A.I. Bakardjiev, In Vivo Virulence Characterization of Pregnancy-Associated *Listeria monocytogenes* Infections, Infect. Immun. 86 (2018). https://doi.org/10.1128/IAI.00397-18.

[4] R.H. Orsi, M. Wiedmann, Characteristics and distribution of *Listeria spp*., including *Listeria* species newly described since 2009, Appl Microbiol Biotechnol. 100 (2016) 5273–5287. https://doi.org/10.1007/s00253-016-7552-2.

[5] J.A. Vazquez-Boland, M. Kuhn, P. Berche, T. Chakraborty, G. Dominguez-Bernal, W. Goebel, B. Gonzalez-Zorn, J. Wehland, J. Kreft, *Listeria* Pathogenesis and Molecular Virulence Determinants, Clinical Microbiology Reviews. 14 (2001) 584–640. https://doi.org/10.1128/CMR.14.3.584-640.2001.

[6] L.T. Matereke, A.I. Okoh, *Listeria monocytogenes* Virulence, Antimicrobial Resistance and Environmental Persistence: A Review, Pathogens. 9 (2020) 528. https://doi.org/10.3390/pathogens9070528.

[7] *Listeria* (Listeriosis), Centers for Disease Control and Prevention, n.d. https://www.cdc.gov/listeria/index.html (accessed March 6, 2022).

[8] Listeria During Pregnancy, American Pregnancy Association, n.d. https://americanpregnancy.org/healthy-pregnancy/pregnancy-concerns/listeria-during-pregnancy/.

[9] M. Lecuit, D.M. Nelson, S.D. Smith, H. Khun, M. Huerre, M.-C. Vacher-Lavenu, J.I. Gordon, P. Cossart, Targeting and crossing of the human maternofetal barrier by *Listeria monocytogenes*: Role of internalin interaction with trophoblast E-cadherin, Proceedings of the National Academy of Sciences. 101 (2004) 6152–6157. https://doi.org/10.1073/pnas.0401434101.

[10] A. Le Monnier, O.F. Join-Lambert, F. Jaubert, P. Berche, S. Kayal, Invasion of the Placenta during Murine Listeriosis, Infection and Immunity. 74 (2006) 663–672. https://doi.org/10.1128/IAI.74.1.663-672.2006.

[11] A. Le Monnier, N. Autret, O.F. Join-Lambert, F. Jaubert, A. Charbit, P. Berche, S. Kayal, ActA Is Required for Crossing of the Fetoplacental Barrier by *Listeria monocytogenes*, Infection and Immunity. 75 (2007) 950–957. https://doi.org/10.1128/IAI.01570-06.

[12] C. Faralla, G.A. Rizzuto, D.E. Lowe, B. Kim, C. Cooke, L.R. Shiow, A.I. Bakardjiev, InlP, a New Virulence Factor with Strong Placental Tropism, Infect. Immun. 84 (2016) 3584–3596. https://doi.org/10.1128/IAI.00625-16.

[13] C. Faralla, E.E. Bastounis, F.E. Ortega, S.H. Light, G. Rizzuto, L. Gao, D.K. Marciano, S. Nocadello, W.F. Anderson, J.R. Robbins, J.A. Theriot, A.I. Bakardjiev, *Listeria monocytogenes* InlP interacts with afadin and facilitates basement membrane crossing, PLoS Pathog. 14 (2018) e1007094. https://doi.org/10.1371/journal.ppat.1007094.

[14] K. Ireton, R. Mortuza, G.C. Gyanwali, A. Gianfelice, M. Hussain, Role of internalin proteins in the pathogenesis of *Listeria monocytogenes*, Mol Microbiol. (2021) mmi.14836. https://doi.org/10.1111/mmi.14836.

[15] B. Kobe, The leucine-rich repeat as a protein recognition motif, Current Opinion in Structural Biology. 11 (2001) 725–732. https://doi.org/10.1016/S0959-440X(01)00266-4.

[16] B. Kobe, J. Deisenhofer, The leucine-rich repeat: a versatile binding motif, Trends in Biochemical Sciences. 19 (1994) 415–421. https://doi.org/10.1016/0968-0004(94)90090-6.

[17] T. Hain, S.S. Chatterjee, R. Ghai, C.T. Kuenne, A. Billion, C. Steinweg, E. Domann, U. Kärst, L. Jänsch, J. Wehland, W. Eisenreich, A. Bacher, B. Joseph, J. Schär, J. Kreft, J. Klumpp, M.J. Loessner, J. Dorscht, K. Neuhaus, T.M. Fuchs, S. Scherer, M. Doumith, C. Jacquet, P. Martin, P. Cossart, C. Rusniock, P. Glaser, C. Buchrieser, W. Goebel, T. Chakraborty, Pathogenomics of *Listeria spp*., International Journal of Medical Microbiology. 297 (2007) 541–557. https://doi.org/10.1016/j.ijmm.2007.03.016.

[18] P. Glaser, L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero, P. Berche, H. Bloecker, P. Brandt, T. Chakraborty, A. Charbit, F. Chetouani, E. Couvé, A. de Daruvar, P. Dehoux, E. Domann, G. Domínguez-Bernal, E. Duchaud, L. Durant, O. Dussurget, K.-D. Entian, H. Fsihi, F.G.-D. Portillo, P. Garrido, L. Gautier, W. Goebel, N. Gómez-López, T. Hain, J. Hauf, D. Jackson, L.-M. Jones, U. Kaerst, J. Kreft, M. Kuhn, F. Kunst, G. Kurapkat, E. Madueño, A. Maitournam, J.M. Vicente, E. Ng, H. Nedjari, G. Nordsiek, S. Novella, B. de Pablos, J.-C. Pérez-Diaz, R. Purcell, B. Remmel, M. Rose, T. Schlueter, N. Simoes, A. Tierrez, J.-A. Vázquez-Boland, H. Voss, J. Wehland, P. Cossart, Comparative Genomics of *Listeria* Species, Science. 294 (2001) 849–852. https://doi.org/10.1126/science.1063447.

[19] E. Gouin, J. Mengaud, P. Cossart, The virulence gene cluster of *Listeria monocytogenes* is also present in *Listeria ivanovii*, an animal pathogen, and *Listeria seeligeri*, a nonpathogenic species, Infect Immun. 62 (1994) 3550–3553. https://doi.org/10.1128/iai.62.8.3550-3553.1994.

[20] J.T. Burke, S.Z. Chen, L.M. Sosinski, J.B. Johnston, J. Ravi, MolEvolvR: A web-app for characterizing proteins using molecular evolution and phylogeny, BioRxiv. (2022). https://doi.org/10.1101/2022.02.18.461833.

[21] N.A. O'Leary, M.W. Wright, J.R. Brister, S. Ciufo, D. Haddad, R. McVeigh, B. Rajput, B. Robbertse, B. Smith-White, D. Ako-Adjei, A. Astashyn, A. Badretdin, Y. Bao, O. Blinkova, V. Brover, V. Chetvernin, J. Choi, E. Cox, O. Ermolaeva, C.M. Farrell, T. Goldfarb, T. Gupta, D. Haft, E. Hatcher, W. Hlavina, V.S. Joardar, V.K. Kodali, W. Li, D. Maglott, P. Masterson, K.M. McGarvey, M.R. Murphy, K. O'Neill, S. Pujar, S.H. Rangwala, D. Rausch, L.D. Riddick, C. Schoch, A. Shkeda, S.S. Storz, H. Sun, F. Thibaud-Nissen, I. Tolstoy, R.E. Tully, A.R. Vatsan, C. Wallin, D. Webb, W. Wu, M.J. Landrum, A. Kimchi, T. Tatusova, M. DiCuccio, P. Kitts, T.D. Murphy, K.D. Pruitt, Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation, Nucleic Acids Res. 44 (2016) D733–D745. https://doi.org/10.1093/nar/gkv1189.

[22] P.D. Karp, R. Billington, R. Caspi, C.A. Fulcher, M. Latendresse, A. Kothari, I.M. Keseler, M. Krummenacker, P.E. Midford, Q. Ong, W.K. Ong, S.M. Paley, P. Subhraveti, The BioCyc collection of microbial genomes and metabolic pathways, Briefings in Bioinformatics. 20 (2019) 1085–1093. https://doi.org/10.1093/bib/bbx085.

[23] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol Syst Biol. 7 (2011) 539. https://doi.org/10.1038/msb.2011.75.

[24] P. Stothard, The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences, Biotechniques. 28 (2000) 1102–1104.

[25] J.J. Campanella, L. Bitincka, J. Smalley, MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences, BMC Bioinformatics. 4 (2003) 29. https://doi.org/10.1186/1471-2105-4-29.

[26] T. Lassmann, E.L. Sonnhammer, Kalign – an accurate and fast multiple sequence alignment algorithm, BMC Bioinformatics. 6 (2005) 298. https://doi.org/10.1186/1471-2105-6-298.

[27] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2--a multiple sequence alignment editor and analysis workbench, Bioinformatics. 25 (2009) 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

[28] G. Yu, D.K. Smith, H. Zhu, Y. Guan, T.T. Lam, ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data, Methods Ecol Evol. 8 (2017) 28–36. https://doi.org/10.1111/2041-210X.12628.

[29] E. Paradis, K. Schliep, ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R, Bioinformatics. 35 (2019) 526–528. https://doi.org/10.1093/bioinformatics/bty633.

[30] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Research. 46 (2018) W296–W303. https://doi.org/10.1093/nar/gky427.

[31] H.M. Berman, The Protein Data Bank, Nucleic Acids Research. 28 (2000) 235–242. https://doi.org/10.1093/nar/28.1.235.

[32] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers, Protein Science. 30 (2021) 70–82. https://doi.org/10.1002/pro.3943.

[33] P. Boerlin, J. Rocourt, F. Grimont, P.A.D. Grimont, C. Jacquet, J.-C. Piffaretti, *Listeria ivanovii* subsp. *londoniensis* subsp. nov., International Journal of Systematic Bacteriology. 42 (1992) 69–73. https://doi.org/10.1099/00207713-42-1-69.

[34] M. Hupfeld, D. Trasanidou, L. Ramazzini, J. Klumpp, M.J. Loessner, S. Kilcher, A functional type II-A CRISPR–Cas system from *Listeria* enables efficient genome editing of large non-integrating bacteriophage, Nucleic Acids Research. 46 (2018) 6920–6933. https://doi.org/10.1093/nar/gky544.

[35] J. Rocourt, H. Hof, A. Schrettenbrunner, R. Malinverni, J. Bille, Acute purulent *Listeria seelingeri* meningitis in an immunocompetent adult, Schweiz Med Wochenschr. 116 (1986) 248–251.

CHAPTER 4

IDENTIFICATION OF NATURALLY OCCURRING INTERNALIN P VARIANTS ACROSS *LISTERIA MONOCYTOGENES* ISOLATES

ABSTRACT

The bacterial pathogen *Listeria monocytogenes* (*Lm*), causative agent of listeriosis, is one of few pathogens able to breach protective host barriers, leading to colonization of the placenta in the pregnant host. Placental listeriosis can lead to various adverse pregnancy outcomes including stillbirth, spontaneous abortion, and miscarriage. Several *Lm* virulence factors, including various internalins, have been identified as essential for placental colonization. Another internalin, InlP, was also previously found to be important for colonization of the placenta, likely by conferring the ability of *Lm* to transcytose protective barriers within the organ. Still, many uncertainties remain regarding InlP and its role in *Lm* pathogenesis. In this study, we use a bioinformatics approach to analyze publicly available assembled *Lm* whole genome sequences (WGS) and identify naturally occurring InlP variants in the *Lm* population. Furthermore, we have paired WGS data with available metadata to delineate links between InlP mutations and *Lm* serovars. To date, we have uncovered two naturally occurring InlP variants of interest in the *Lm* population. The first results from a start codon point mutation in the *inlP* gene, which is exclusive to the *Lm* 1/2b serovar. This variant is computationally predicted to produce a truncated and potentially nonfunctional InlP protein product. Additionally, we have identified *Lm* isolates harboring proline to serine substitutions in an InlP calcium binding loop, which has been hypothesized previously to play a InlP signaling and/or stabilization with its host binding partner. This mutation may result in altered Ca^{2+} binding affinities in InlP. These computational results are currently being validated experimentally by testing their effects on *Lm* transcytosis and *inlP* expression compared to wild-type *Lm* 10403S, following their substitution for native InlP in this strain. Together, these results provide two avenues for further investigation of InlP regulation and function. Furthermore, our

identification of naturally occurring InlP variants suggests the potential for InlP-dependent variation in placental colonization potential across *Lm* isolates.

INTRODUCTION

Listeria monocytogenes (*Lm*) is a Gram-positive saprophytic bacterium, and an opportunistic pathogen [1]. *Lm* can result in listeriosis in vulnerable populations and is one of few pathogens that can colonize the placenta, making *Lm* a particular threat to pregnant women [2,3]. Exposure to *Lm* typically happens through consumption of contaminated food items. Once in the gut, *Lm* utilizes a large repertoire of virulence factors to traffic to and colonize other organ systems [1]. Of distinct importance are the internalins, a large protein family found within all *Listeria* species [4]. The internalins encompass over twenty proteins in *Lm* alone, most of which remain uncharacterized [4]. Among the best understood internalins are Internalin A (InlA) and Internalin B (InlB) which are critical for invasion of the gut epithelium and hepatocytes, respectively. This is accomplished through respective binding of E-Cadherin and C-Met on host cells, which interact with the leucine rich repeat (LRR) regions of the internalins [5–7].

Previous studies have demonstrated variability in pathogenic tropism between different *Lm* serovars [8,9]. Altogether, there are 14 recognized *Lm* serovars, but approximately 95% of all human cases are attributed to only three of these: 1/2a, 1/2b, and 4b [10]. Furthermore, while serovar 1/2 strains predominate food-contaminating isolates, serovar 4b strains are responsible for >50% of all human infections [10]. Cases of placental listeriosis are strongly associated with serovar 4b strains, which are found more frequently in pregnancy-associated cases than in cases not associated with pregnancy [10,11]. This variability in tropism has been the topic of previous studies and has been speculated to be due in part to differences in expression of key virulence genes in *Lm* isolates [12]. Furthermore, previous studies have identified naturally occurring variants of InlA and InlB which have been shown to confer phenotypic differences [13–15].

Another internalin, Internalin P (InlP), has been identified as a key *Lm* virulence factor for promotion of placental colonization [16,17]. Prior characterization of InlP revealed several key features of this protein: nine LRR domains (LRR1-9), one calcium binding loop in LRR3, a signal peptide, and a C-terminal IgG-like domain which is characteristic of internalins [17]. While the Δ*inlP* strain of *Lm* showed moderately reduced placental colonization of the liver and spleen, there was a nearly 4-log decrease in placental colonization compared to wild-type, suggesting that *inlP* confers placental tropism [16]. It is hypothesized that this is accomplished through InlP's binding of afadin, a eukaryotic protein shown to be a binding partner of InlP, which enhances transcytosis through placental layers through an unknown mechanism [17]. Despite this knowledge, much remains to be elucidated regarding InlP and its structure, function, regulation, and binding partner(s).

We hypothesized that naturally occurring variants of InlP also exist in the *Lm* population and that variants could inform on open questions surrounding InlP. In this study, we have identified 95 *Lm inlP* sequences from GenBank and PATRIC databases containing *Lm* whole genome sequences (WGS). We extracted the *inlP* sequences from each of these WGS and translated them to InlP amino acid sequences *in silico* and explored these sequences for variants of interest. Pairing these sequences with corresponding serovar and isolate source data allowed for investigation of association between variants and serovar/source. We identified two variants of interest, including a truncated variant strongly associated with serovar 1/2b and a P128S substitution in the InlP Ca^{2+} binding loop, present in approximately half of our representative isolates. We introduced these mutations of interest into the *Lm* 10403S and Xen32 background laboratory strains to allow for investigation of their phenotypic consequences. To the authors' knowledge, this is the first study

identifying naturally occurring variants of InlP. Together, our data begin to shed light on InlP, and our study provides new tools for studying InlP function.

MATERIALS AND METHODS

L. monocytogenes Sequence Search

To begin investigating the possibility of internalin P variants occurring within diverse isolates of *Lm*, we queried GenBank and PATRIC databases for assembled whole genome sequences of *Lm* [18,19]. We then downloaded accession numbers and all associated metadata for each of these sequences. We filtered our sequences, only keeping those with an associated serovar and source, which resulted in 94 total isolates of interest (**Table 4.1**).

Extraction of *inlP* Sequences and Subsequent Translation

To extract the *inlP* gene sequence from each assembled *Lm* whole genome sequence, we utilized BLAST [20]. Each sequence of interest was aligned with the *Lm* 10403S *inlP* sequence (GenBank Accession: CP002002.1, locus tag lmrg_01778) and the resulting nucleotide sequences were saved in a multi-FASTA file. Nucleotide sequences were translated to amino acid sequences by the Expasy Translate Tool using the standard genetic code [21].

Multiple Sequence Alignment and Identification of Variants of Interest

All InlP amino acid sequences were aligned using Kalign with default settings [22]. This alignment was visualized using Jalview, which allowed for visualization of individual nucleotide conservation [23]. Regions of low conservation were compared to previous literature discussing InlP to determine which of these regions could help inform on the function and activity of InlP, thus worth continuing with further characterization.

Phylogenetic Trees

Phylogenetic neighbor joining trees for all identified InlP variant amino acid sequences were constructed using the Simple Phylogeny tool from EMBL-EBI using default settings following multiple sequence alignment (see *Multiple Sequence Alignment and Identification of Variants of Interest*) [24]. Trees were visualized and annotated using Interactive Tree of Life [25].

Open Reading Frame Prediction and Protein Modeling

For InlP sequences with predicted truncated amino acid sequences compared to the reference, we completed open reading frame prediction using ORFfinder with the standard genetic code [26]. Protein modeling for InlP variants of interest was conducted using SWISS-MODEL with the previously resolved InlP crystal structure (PDB: 5hl3) as a template [27,28]. Protein models were visualized using ChimeraX [29].

InlP Variant Construction

For phenotypic studies of the identified InlP variants of interest, we introduced two variants (InlP.2, harboring a start codon point mutation in the *inlP* gene and InlP.3, harboring a proline to serine residue substitution in the Ca^{2+}-binding loop of InlP's LRR3) into the background of the *Lm* 10403S lab strain [30]. Additionally, we constructed an in-frame deletion of *inlP* in this background strain to use as a control. To accomplish this, we used the pKSV7x plasmid and designed all constructs *in silico* using SnapGene [31]. First, pKSV7x was PCR linearized by inverse PCR with Q5 polymerase (NEB) (**Primers: Table 4.2**). The linearized PCR product was run on a 1% agarose gel and purified using the Qiagen Qiaquick Gel Extraction Kit according to manufacturer protocols. To introduce the appropriate mutations in *inlP* for construction of variants, we used 10403S genomic DNA as a template and PCR amplified *inlP* fragments with primers designed to introduce the correct mutations (**Table 4.2**). Primers were designed to allow for amplification of ~500bp flanking the *inlP* gene to allow for downstream allelic exchange in *Lm* 10403S, as well as ~15bp of homology with pKSV7x to allow for construction by Gibson Assembly. Final constructs were assembled by Gibson Assembly using the NEB Gibson Assembly Cloning Kit, transformed into *E. coli* DH5α, and transformations were plated on LB + 50 μg/mL carbenicillin for selection. Resulting transformants were screened for the presence of *inlP* by colony PCR, and putative hits were grown in liquid culture and miniprepped using the Qiagen QIAprep Spin Miniprep Kit. Resulting constructs were screened by restriction digest and confirmed by Sanger sequencing (GeneWiz).

inlP Allelic Exchange

Correct constructs confirmed by Sanger sequencing were transformed into *E. coli* SM10, an F+ strain that allows for transconjugation. Resulting transformants were grown in liquid culture, miniprepped using the QIAprep Spin Miniprep Kit, and screened via restriction digest. Confirmed transformants were saved as glycerol stocks for downstream transconjugation experiments. For transconjugation of each construct, the *E. coli* SM10 strain containing our construct and *Lm* 10403S or *Lm* Xen32 (a bioluminescent *Lm* strain generated in the 10403S background) were grown to mid-log phase. Equal volumes of SM10 and *Lm* were combined then filtered onto a 0.22 μM syringe filter. This filter was incubated on non-selective BHI agar overnight at 30°C. The following day, the filter was scraped with a sterile inoculation loop which was then used to streak for isolation on BHI agar with 200 μg/mL streptomycin (Strep) to select against *E. coli* (*Lm* 10403S is streptomycin resistant) and 7.5 μg/mL chloramphenicol (Cm) to select for *Lm* containing our construct of interest. Plates were incubated at 30°C until single colonies appeared (~48-72H). One transconjugant colony was chosen to re-streak in duplicate on selective BHI with 200 μg/mL Strep and 7.5 μg/mL Cm; one plate was incubated at 30°C (permissive for pKSV7x) while the other was incubated at 42°C (non-permissive for pKSV7x to select for integrants). One putative integrant colony was chosen and restreaked on selective BHI with 200 μg/mL Strep and 7.5 μg/mL Cm and incubated at 42°C until single colonies appeared (~48H). This was repeated one final time to purify our integrant. To cure mutant strains of the pKSV7x plasmid, a single colony was used to inoculate BHI broth with 200 μg/mL streptomycin for selection of *Lm*. Liquid cultures were passed by dilution 1:1000 in fresh medium twice per day for five days. Beginning with passage four, we plated on selective BHI agar with 200 μg/mL Strep. Single colonies were patched in duplicate on

selective BHI agar with 200 μg/mL Strep and a second selective BHI plate with 200 μg/mL Strep and 7.5 μg/mL Cm to screen for the expected $Strep^R/Cm^S$ phenotype of pKSV7x-cured integrants. Clones expressing the $Strep^R/Cm^S$ phenotype were grown in liquid culture and desired mutations were confirmed by Sanger sequencing of the appropriate *inlP* region (**primers: Table 4.2**) (GeneWiz).

Growth Curves

For growth curves of newly generated *Lm* mutants, each mutant strain was grown overnight in BHI broth supplemented with 200 μg/mL Strep for selection. The following morning, cultures were back-diluted in selective BHI broth (200 μg/mL Strep) to a starting OD_{580} of 0.01. Cultures were plated in triplicate in a 96-well plate and OD_{580} measurements were recorded using the PerkinElmer VICTOR Nivo microplate reader every hour for twelve hours. Absorbance values for sterile medium were subtracted from each experimental absorbance value.

Bacterial Strains and Growth Conditions

Lm 10403S and Xen32 strains were grown directly from glycerol stocks in BHI medium, shaking at 37°C unless otherwise indicated. *E. coli* DH5α and SM10 strains were grown directly from glycerol stocks in BHI medium, shaking at 37°C unless otherwise indicated. Where selection was appropriate for variant construction and allelic exchange, antibiotic name and concentration are indicated in those respective methods sections.

RESULTS

Identification and classification of *Listeria monocytogenes* sequences of interest

To begin investigating the possibility of heterogeneity in the *inlP* gene across *Lm* isolates, we queried GenBank and PATRIC databases for assembled whole genome sequences (WGS) of *Lm* [18,19]. To identify any associations between variants and serovar or isolate source, we filtered the resulting assembled sequences, only keeping those with serovar and source metadata available. This resulted in 95 total sequences of interest. To extract the *inlP* sequence, we first obtained the *inlP* sequence from *Lm* 10403S (GenBank accession CP002002.1) and used BLAST to align this sequence with each WGS previously identified [18,20]. Resulting *inlP* nucleotide sequences were saved in a multi-FASTA file. Unsurprisingly, the majority of our sequences fell within serovars 1/2a and 4b (**Fig. 4.1A**). Less common serovars (1/2b, 1/2c, 3c, 4a, and 4d) were also represented in our study (**Fig. 4.1A**). Isolate source metadata entered by those submitting sequences can range drastically in specificity. Therefore, to simplify our analysis, we assigned each sequence to one of five source categories based on the available source metadata: clinical (human), veterinary, food, environmental, or laboratory. Overall, most of our samples were human clinical or food isolates (**Fig. 4.1B**).

Identification of variants and InlP phylogeny

To begin identifying variants of InlP, we first translated the *inlP* nucleotide sequences to amino acid sequences using Expasy (see *Methods*) [21]. Amino acid sequences were aligned using Kalign and the resulting alignment was used to produce a neighbor joining tree of InlP sequences (**Fig. 4.2**) [22,24]. This neighbor joining tree illustrated three distinct clades, which seemed to

group by serovar (**Fig. 4.2A**). Coloring sequences by lineage revealed that these three clades were distinctly *Lm* evolutionary lineages I, II, and III (**Fig. 4.2B**). Visualization of the multiple sequence alignment in Jalview allowed us to identify regions of low conservation to determine how InlP sequences differ from one another [23]. Because one or more LRR domains are hypothesized to serve as binding sites for InlP and its binding partner(s), we were interested to also find an area of low conservation in LRR5 (**Fig. 4.3A**).

The InlP Ca^{2+} binding loop harbors a P128S substitution in approximately half of all isolates

Further investigation of InlP variants revealed the noteworthy discovery that approximately half of the representative isolates in this study harbor a proline to serine substitution at amino acid 128, corresponding to a Ca^{2+} binding loop in LRR3 (**Fig. 4.3B**). This is of particular interest due to this region's predicted role in InlP activity and/or protein-protein stabilization between InlP and its binding partner(s). To further investigate this substitution's effect on InlP's calcium-binding ability, we generated a model of $InlP^{P128S}$ using SWISS-MODEL and found that this residue substitution results in altered bond angle measurements in the Ca^{2+} binding loop (**Fig. 4.4**) [27].

To create the possibility of validating computational predictions with wet lab experiments, we generated mutants of *Lm* 10403S and Xen32 expressing the $InlP^{P128S}$. In growth curve experiments, both mutants grew similarly to their wild type counterparts (**Fig. 4.9**).

Identification of a truncated InlP variant

We noted that 14 of the isolates in our analysis contained a point mutation in the *inlP* start codon, which is predicted to result in a frameshift and truncated protein product (**Fig. 4.5, Fig. 4.10, Table 4.1**). This truncation results in a complete loss of the InlP signal peptide region and a

partial loss of the N-terminal region (**Fig. 4.5**). Interestingly, this mutation was strongly associated with the 1/2b serovar; out of 14 isolates harboring this mutation, 13 were indicated to be serovar 1/2b (**Table 4.1**).

DISCUSSION

Placental infection by *Lm* can be devastating for both mother and infant, often resulting in stillbirth, miscarriage, developmental delay, and/or deadly neonatal meningitis. Pregnant mothers are approximately 10 times more likely to contract listeriosis than non-pregnant people, and roughly 17% of all cases of listeriosis can be attributed to pregnancy [3]. While still considered relatively rare, there are concerns of increasing prevalence of listeriosis due to the expansion of processed food production and the *Lm* ability to survive on food production equipment. The largest listeriosis outbreak in history occurred recently (2018) in South Africa and was the result of large-scale food contamination [32]. Devastatingly, approximately half of all cases in this outbreak were pregnancy-associated.

Until the identification of InlP, no virulence factors conferring specific placental tropism had been described in *Lm*. The discovery of InlP has opened the door for studies of placental listeriosis that will further elucidate the mechanisms by which *Lm* colonizes this organ that is largely restrictive to pathogens. Furthermore, future studies of InlP could offer insight to the mechanisms used by other bacterial pathogens that invade the placenta.

This study describes naturally occurring variants of InlP and their associations (or lack thereof) with *Lm* serovar and source type. We have identified several variants of interest that have the potential to inform on key aspects of InlP regulation and function. Phylogenetic analysis of the 94 InlP amino acid sequences used in our study suggest distinct groupings by both *Lm* serovar and

lineage. Alignment of these sequences allowed us to assess conservation of each amino acid residues and identify regions of low conservation. We were particularly interested in poorly conserved regions in the LRR domains, as these have been hypothesized to be the site(s) of InlP interaction with its host binding partner, afadin. One variant of interest identified in this study is $InlP^{P128S}$ which harbors a proline to serine substitution in the LRR3 Ca^{2+} binding loop of InlP. Because proline is known to be a key structural residue, we modeled this substitution using WT InlP as a template, and found that this substitution alters the bond angles within the Ca^{2+} binding loop. This is a key finding due to the Ca^{2+} binding loop's hypothesized role in InlP activation and/or stabilization. We have introduced this mutation into the *Lm* 10403S and Xen32 background strains to allow for further investigation of this mutation's phenotypic effects.

Another variant of interest was a truncated InlP resulting from a point mutation in the *inlP* start codon and predicted frameshift. This truncation results in complete loss of the InlP signal peptide and partial loss of its N-terminal domain. Interestingly, this variant was strongly associated with the 1/2b serovar, suggesting one likely explanation for the previously observed differences in *Lm* pathogenic tropisms. Because the mechanism by which InlP is secreted from *Lm* remains to be identified, it is difficult to predict the phenotypic effect of this truncation. This mutation has been introduced into the *Lm* 10403S and Xen32 background strains to allow for further study. We believe it is noteworthy that the pKSV7x construct containing this mutation resulted in apparent partial lysis of *E. coli* DH5α when grown at 37ºC during the cloning process (**Fig. 4.11**). This effect was avoided when growing the transformants at 30ºC and was not observed when growing pKSV7x harboring WT *inlP* in *E. coli* DH5α. Together, this suggests that introduction of this point mutation results in a change to InlP that is toxic to *E. coli*.

In addition to the InlP variants reported in this study, we recently reported the discovery of InlP homologs encoded by other *Listeria* species apart from *Lm*. *L. seeligeri*, *L. ivanovii londoniensis*, *L. innocua*, and *L. costaricensis* are not known to colonize the placenta but encode InlP homologs [33]. Interestingly, these homologs lack the full-length LRR6 and LRR7 domains [33]. Because the site of interaction between InlP and its binding partner, afadin, has not yet been identified, it will be interesting to assess the afadin-binding ability of these homologs. Lack of binding would suggest a potential role for LRR6 and/or LRR7 in the InlP-afadin interaction.

In summary, we have identified and described naturally occurring InlP variants in *Lm*. We have analyzed these variants computationally for their effects on InlP structure, allowing us to speculate on their potential phenotypic effects. We have introduced two of these variants of interest into the *Lm* 10403S and Xen32 laboratory strains to allow for future their future experimental analysis. This study offers new information regarding a key *Lm* virulence factor and has generated new tools for its study in the laboratory. Future studies should address how InlP variants contribute to differences in *Lm* pathogenic and tissue tropism, and how InlP mutations affect phenotype.

DECLARATIONS

Competing Interests

The authors declare that there are no competing interests.

Funding

KNC is supported by Michigan State University (MSU) Microbiology and Molecular Genetics departmental fellowships. This work was supported by MSU start-up funds granted to JWH.

Contributions

Conceptualization: KNC. *Methodology:* KNC. *Investigation:* KNC. *Resources:* JWH. *Data Curation:* KNC. *Writing:* KNC, JWH. *Visualization:* KNC. *Supervision:* JWH. *Project Administration:* JWH. *Funding:* JWH.

Acknowledgements

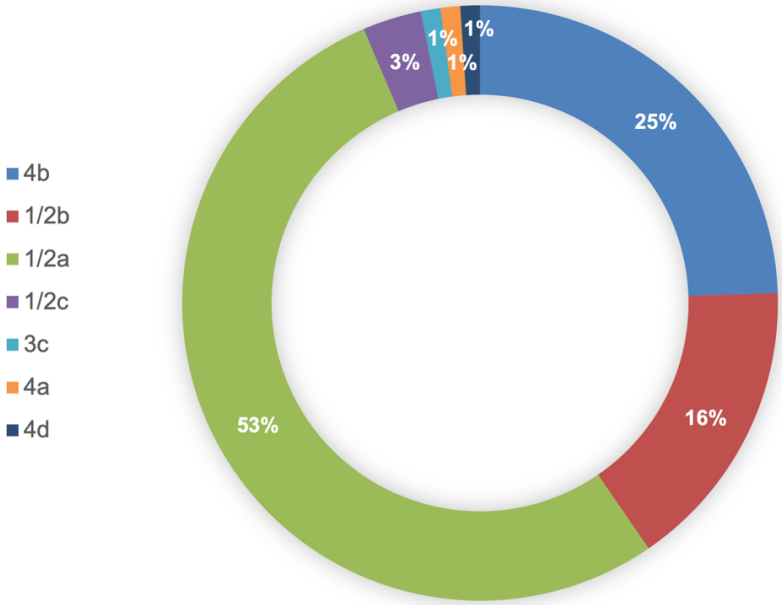
The authors would like to acknowledge the Portnoy Laboratory and Dr. Preethi Ragunathan at University of California Berkeley for generously providing the pKSV7x plasmid used in our allelic exchange experiments and for providing troubleshooting assistance with construct generation and allelic exchange.

APPENDIX

FIGURES AND TABLES

Figure 4.1 Breakdown of study isolates by serovar and source. In this study, we extracted and translated the InlP sequence from 94 publicly available *Lm* sequences with serovar and source metadata available. The following display the breakdown of these isolates by serovar (A) and source (B).

A.



B.

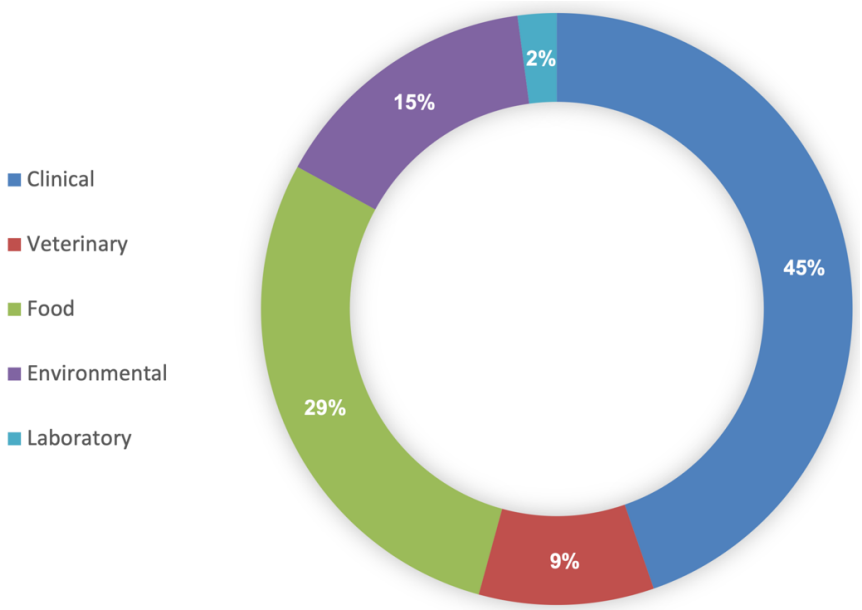


Figure 4.2 Neighbor-joining trees of InlP sequences. InlP amino acid sequences were used to generate a neighbor joining tree using Simple Phylogeny. The same tree is presented twice below, annotated based on *Lm* serovar (A) and *Lm* lineage (B). Annotations were completed using Interactive Tree of Life. Labels are GenBank accession numbers for the corresponding nucleotide sequence.

A.

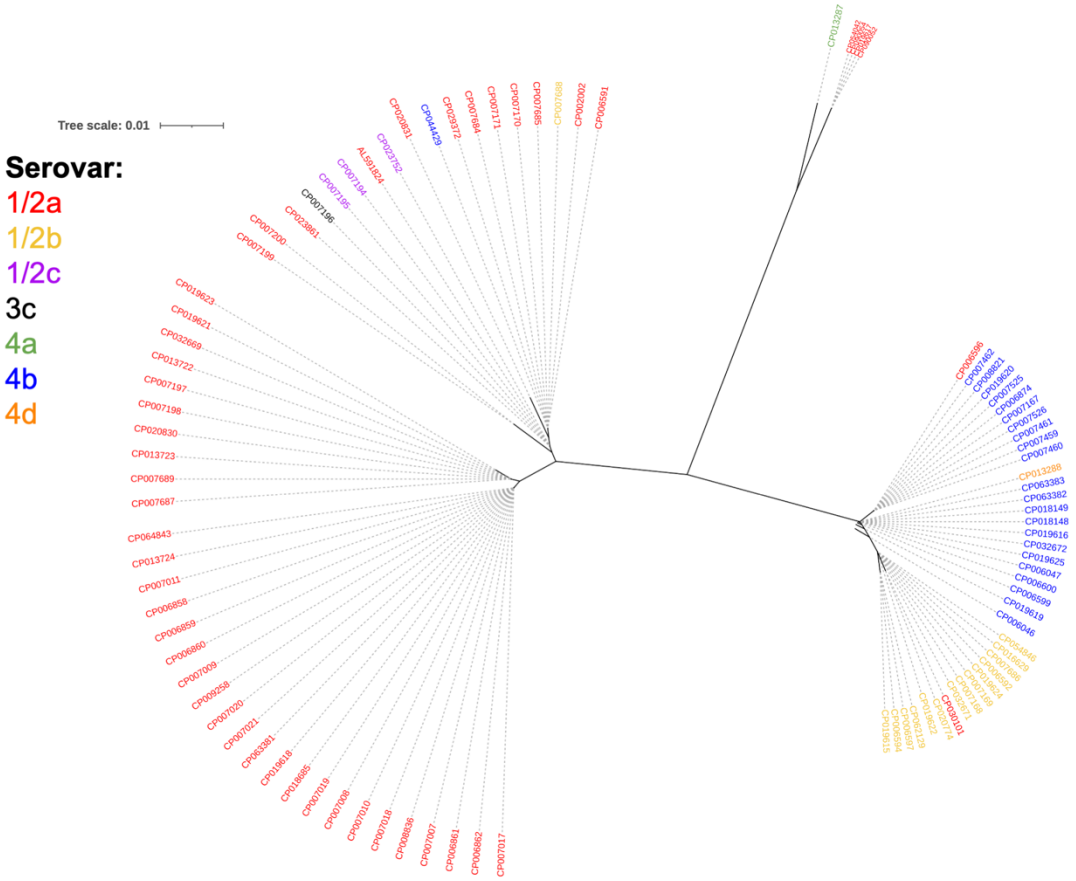


Figure 4.2 (Cont'd)

B.

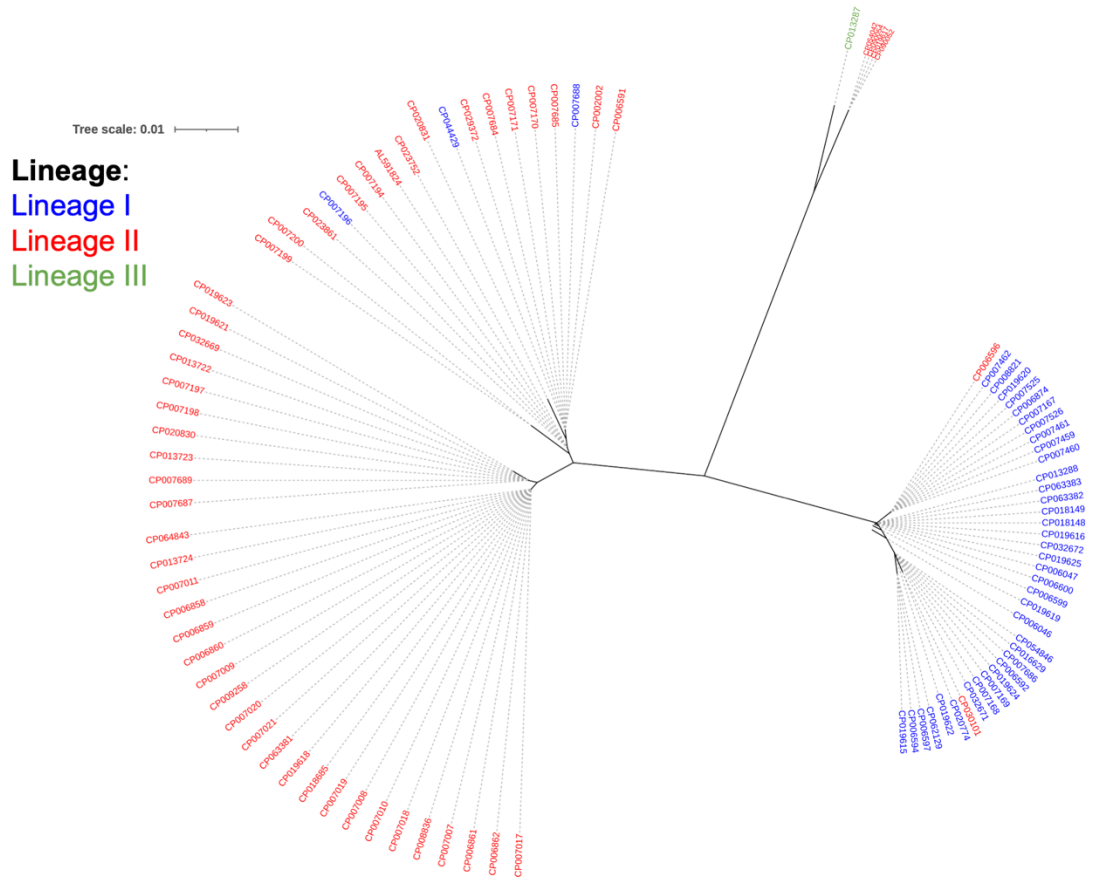
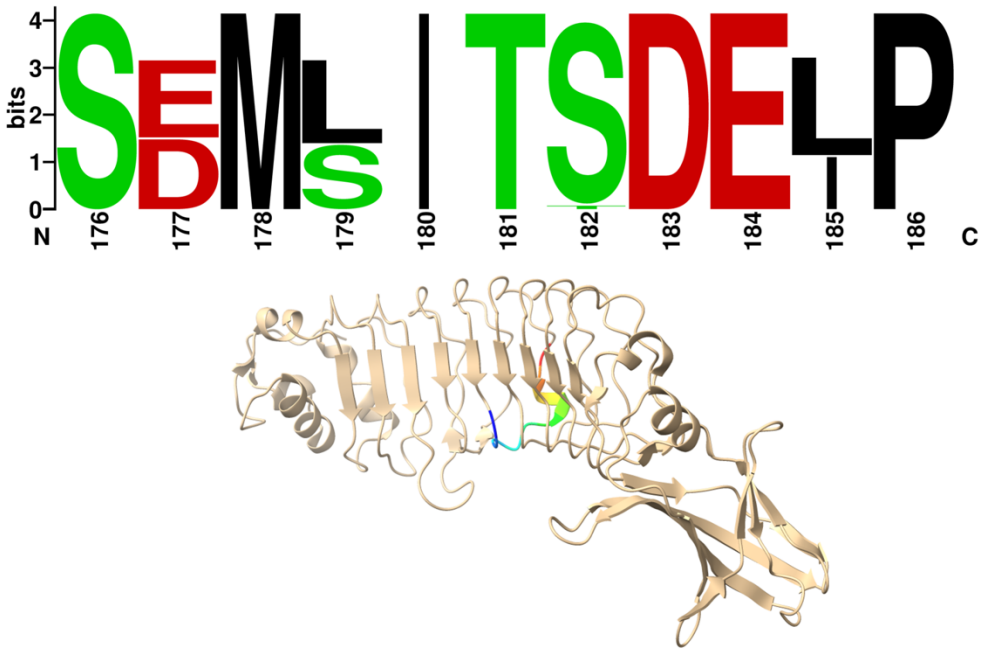


Figure 4.3 Logo diagrams and models of variants of interest. Three variants of interest are represented below as logo diagrams and corresponding protein models. One region of interest in LRR5 harbored three poorly conserved residues (A). Approximately half of the isolates in our study harbored a proline to serine substitution at residue 128, corresponding to the InlP Ca^{2+} binding loop (B). Logo diagrams were generated using WebLogo. The x-axis displays each residue number, and the y-axis represents each residue's relative occurrence in our samples. Protein model images were generated using ChimeraX to visualize the InlP crystal structure (PDB: 5hl3), and regions of low conservation are highlighted in rainbow coloring.

A.



B.

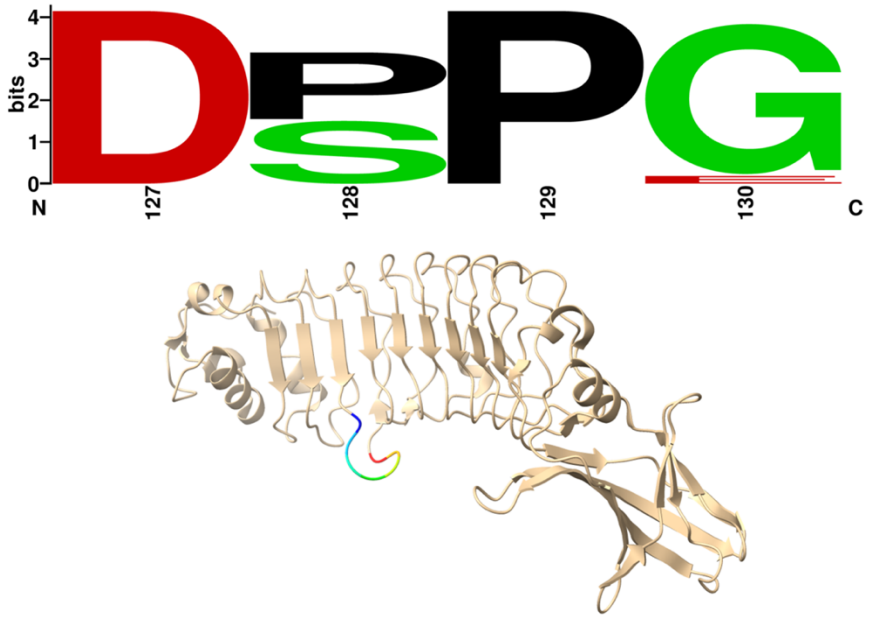
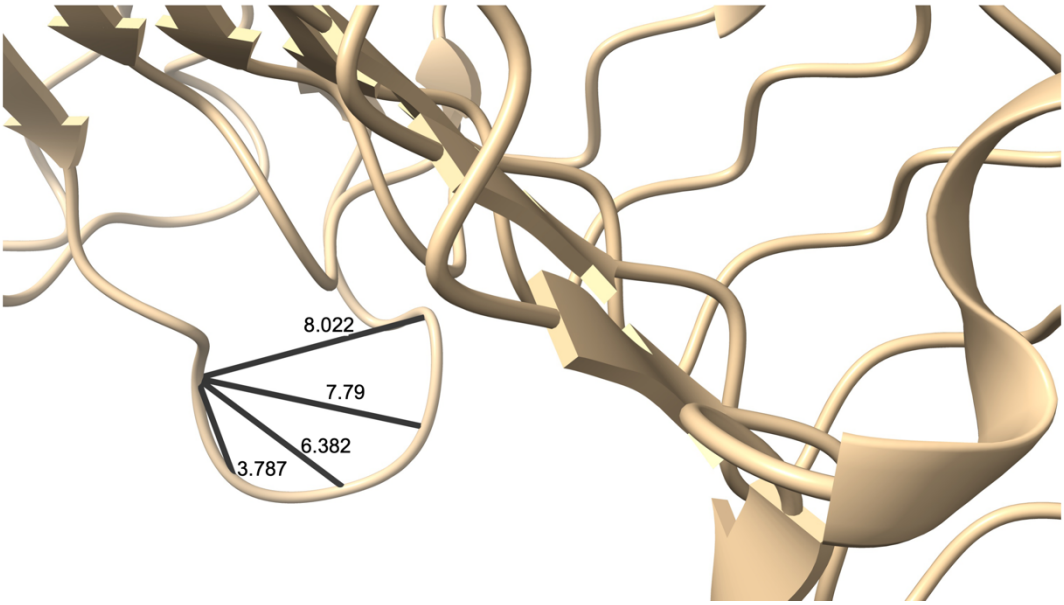


Figure 4.4 InlP^{P128S}: A substitution in the InlP Ca^{2+}-binding loop. After identifying the P128S substitution in InlP, we generated a model of this variant using SWISS-MODEL [27]. The Ca^{2+} binding loop regions of the wild-type InlP (A) protein and InlP^{P128S} (B) were visualized using ChimeraX and measurements were taken of distances between residue 128 and each other residue in the Ca^{2+}-binding loop [29]. Distances are expressed in angstroms.

A.



B.

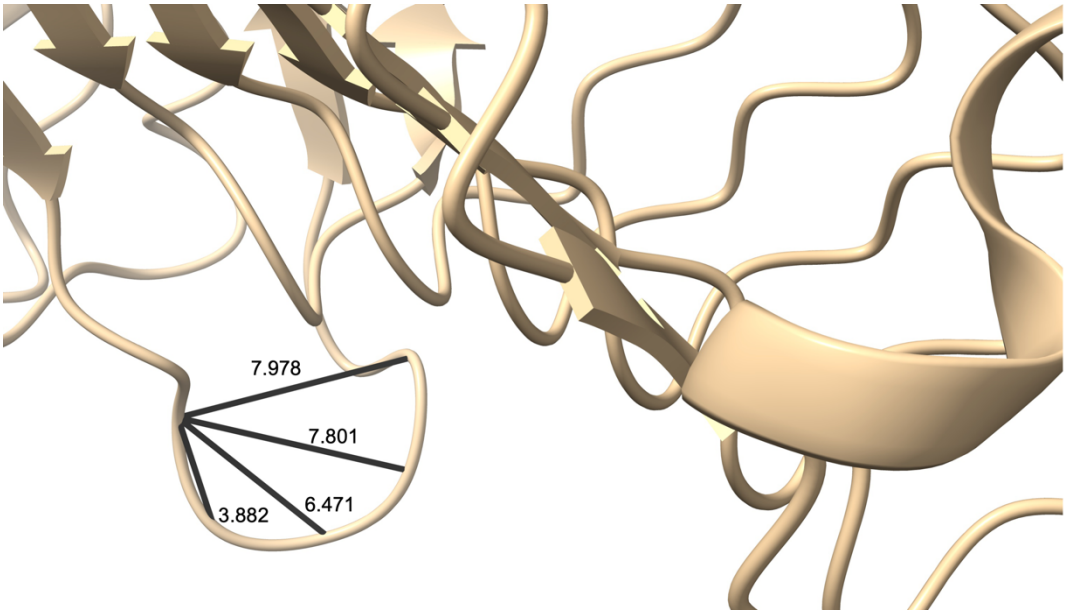


Figure 4.5 A truncated InlP variant. In our analysis, we identified several isolates harboring an InlP variant with a predicted 50 amino acid truncation on the C-terminus. A model of this variant was generated with SWISS-MODEL and visualized using ChimeraX (A). The wild-type InlP protein model is also presented with the truncated region highlighted in rainbow (B).

A.

B.

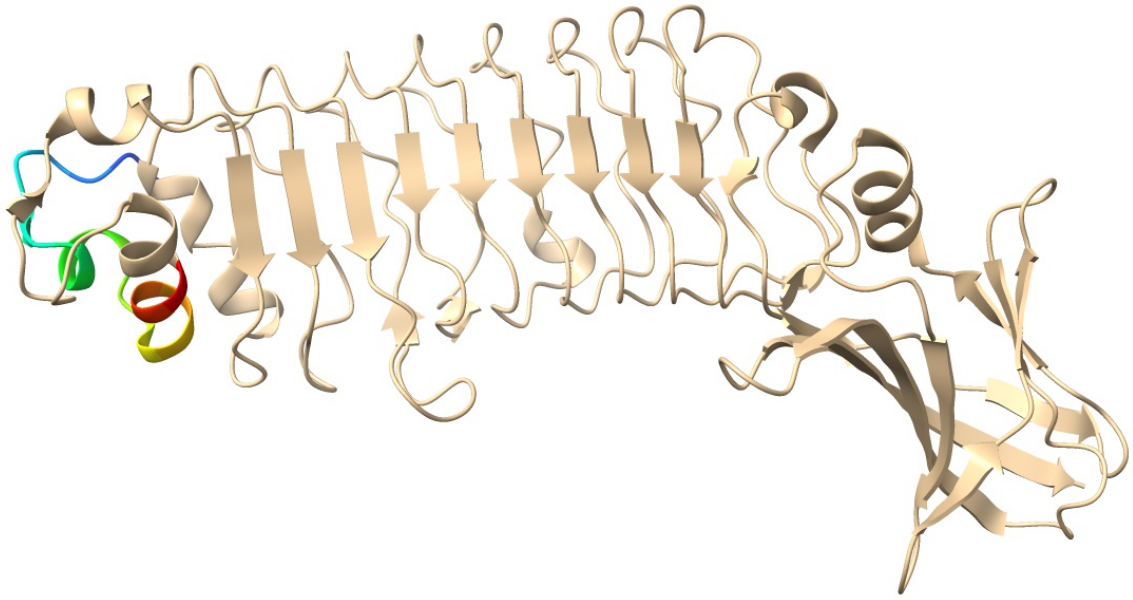


Figure 4.6 Construct maps for generation of InlP variant mutant strains. The pKSV7x plasmid was used to generate in-frame substitutions for each InlP variant of interest, as well as an in-frame deletion of the *inlP* gene. Maps for the pKSV7x vector (A), the pKSV7x-InlP.3 construct (B), the pKSV7x-InlP.2 construct (C), and the pKSV7x-InlP-Knockout construct (D) were generated with SnapGene.

A.

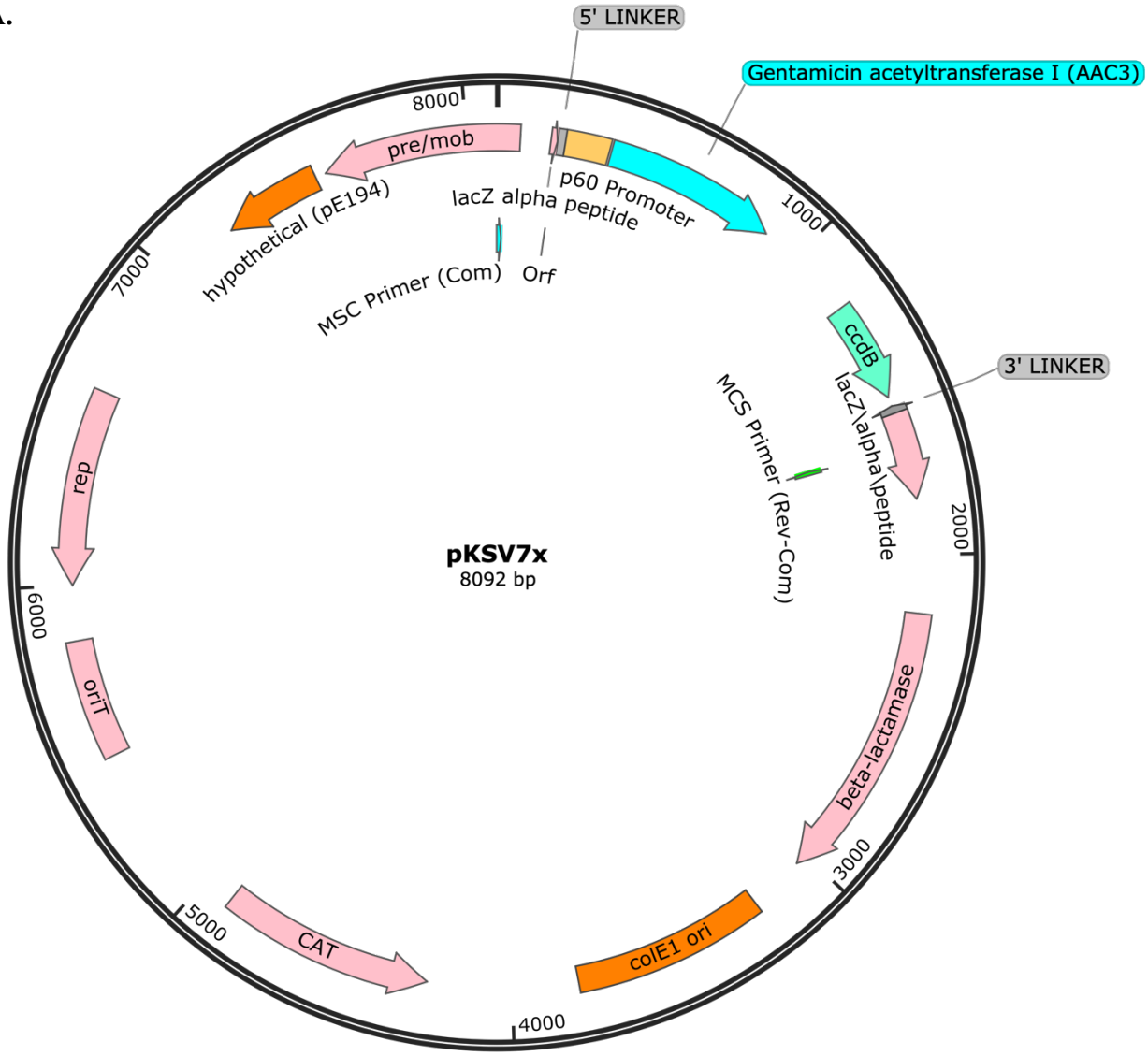


Figure 4.6 (Cont'd)

B.

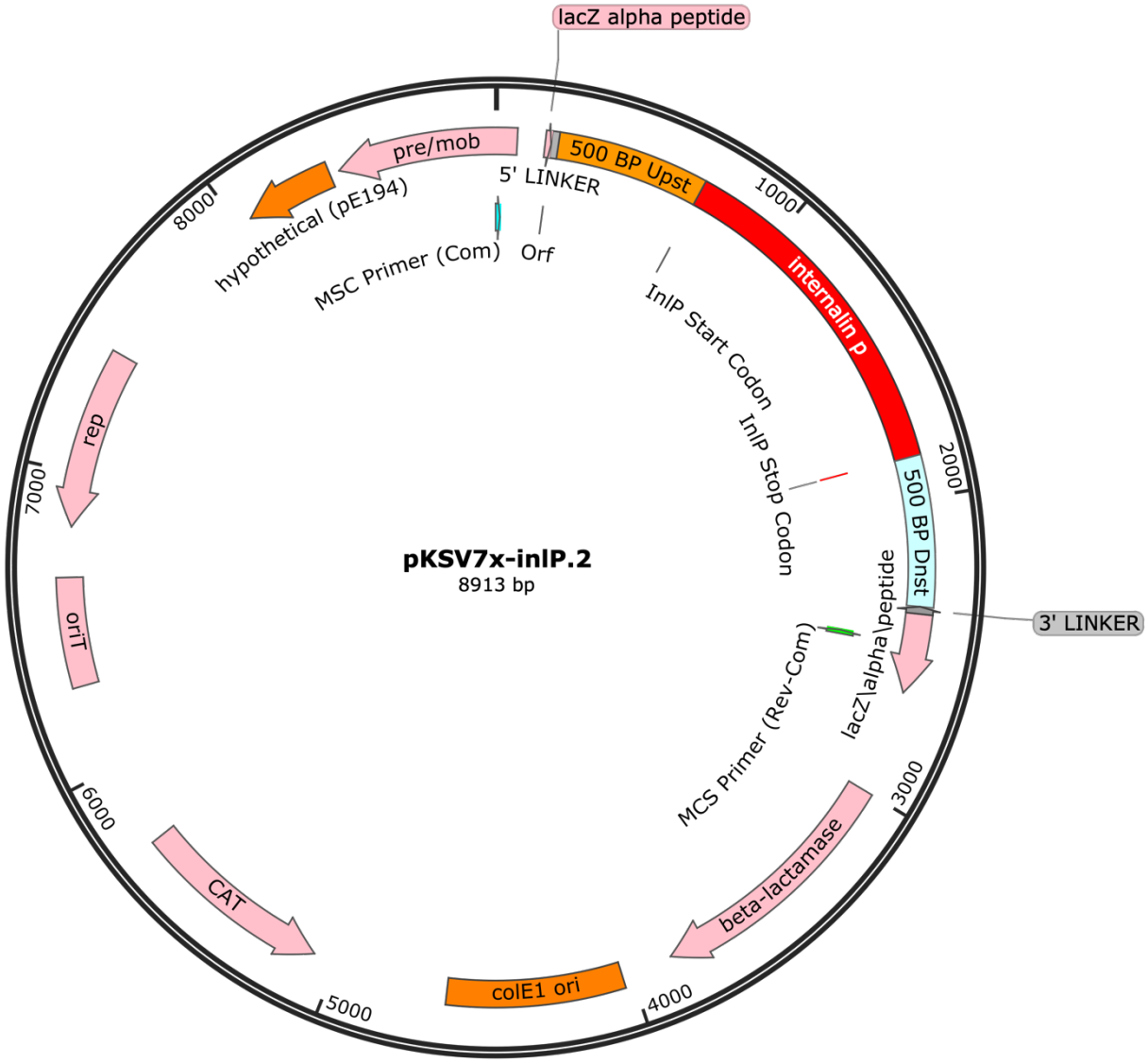


Figure 4.6 (Cont'd)

C.

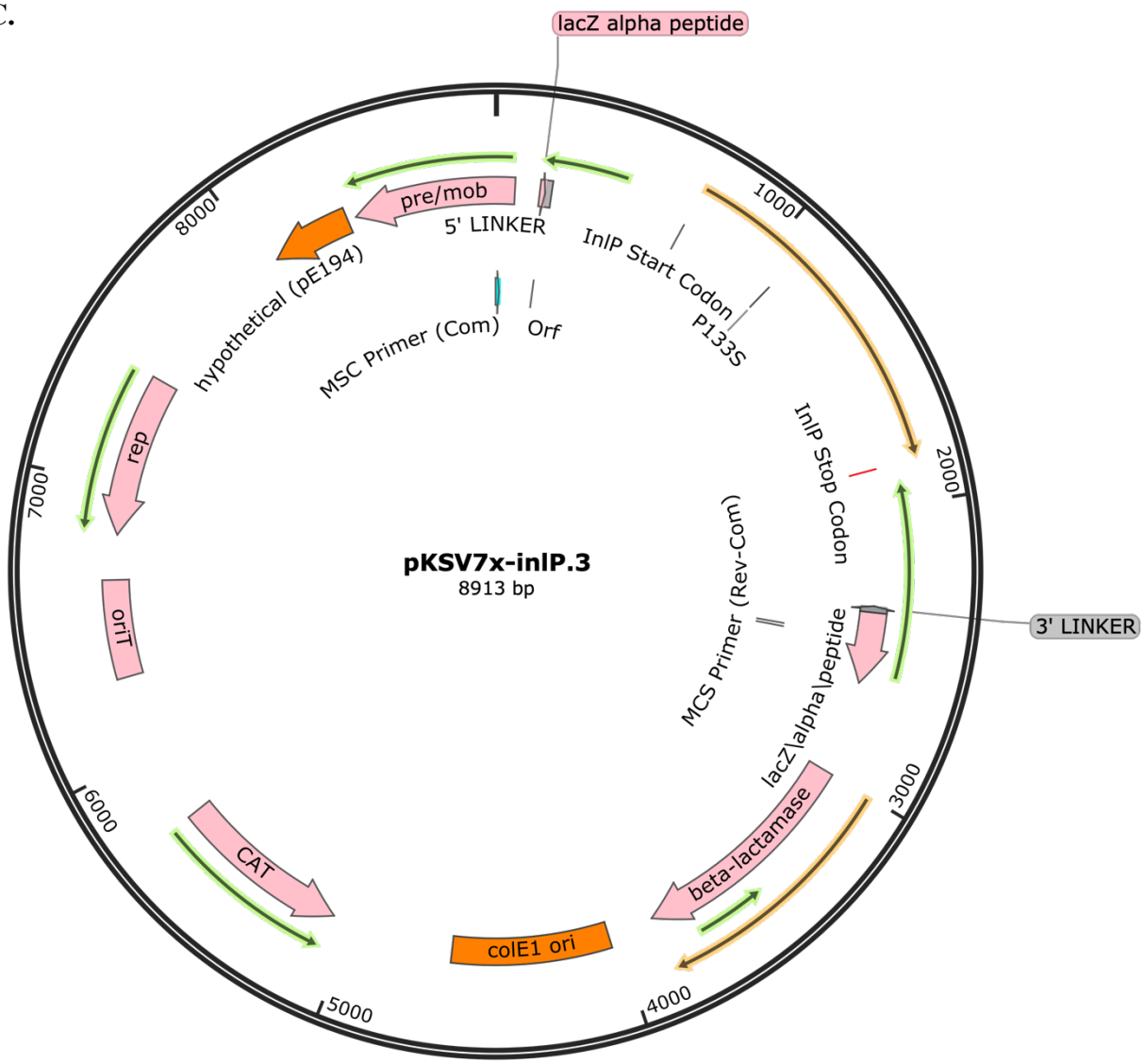


Figure 4.6 (Cont'd)

D.

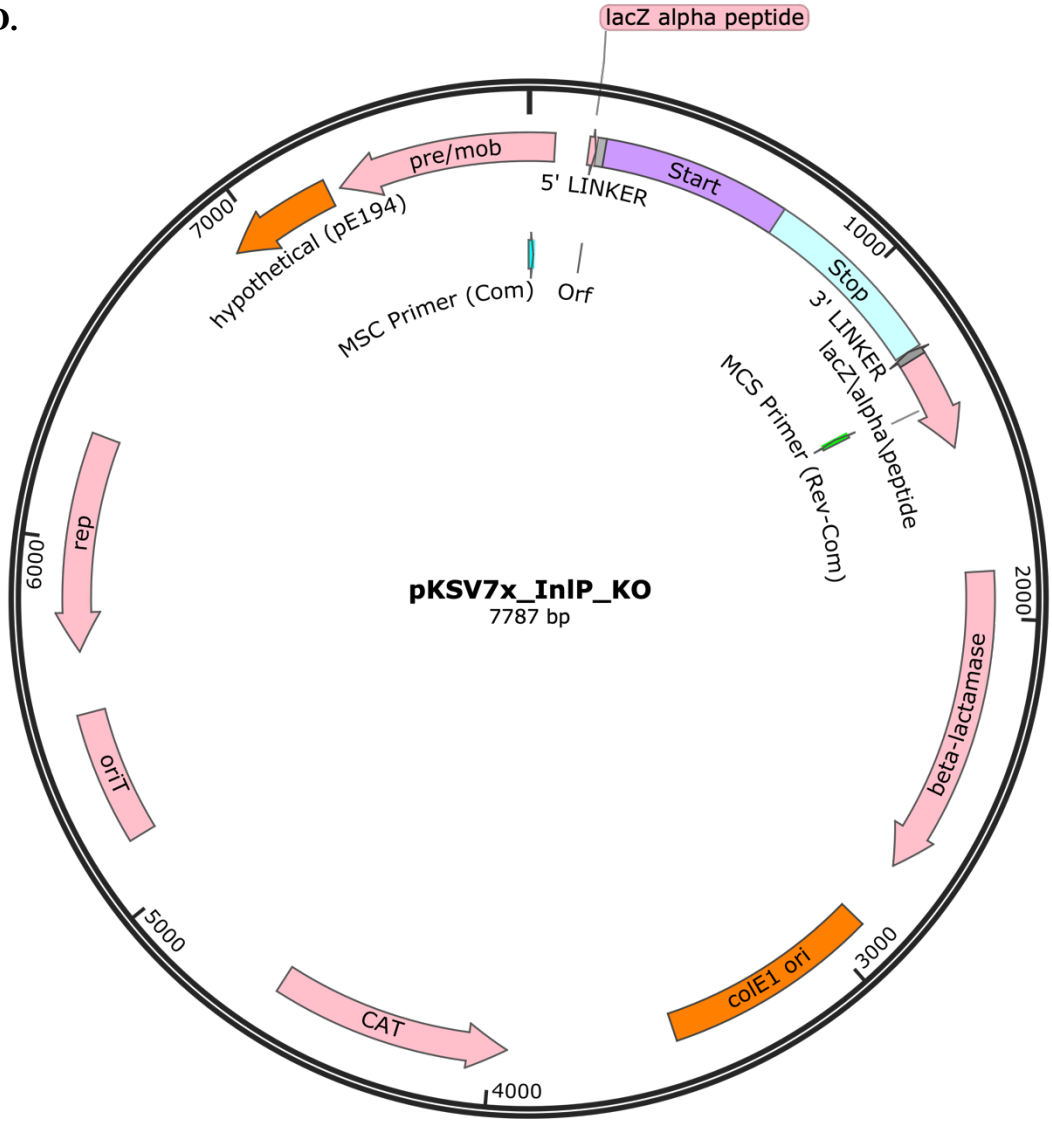
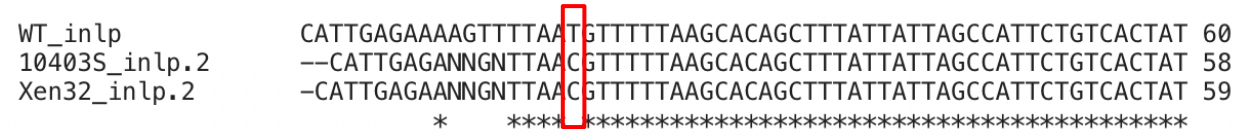


Figure 4.7 Sanger sequencing confirmation of mutations of interest. Upon completion of transconjugation of our mutant constructs into *Lm* 10403S and *Lm* Xen32, purified genomic DNA was sent to GeneWiz for Sanger sequencing confirmation that our *inlP* mutations of interest were introduced. Partial sequences of the mutated regions for *inlP*.2 (A) and *inlP*.3 (B) in both 10403S and Xen32 were aligned with the wild type *inlP* gene sequence from *Lm* 10403S to verify appropriate mutations, outlined in red below.

A.



B.

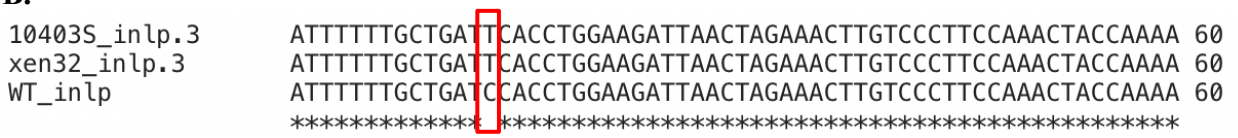


Figure 4.8 Confirmation of successful *inlP* knockout in *Lm* 10403S and Xen32. To confirm that the *inlP* gene was successfully deleted from our *Lm* 10403S (A) and Xen32 (B) background strains, we amplified the entire *inlP* region in both wild type and putative knockout strains.

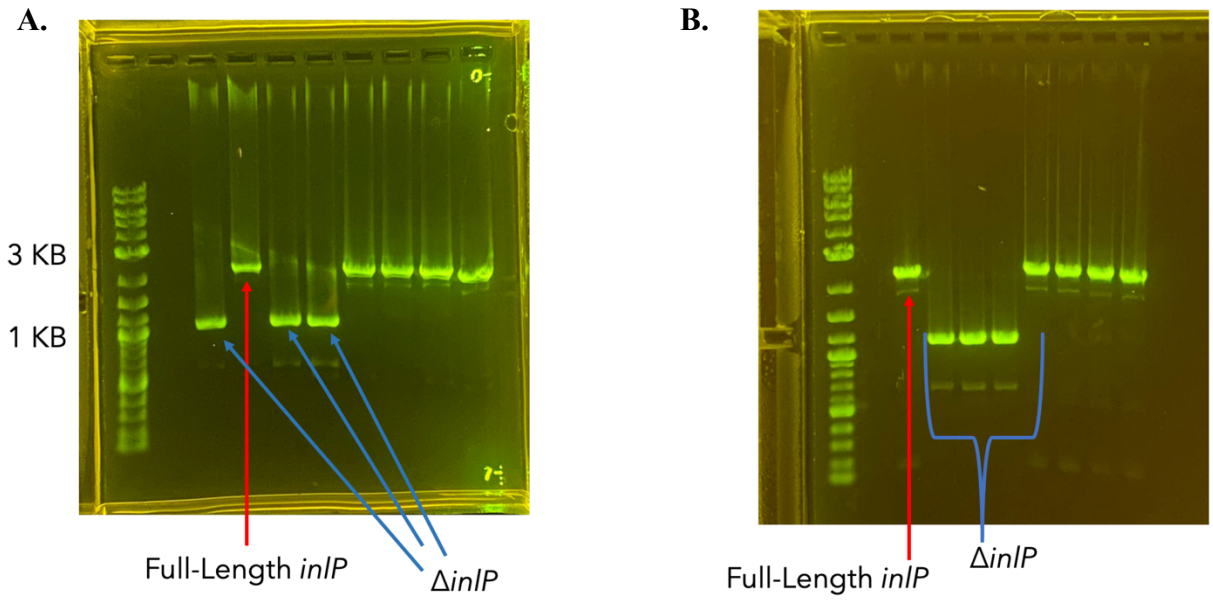
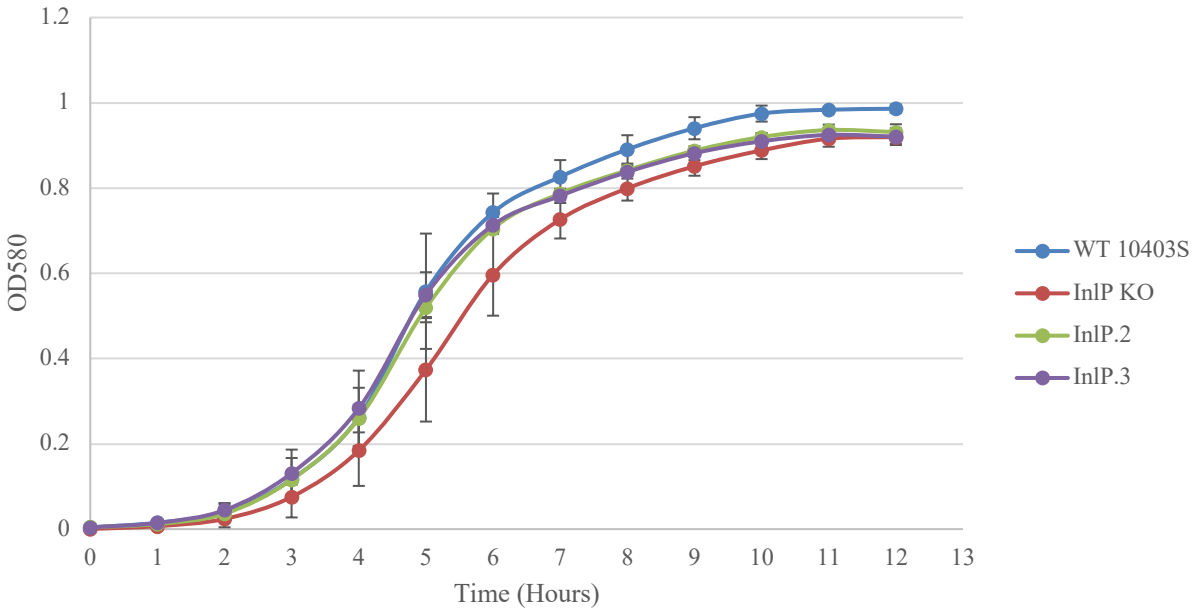


Figure 4.9 Growth curves for mutants generated in this study. Growth curves were carried out for all mutants generated in this study as well as their wild-type counterparts for both 10403S (A) and Xen32 (B) background strains to ensure that the introduced mutations did not confer a growth defect. Error bars are standard deviation (n = 3).

A.



B.

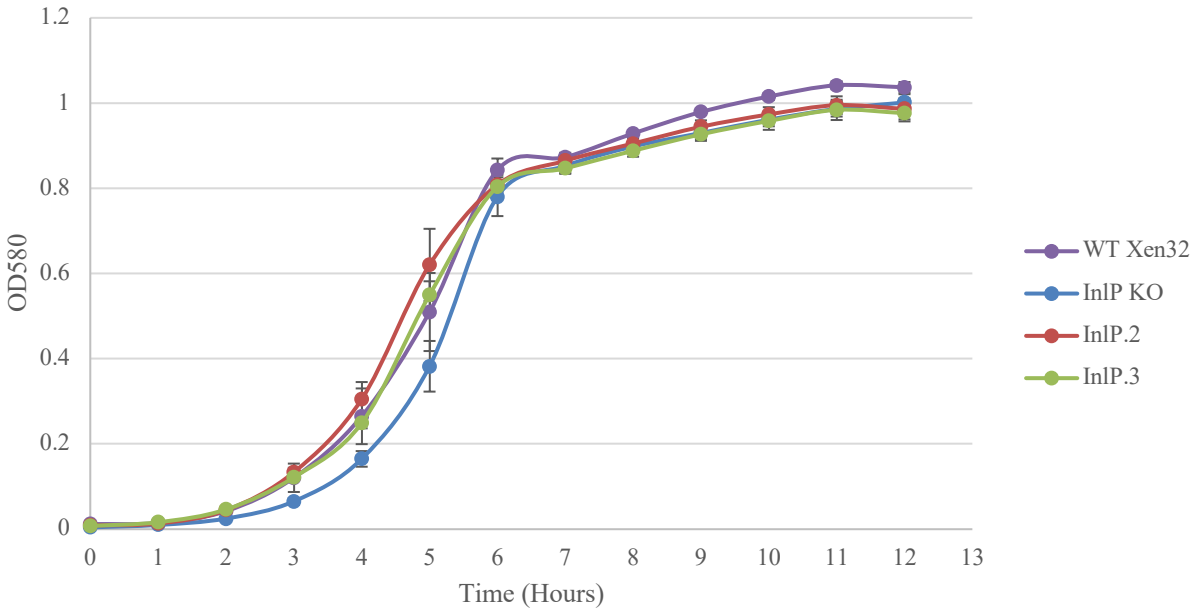


Figure 4.10 Open reading frame prediction for truncated InlP variants. To further investigate the effect of the start codon point mutation harbored by 14 of our isolates, we submitted a representative *inlP* gene sequence from one of these isolates (Accession: CP019622) (A) to the NCBI open reading frame finder tool and compared the putative ORFs with those from the *inlP* gene region in *Lm* 10403S (B).

A.

B.

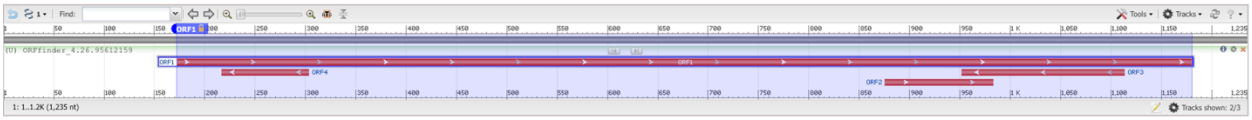


Figure 4.11 Lysis of *E. coli* DH5α transformed with pKSV7x-InlP.2. During transconjugation experiments, we noted that while *E. coli* DH5α readily grew at 37ºC with pKSV7x-InlP.3 as well as pKSV7x-InlP-Knockout, there appeared to be partial lysis of this strain when transformed with pKSV7x-InlP.2. This was observed when the strain was grown at 37ºC, but was rescued by growing the strain at 30ºC.

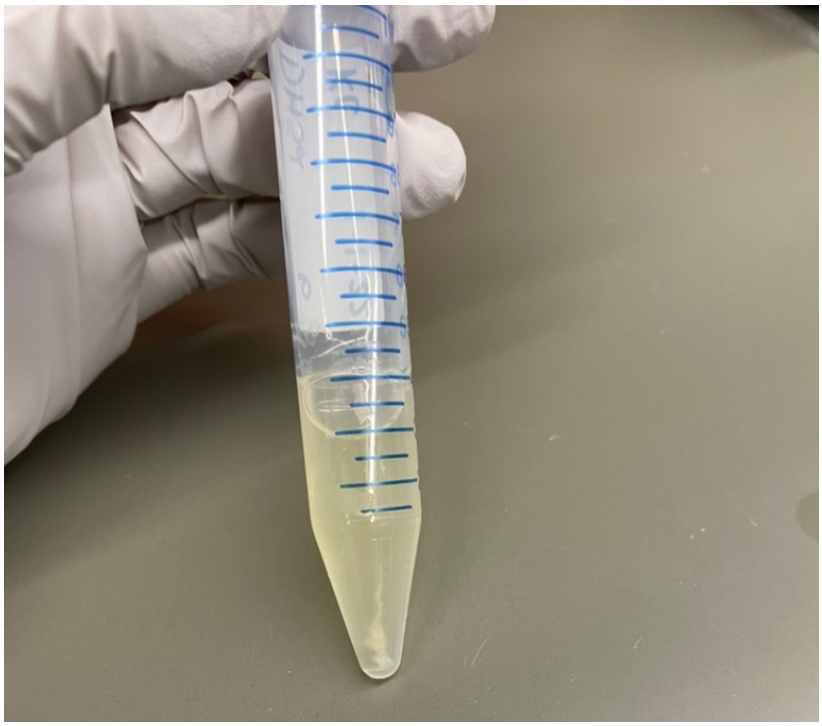


Table 4.1 Sequences used in this study

Note: Isolates harboring a truncated InlP variant are listed in **red and bold** text.

Strain	Serovar	GenBank Accession	Source
LIS0087	4b	CP044429	Environmental
81-0558	4b	CP007525	Clinical
81-0592	4b	CP007526	Clinical
10-0809	4b	CP007167	Clinical
10-0810	**1/2b**	**CP007168**	**Clinical**
1/2b	**1/2b**	**CP007169**	**Food**
1/2a	1/2a	CP007170	Food
10-0813	1/2a	CP007171	Clinical
10-4754	1/2a	CP007197	Clinical
1/2a	1/2a	CP007198	Food
10-0933	1/2a	CP007199	Clinical
1/2a	1/2a	CP007200	Food
02-1103	4b	CP007459	Clinical
02-1289	4b	CP007460	Clinical
02-1792	4b	CP007461	Food
02-6679	4b	CP008821	Clinical
02-6680	4b	CP007462	Food
95-0093	1/2a	CP007019	Clinical
98-2035	1/2a	CP007020	Clinical
99-6370	1/2a	CP007021	Clinical
02-5993	1/2a	CP007007	Clinical
04-5457	1/2a	CP007008	Clinical
1/2a	1/2a	CP007009	Food
08-7374	1/2a	CP007010	Food

Table 4.1 (Cont'd)

08-7669	1/2a	CP007011	Clinical
10-0814	1/2a	CP008836	Food
10-1046	1/2a	CP007017	Clinical
10-1321	1/2a	CP007018	Clinical
1/2c	1/2c	CP007194	Food
1/2c	1/2c	CP007195	Environmental
3c	3c	CP007196	Environmental
WSLC 1020	4a	CP013287	Veterinary
Lm 3163	1/2a	CP013722	Clinical
Lm 3136	1/2a	CP013723	Clinical
Lm N1546	1/2a	CP013724	Clinical
FSL-N1-304	1/2a	CP090052	Environmental
FSL-N1-334	1/2a	CP090054	Environmental
VIMVR081	4b	CP018148	Veterinary
VIMHA007	4b	CP018149	Clinical
10-092876-0168	**1/2b**	**CP019615**	**Food**
10-092876-1063 LM3	4b	CP019616	Food
10-092876-0055 LM4	1/2a	CP019617	Environmental
10-092876-0731 LM5	1/2a	CP019618	Environmental
10-092876-1155 LM6	4b	CP019619	Environmental
10-092876-1547 LM7	4b	CP019620	Environmental
10-092876-1235 LM8	1/2a	CP019621	Food
10-092876-0145 LM9	**1/2b**	**CP019622**	**Food**
10-092876-1763 LM10	1/2a	CP019623	Food
10-092876-1016 LM11	**1/2b**	**CP019624**	**Food**

Table 4.1 (Cont'd)

10-092876-0769 LM12	4b	CP019625	Environmental
H34	**1/2b**	**CP020774**	**Clinical**
MOD1_LS152	1/2a	CP020830	Clinical
PNUSAL000144	1/2a	CP020831	Clinical
FORC_049	**1/2b**	**CP016629**	**Food**
AT3E	1/2c	CP023752	Food
2018TE5305-1-4	1/2a	CP029372	Clinical
52869	1/2a	CP032669	Veterinary
52859	**1/2b**	**CP032671**	**Veterinary**
52854	4b	CP032672	Veterinary
R2-502	**1/2b**	**CP006594**	**Food**
C1-387	1/2a	CP006591	Food
J2-064	**1/2b**	**CP006592**	**Veterinary**
J2-1091	1/2a	CP006596	Veterinary
N1-011A	**1/2b**	**CP006597**	**Environmental**
J1817	4b	CP006599	Environmental
J1926	4b	CP006600	Food
PNUSAL000009	1/2a	CP054042	Clinical
BfR-LI-00752	**1/2b**	**CP054846**	**Food**
clinical isolate	4b	CP063382	Clinical
clinical isolate	4b	CP063383	Clinical
clinical isolate	1/2a	CP063381	Clinical
clinical isolate	1/2a	CP064843	Clinical
Lm60	1/2a	CP009258	Clinical
L2074	1/2a	CP007689	Clinical

Table 4.1 (Cont'd)

L1846	1/2b	CP007688	Clinical
L2625	1/2a	CP007687	Clinical
L2624	**1/2b**	**CP007686**	**Clinical**
L2676	1/2a	CP007685	Clinical
L2626	1/2a	CP007684	Clinical
EGD-e	1/2a	AL591824	Laboratory
EGD-e	1/2a	CP023861	Veterinary
FSL J1-175	**1/2b**	**CP062129**	**Environmental**
WSLC 1033	4d	CP013288	Veterinary
08-6569	1/2a	CP006858	Food
08-6997	1/2a	CP006859	Clinical
10-0815	1/2a	CP006860	Food
J1-220	4b	CP006046	Clinical
HPB913	1/2a	CP018685	Food
10-1047	1/2a	CP006861	Clinical
88-0478	1/2a	CP006862	Clinical
81-0861	4b	CP006874	Food
J1816	4b	CP006047	Food
FDAARGOS_57	**1/2a**	**CP030101**	**Environmental**
10403S	1/2a	CP002002	Laboratory

Table 4.2 Primers used in this study

Primer Set Name	Purpose	Primer Sequence
pKSV7x_Lin	Linearization of pKSV7x	F: 5'-GCTGCAGGAGGCAGTGGAG-3' R: 5'-GGATCCAGCGCCGCT-3'
InlP.2_Upst	Upstream Fragment for *inlP*.2 construction (mutation in red; Gibson Assembly overhangs underlined)	F: 5'-GGGTCCAGCGGCGCTGGATCCTGCGACTGGTAATTTAGAAGC-3' R: 5'-TGTGCTTAAAAACGTTAAAACTTTTCTCAA-3'
InlP.2_Dnst	Downstream Fragment for *inlP*.2 construction	F: 5'-ATTGAGAAAAGTTTTAACGTTTTTAAGCACA-3' R: 5'-GCTCGCTCCACTGCCTCCTGCAGCAATCGATATAAAATTTTAAGTGATATTATTAAAGCA-3'
InlP.3_Upst	Upstream Fragment for *inlP*.3 construction	F: 5'-GGGTCCAGCGGCGCTGGATCCTGCGACTGGTAATTTAGAAGC-3' R: 5'-ATCTTCCAGGTGAATCAGCAAA-3'
InlP.3_Dnst	Downstream Fragment for *inlP*.3 construction (mutation in red; Gibson Assembly overhangs underlined)	F: 5'-TTTTGCTGATTCACCTGGAAGAT-3' R: 5'-GCTCGCTCCACTGCCTCCTGCAGCAATCGATATAAAATTTTAAGTGATATTATTAAAGCA-3'
InlP_KO_Upst	Upstream Fragment for in-frame *inlP* Deletion (Gibson Assembly overhangs underlined)	F: 5'-GGTCCAGCGGCGCTGGATCCCTGCGACTGGTAATTTAGAAGC-3' R: 5'-GAAATTAATTAATAGTTACAGCTTAAAAACATTAAAACTTTTC-3'
InlP_KO_Dnst	Downstream Fragment for in-frame *inlP* Deletion (Gibson Assembly overhangs underlined)	F: 5'-AAGTTTTAATGTTTTTAAGCTGTAACTATTAATTAATTTCTACTAAAAAAGCTGGA-3' R: 5'-GCTCCACTGCCTCCTGCAGCCGATATAAAATTTTAAGTGATATTATTAAAGCAGTGAAG-3'
InlP.2_Seq	Sequencing of *inlP*.2 Mutant	F: 5'-GTTTTTTCGTGTTATTCTTTAGACC-3' R: 5'-CAGAAGCAGCTTTTGCCTTC-3'
InlP.3_Seq	Sequencing of *inlP*.3 Mutant	F: 5'-GGAATTTGGGGCTAAACTAACG-3' R: 5'-CCAGTGAAATCAGGAATAGAACC-3'

REFERENCES

REFERENCES

[1] J.A. Vazquez-Boland, M. Kuhn, P. Berche, T. Chakraborty, G. Dominguez-Bernal, W. Goebel, B. Gonzalez-Zorn, J. Wehland, J. Kreft, *Listeria* Pathogenesis and Molecular Virulence Determinants, Clinical Microbiology Reviews. 14 (2001) 584–640. https://doi.org/10.1128/CMR.14.3.584-640.2001.

[2] J.R. Robbins, A.I. Bakardjiev, Pathogens and the placental fortress, Curr. Opin. Microbiol. 15 (2012) 36–43. https://doi.org/10.1016/j.mib.2011.11.006.

[3] *Listeria* During Pregnancy, American Pregnancy Association, n.d. https://americanpregnancy.org/healthy-pregnancy/pregnancy-concerns/listeria-during-pregnancy/.

[4] K. Ireton, R. Mortuza, G.C. Gyanwali, A. Gianfelice, M. Hussain, Role of internalin proteins in the pathogenesis of *Listeria monocytogenes*, Mol Microbiol. (2021) mmi.14836. https://doi.org/10.1111/mmi.14836.

[5] J. Mengaud, H. Ohayon, P. Gounon, R.-M. Mège, P. Cossart, E-Cadherin Is the Receptor for Internalin, a Surface Protein Required for Entry of *L. monocytogenes* into Epithelial Cells, Cell. 84 (1996) 923–932. https://doi.org/10.1016/S0092-8674(00)81070-3.

[6] Y. Shen, M. Naujokas, M. Park, K. Ireton, InlB-Dependent Internalization of *Listeria* Is Mediated by the Met Receptor Tyrosine Kinase, Cell. 103 (2000) 501–510. https://doi.org/10.1016/S0092-8674(00)00141-0.

[7] M. Lecuit, H. Ohayon, L. Braun, J. Mengaud, P. Cossart, Internalin of *Listeria monocytogenes* with an intact leucine-rich repeat region is sufficient to promote internalization, Infect Immun. 65 (1997) 5309–5319. https://doi.org/10.1128/iai.65.12.5309-5319.1997.

[8] M.M. Maury, Y.-H. Tsai, C. Charlier, M. Touchon, V. Chenal-Francisque, A. Leclercq, A. Criscuolo, C. Gaultier, S. Roussel, A. Brisabois, O. Disson, E.P.C. Rocha, S. Brisse, M. Lecuit, Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity, Nat Genet. 48 (2016) 308–313. https://doi.org/10.1038/ng.3501.

[9] A. Moura, A. Criscuolo, H. Pouseele, M.M. Maury, A. Leclercq, C. Tarr, J.T. Björkman, T. Dallman, A. Reimer, V. Enouf, E. Larsonneur, H. Carleton, H. Bracq-Dieye, L.S. Katz, L. Jones, M. Touchon, M. Tourdjman, M. Walker, S. Stroika, T. Cantinelli, V. Chenal-Francisque, Z. Kucerova, E.P.C. Rocha, C. Nadon, K. Grant, E.M. Nielsen, B. Pot, P. Gerner-Smidt, M. Lecuit, S. Brisse, Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*, Nat Microbiol. 2 (2017) 16185. https://doi.org/10.1038/nmicrobiol.2016.185.

[10] *Listeria* (Listeriosis), Centers for Disease Control and Prevention, n.d. https://www.cdc.gov/listeria/index.html (accessed March 6, 2022).

[11] J.A. Vázquez-Boland, M. Wagner, M. Scortti, Why Are Some *Listeria monocytogenes* Genotypes More Likely To Cause Invasive (Brain, Placental) Infection?, MBio. 11 (2020) e03126-20. https://doi.org/10.1128/mBio.03126-20.

[12] S. Kathariou, P. Evans, V. Dutta, Strain-Specific Virulence Differences in *Listeria monocytogenes*: Current Perspectives in Addressing an Old and Vexing Issue, in: J.B. Gurtler, M.P. Doyle, J.L. Kornacki (Eds.), Foodborne Pathogens, Springer International Publishing, Cham, 2017: pp. 61–92. https://doi.org/10.1007/978-3-319-56836-2_3.

[13] K.K. Nightingale, K. Windham, K.E. Martin, M. Yeung, M. Wiedmann, 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. 71 (2005) 8764–8772. https://doi.org/10.1128/AEM.71.12.8764-8772.2005.

[14] A. Holch, H. Ingmer, T.R. Licht, L. Gram, *Listeria monocytogenes* strains encoding premature stop codons in inlA invade mice and guinea pig fetuses in orally dosed dams, J Med Microbiol. 62 (2013) 1799–1806. https://doi.org/10.1099/jmm.0.057505-0.

[15] K. Sobyanin, E. Sysolyatina, M. Krivozubov, Y. Chalenko, A. Karyagina, S. Ermolaeva, Naturally occurring InlB variants that support intragastric *Listeria monocytogenes* infection in mice, FEMS Microbiology Letters. (2017) fnx011. https://doi.org/10.1093/femsle/fnx011.

[16] C. Faralla, G.A. Rizzuto, D.E. Lowe, B. Kim, C. Cooke, L.R. Shiow, A.I. Bakardjiev, InlP, a New Virulence Factor with Strong Placental Tropism, Infect. Immun. 84 (2016) 3584–3596. https://doi.org/10.1128/IAI.00625-16.

[17] C. Faralla, E.E. Bastounis, F.E. Ortega, S.H. Light, G. Rizzuto, L. Gao, D.K. Marciano, S. Nocadello, W.F. Anderson, J.R. Robbins, J.A. Theriot, A.I. Bakardjiev, *Listeria monocytogenes* InlP interacts with afadin and facilitates basement membrane crossing, PLoS Pathog. 14 (2018) e1007094. https://doi.org/10.1371/journal.ppat.1007094.

[18] K. Clark, I. Karsch-Mizrachi, D.J. Lipman, J. Ostell, E.W. Sayers, GenBank, Nucleic Acids Res. 44 (2016) D67–D72. https://doi.org/10.1093/nar/gkv1276.

[19] J.J. Davis, A.R. Wattam, R.K. Aziz, T. Brettin, R. Butler, R.M. Butler, P. Chlenski, N. Conrad, A. Dickerman, E.M. Dietrich, J.L. Gabbard, S. Gerdes, A. Guard, R.W. Kenyon, D. Machi, C. Mao, D. Murphy-Olson, M. Nguyen, E.K. Nordberg, G.J. Olsen, R.D. Olson, J.C. Overbeek, R. Overbeek, B. Parrello, G.D. Pusch, M. Shukla, C. Thomas, M. VanOeffelen, V. Vonstein, A.S. Warren, F. Xia, D. Xie, H. Yoo, R. Stevens, The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities, Nucleic Acids Res. 48 (2020) D606–D612. https://doi.org/10.1093/nar/gkz943.

[20] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool, Journal of Molecular Biology. 215 (1990) 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

[21] S. Duvaud, C. Gabella, F. Lisacek, H. Stockinger, V. Ioannidis, C. Durinx, Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users, Nucleic Acids Research. 49 (2021) W216–W227. https://doi.org/10.1093/nar/gkab225.

[22] T. Lassmann, E.L. Sonnhammer, Kalign – an accurate and fast multiple sequence alignment algorithm, BMC Bioinformatics. 6 (2005) 298. https://doi.org/10.1186/1471-2105-6-298.

[23] A.M. Waterhouse, J.B. Procter, D.M.A. Martin, M. Clamp, G.J. Barton, Jalview Version 2--a multiple sequence alignment editor and analysis workbench, Bioinformatics. 25 (2009) 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

[24] F. Madeira, M. Pearce, A.R.N. Tivey, P. Basutkar, J. Lee, O. Edbali, N. Madhusoodanan, A. Kolesnikov, R. Lopez, Search and sequence analysis tools services from EMBL-EBI in 2022, Nucleic Acids Research. (2022) gkac240. https://doi.org/10.1093/nar/gkac240.

[25] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation, Bioinformatics. 23 (2007) 127–128. https://doi.org/10.1093/bioinformatics/btl529.

[26] D.L. Wheeler, D.M. Church, S. Federhen, A.E. Lash, T.L. Madden, J.U. Pontius, G.D. Schuler, L.M. Schriml, E. Sequeira, T.A. Tatusova, L. Wagner, Database resources of the National Center for Biotechnology, Nucleic Acids Res. 31 (2003) 28–33. https://doi.org/10.1093/nar/gkg033.

[27] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F.T. Heer, T.A.P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, Nucleic Acids Research. 46 (2018) W296–W303. https://doi.org/10.1093/nar/gky427.

[28] H.M. Berman, The Protein Data Bank, Nucleic Acids Research. 28 (2000) 235–242. https://doi.org/10.1093/nar/28.1.235.

[29] E.F. Pettersen, T.D. Goddard, C.C. Huang, E.C. Meng, G.S. Couch, T.I. Croll, J.H. Morris, T.E. Ferrin, UCSF ChimeraX: Structure visualization for researchers, educators, and developers, Protein Science. 30 (2021) 70–82. https://doi.org/10.1002/pro.3943.

[30] C. Bécavin, C. Bouchier, P. Lechat, C. Archambaud, S. Creno, E. Gouin, Z. Wu, A. Kühbacher, S. Brisse, M.G. Pucciarelli, F. García-del Portillo, T. Hain, D.A. Portnoy, T. Chakraborty, M. Lecuit, J. Pizarro-Cerdá, I. Moszer, H. Bierne, P. Cossart, Comparison of Widely Used *Listeria monocytogenes* Strains EGD, 10403S, and EGD-e Highlights Genomic

Differences Underlying Variations in Pathogenicity, MBio. 5 (2014) e00969-14. https://doi.org/10.1128/mBio.00969-14.

[31] A. Camilli, L.G. Tilney, D.A. Portnoy, Dual roles of plcA in *Listeria monocytogenes* pathogenesis, Mol Microbiol. 8 (1993) 143–157. https://doi.org/10.1111/j.1365-2958.1993.tb01211.x.

[32] J. Thomas, N. Govender, K.M. McCarthy, L.K. Erasmus, T.J. Doyle, M. Allam, A. Ismail, N. Ramalwa, P. Sekwadi, G. Ntshoe, A. Shonhiwa, V. Essel, N. Tau, S. Smouse, H.M. Ngomane, B. Disenyeng, N.A. Page, N.P. Govender, A.G. Duse, R. Stewart, T. Thomas, D. Mahoney, M. Tourdjman, O. Disson, P. Thouvenot, M.M. Maury, A. Leclercq, M. Lecuit, A.M. Smith, L.H. Blumberg, Outbreak of Listeriosis in South Africa Associated with Processed Meat, N Engl J Med. 382 (2020) 632–643. https://doi.org/10.1056/NEJMoa1907462.

[33] K.N. Conner, J.T. Burke, J.W. Hardy, J. Ravi, Novel Internalin P homologs in *Listeria*, Microbiology, 2022. https://doi.org/10.1101/2022.01.19.476994.

[34] G.E. Crooks, G. Hon, J.-M. Chandonia, S.E. Brenner, WebLogo: a sequence logo generator, Genome Res. 14 (2004) 1188–1190. https://doi.org/10.1101/gr.849004.

CHAPTER 5

CONCLUSIONS AND FUTURE DIRECTIONS

Adverse pregnancy outcomes resulting from placental infection remain a major public health issue. Specifically, preterm birth is of concern due to its contribution to developmental delays and abnormalities. In the United States, the preterm birth rate in 2019 was 10.2% [1]. It is estimated that preterm birth and low birth weight contributed to approximately 17% of infant deaths in the same year [1]. Placental infection is one of many factors that can induce preterm birth, and several bacterial, viral, and parasitic pathogens are able to colonize this important organ [2,3]. While *Listeria monocytogenes* (*Lm*) has long been recognized as a placental pathogen able to induce preterm labor and other adverse pregnancy outcomes, the mechanisms driving these outcomes remain unclear.

This dissertation addresses knowledge gaps in the field of placental listeriosis and listeriosis-induced preterm labor. In chapter 2, I outlined our analysis on the placental transcriptome and eicosanome following infection with *Lm* in a pregnant mouse model. Through this study, we concluded that placental gene expression is altered due to infection and identified gene expression signatures associated with placental listeriosis. Most interestingly, we identified an enrichment in genes associated with the eicosanoid pathway. Due to this pathway's known critical functions in inflammation and temporal regulation of labor, we measured the concentrations of various eicosanoids in infected and uninfected placentas. Infected placentas showed significant increases in concentrations of leukotriene B_4 (LTB_4), lipoxin A_4 (LXA_4), prostaglandin A_2 (PGA_2), prostaglandin D_2 (PGD_2), and eicosatrienoic acid. Previous studies have established associations between placental pathology and increased LTB_4, LXA_4, and PGD_2 [4–6]. To our knowledge, this is the first study identifying an association between increased PGA_2 levels and placental infection or preterm labor. This was an intriguing finding, as PGA_2 is a known degradation product of the less-stable eicosanoid prostaglandin E_2 (PGE_2). PGE_2 is a known

temporal mediator of labor and has been used clinically under the name Dinoprostone to induce labor at term. Increased concentrations of PGA_2 suggest increased production of PGE_2 upstream, which could contribute to the preterm labor phenotype observed in our pregnant mouse model.

Further experimentation could provide more mechanistic detail for the dysregulation of eicosanoids in placental listeriosis. Prior to the COVID-19 pandemic, we established experimental goals to begin delineating these details. Completing these goals proved to be challenging due to the stay-at-home order, and it was further complicated by supply chain disruptions and cell line contamination issues upon return to the lab. Given the opportunity, I would have liked to have completed the additional experimental objectives: 1) flow cytometry analysis of immune cell populations within infected and uninfected placentas and 2) analysis of cytosolic phospholipase A_2 ($cPLA_2$) levels and activity in infected and uninfected placentas.

Completing flow cytometry analysis of cell populations in infected and control placentas would allow for more comprehensive characterization of placental inflammation following infection. Increased eicosanoid concentrations within infected placentas suggests an inflammatory state of the organ, but the cell types responsible for this eicosanoid production remain unknown. Nearly all cells in the body, including immune cells, produce eicosanoids [7]. It is possible that the immune response to *Lm* within the placenta results in secretion of eicosanoids by immune cells, particularly macrophages, neutrophils, and natural killer cells, thus driving a pro-inflammatory state. It is also possible that resident placenta cells, like trophoblasts, contribute to eicosanoid production in response to *Lm* invasion.

Another possible explanation for the increase in placental eicosanoid levels is an overall increase in expression and/or activity of $cPLA_2$ within the organ following infection. This enzyme is responsible for the initial liberation of free arachidonic acid from phospholipid bilayers, and

thus, is key in driving the eicosanoid pathway [8]. A previous study by Noor *et al.* found that *Lm* infection of peritoneal macrophages induces arachidonic acid release [9]. This effect is due to induction of $cPLA_2$, and can be blocked using the specific $cPLA_2$ inhibitor, pyrrolidine [9]. This effect was partially dependent upon the *Lm* virulence factor listeriolysin O [9]. Therefore, we hypothesize that specific interactions between *Lm* and host cells could drive increased expression or activity of $cPLA_2$ within the placenta. Our RNAseq analysis (chapter 2) did not indicate a significant over- or underexpression of genes associated with $cPLA_2$ subtypes, but it is known that $cPLA_2$ activity is heavily mediated by phosphorylation and calcium binding [8]. Prior to the COVID-19 pandemic, we began optimizing a western blot protocol to assess $cPLA_2$ and phospho-$cPLA_2$ levels in mouse placentas. Our preliminary data for this was largely inconclusive and suggested that further optimization of our protocol will be required. Further investigation is needed to determine if *Lm* drives $cPLA_2$ activity in the placenta as it does in peritoneal macrophages.

The COVID-19 pandemic offered unique challenges for wet lab experiments. The inability to work in-person hindered my ability to carry out the aims set forth in my prelim proposal. Due to this, we shifted focus away from the proposed experiments and toward bioinformatics-based approaches to characterize an important *Lm* virulence factor for placental colonization, Internalin P (InlP). Prior to our studies, only two published studies of InlP existed. In 2016, Faralla *et al*. described a newly identified *Lm* virulence factor conferring placental tropism [10]. In 2018, they followed with structural and functional characterization of InlP [11]. Still, many questions remained regarding the regulation, function, and evolutionary conservation of InlP.

In chapter 3, I outline our study in which we identified InlP homologs in other *Listeria* species apart from *Lm*. The study described in this chapter uses the novel web-app, MolEvolvR, developed by the Ravi lab for studies of molecular evolution and phylogeny [12]. We used this

novel approach to identify InlP homologs across the domains of life. Of particular interest were InlP homologs and InlP-like homologs in other *Listeria* species: *L. ivanovii londoniensis*, *L. seeligeri*, *L. innocua*, and *L. costaricensis*. We noted that *L. seeligeri* encoded three copies of InlP homologs subsequently in its genome, and to our knowledge, is the only *Listeria* species to do so. The newly identified InlP homologs lack the full-length leucine rich repeat 6 (LRR6) and 7 (LRR7) domains present in *Lm* InlP. LRRs are known sites of protein-protein interactions, and the site of interaction between InlP and its host binding partner(s) is hypothesized to lie within its nine LRR regions [13,14]. The specific binding site for InlP and its only identified binding partner, afadin, remains to be identified.

This study of InlP homologs serves as a springboard for future studies of InlP function and evolution. More broadly, similar future studies may begin answering long-asked questions regarding the evolution of the internalin family of proteins. While previous studies identified afadin, a eukaryotic cell-cell junction protein that promotes cellular adhesion, they did not identify the site of InlP-afadin interaction [11]. We have begun generating mutants of our *Lm* laboratory strains with in-frame substitutions of identified InlP homologs from *L. ivanovii londoniensis* and *L. seeligeri* in place of the endogenous InlP. Future experiments will determine if A) these mutants are able to colonize the placenta as well as wild-type *Lm* and B) these homologs are able to bind afadin and/or other eukaryotic binding partners.

Our study establishes MolEvolvR as a useful tool for identifying internalin homologs within *Listeria* and across the domains of life. While we focused on InlP homologs in other *Listeria* species, our MolEvolvR search returned homologs across the domains of life. Of particular interest were numerous homologous hits in cyanobacteria. Future studies will analyze the conserved domains between InlP and identified homologs. Additionally, this analysis will be expanded to

include other internalins from *Listeria* to answer questions about their evolution – did the internalins originate elsewhere? What are the major structural differences between the internalins? Have duplication events contributed to the expansion of the internalin family within *Listeria*?

While there are certainly compelling questions remaining to be answered regarding the internalin family, there are additional questions surrounding their core LRR domains. LRRs are conserved across the domains of life, and are found in numerous organisms including humans, *Drosophila*, *Saccharomyces*, *Yersinia*, and of course, *Listeria* [13,14]. The LRR motif is characterized by variable 20-30 amino acid stretches that are rich in leucine [15]. LRRs are known to aid in forming protein-protein interactions [13,14]. There have been few published studies focusing on the evolution of the LRR motif, and information regarding its evolutionary lineage is lacking.

In chapter 4, I continued discussion of InlP; specifically, I focused on naturally occurring InlP variants across *Lm* isolates. This study used computational approaches followed by experimental validation. By analyzing 95 publicly available whole genome sequences (WGS) of *Lm*, we were able to identify two InlP variants of interest. The first, denoted as InlP.2, results a point mutation in the *inlP* start codon. The computationally predicted InlP.2 protein product is truncated due to a frameshift, resulting in loss of the InlP signal peptide. Interestingly, analysis of available metadata indicated a significant association of this variant with the *Lm* serovar 1/2b. The second variant of interest, denoted InlP.3, was present in approximately half of the isolates we analyzed and is defined by a proline to serine substitution within the InlP calcium binding loop.

Our discovery of these two variants leads to questions regarding their effect on virulence. Previous studies have identified variants of the better characterized internalins, InlA and InlB, and have established that these variants can contribute to *Lm* virulence [16–18]. Future experiments

will address the function and structure of InlP.2 and InlP.3. Because the InlP.2 variant is predicted to be truncated and lacking the InlP signal peptide, we predict that it will exhibit a defect in secretion to the extracellular space. The secretion pathway utilized by InlP remains unidentified, and its identification would allow for more specific mechanistic predictions of InlP.2 defects. To begin addressing this question, we attempted mass spectrometry analysis of the *Lm* secretome in hopes of quantifying InlP concentrations under various growth conditions. While we were able to detect other secreted internalins in our preparations, we did not detect InlP (data not shown). It is possible that InlP is produced below the limit of detection or that this protein is not produced *in vitro* at all.

The InlP.3 variant is of interest due to its association with the InlP calcium binding loop. This motif has been predicted to play a role in InlP protein-protein interactions and/or activation [11]. Proline offers structural rigidity to proteins, which impacts interactions with ligands (like Ca^{2+}) and other proteins [19]. Additionally, changes in hydrophobicity and charge must be considered when making predictions about the effects of this substitution. While many of these considerations are beyond the scope of our laboratory, future collaborations with structural biochemists will provide insights to the functional consequences of this substitution. Additionally, future studies should include isothermal titration calorimetry on the purified wild-type InlP and InlP.2 protein products to assess calcium binding ability.

Faralla *et al*. hypothesized that InlP contributes to placental colonization by enhancing *Lm* ability to transcytose through the layers of the placenta [11]. They utilized *in vitro* transcytosis assays to assess the transcytosis ability of wild-type *Lm* and a Δ*inlP* mutant of *Lm* in epithelial monolayers. We have begun optimizing this assay in the Hardy laboratory to allow for future

assessment of the mutants we have generated. Continued optimization is required to determine the appropriate multiplicity of infection and infection duration for this assay.

One final, but major, question regarding InlP surrounds its regulation. The master virulence regulator in *Lm*, PrfA, controls transcription of several virulence factors, including partial regulation of InlA, InlB, and InlC [20–22]. The promoter for InlP remains unidentified, adding to difficulty in characterizing its regulation. Computational promoter prediction suggests that InlP does not lie within an operon, and the top predicted promoters are not predicted to be PrfA-dependent (data not shown). Future experiments should focus on addressing InlP regulation and assess its expression in the Δ*prfA* and *prfA** strains of *Lm*. Additionally, electrophoresis mobility shift assays could address PrfA binding at the *inlP* promoter more directly.

This dissertation has addressed several knowledge gaps in the field of placental infection by *Lm*. To our knowledge, our study outlined in chapter 2 is the first study associating increased placental eicosanoid concentrations with *Lm* infection. Additionally, it is the first study assessing the gene expression signatures of infected placentas in a mouse model. Our large RNAseq data set offers a starting point for other investigators in the field to continue addressing placental responses to listeriosis. Our studies of InlP are impactful due to the lack of available information regarding InlP. Prior to our study identifying InlP homologs, only two other published studies of InlP existed. We hope to add to this pool of knowledge once again following further characterization of InlP.2 and InlP.3. Finally, we have generated and validated tools for the study of InlP and other internalins (mutant strains; MolEvolvR) which will be pivotal in their continued investigation.

REFERENCES

REFERENCES

[1] Preterm Birth, Centers for Disease Control and Prevention, n.d. https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pretermbirth.htm (accessed June 2, 2022).

[2] J.R. Robbins, A.I. Bakardjiev, Pathogens and the placental fortress, Curr. Opin. Microbiol. 15 (2012) 36–43. https://doi.org/10.1016/j.mib.2011.11.006.

[3] R.L. Goldenberg, J.F. Culhane, J.D. Iams, R. Romero, Epidemiology and causes of preterm birth, Lancet. 371 (2008) 75–84. https://doi.org/10.1016/S0140-6736(08)60074-4.

[4] A. López Bernal, D.J. Hansell, T.Y. Khong, J.W. Keeling, A.C. Turnbull, Placental leukotriene B4 release in early pregnancy and in term and preterm labour, Early Human Development. 23 (1990) 93–99. https://doi.org/10.1016/0378-3782(90)90132-3.

[5] L.O. Perucci, P.C. Santos, L.S. Ribeiro, D.G. Souza, K.B. Gomes, L.M.S. Dusse, L.P. Sousa, Lipoxin A4 Is Increased in the Plasma of Preeclamptic Women, AJHYPE. 29 (2016) 1179–1185. https://doi.org/10.1093/ajh/hpw053.

[6] S. Kumar, T. Palaia, C.E. Hall, L. Ragolia, Role of Lipocalin-type prostaglandin D2 synthase (L-PGDS) and its metabolite, prostaglandin D2, in preterm birth, Prostaglandins & Other Lipid Mediators. 118–119 (2015) 28–33. https://doi.org/10.1016/j.prostaglandins.2015.04.009.

[7] E. Ricciotti, G.A. FitzGerald, Prostaglandins and Inflammation, Arterioscler Thromb Vasc Biol. 31 (2011) 986–1000. https://doi.org/10.1161/ATVBAHA.110.207449.

[8] A.M. Vasquez, V.D. Mouchlis, E.A. Dennis, Review of four major distinct types of human phospholipase A2, Advances in Biological Regulation. 67 (2018) 212–218. https://doi.org/10.1016/j.jbior.2017.10.009.

[9] S. Noor, H. Goldfine, D.E. Tucker, S. Suram, L.L. Lenz, S. Akira, S. Uematsu, M. Girotti, J.V. Bonventre, K. Breuel, D.L. Williams, C.C. Leslie, Activation of Cytosolic Phospholipase A $_2$ α in Resident Peritoneal Macrophages by *Listeria monocytogenes* Involves Listeriolysin O and TLR2, J. Biol. Chem. 283 (2008) 4744–4755. https://doi.org/10.1074/jbc.M709956200.

[10] C. Faralla, G.A. Rizzuto, D.E. Lowe, B. Kim, C. Cooke, L.R. Shiow, A.I. Bakardjiev, InlP, a New Virulence Factor with Strong Placental Tropism, Infect. Immun. 84 (2016) 3584–3596. https://doi.org/10.1128/IAI.00625-16.

[11] C. Faralla, E.E. Bastounis, F.E. Ortega, S.H. Light, G. Rizzuto, L. Gao, D.K. Marciano, S. Nocadello, W.F. Anderson, J.R. Robbins, J.A. Theriot, A.I. Bakardjiev, *Listeria monocytogenes* InlP interacts with afadin and facilitates basement membrane crossing, PLoS Pathog. 14 (2018) e1007094. https://doi.org/10.1371/journal.ppat.1007094.

[12] J.T. Burke, S.Z. Chen, L.M. Sosinski, J.B. Johnston, J. Ravi, MolEvolvR: A web-app for characterizing proteins using molecular evolution and phylogeny, BioRxiv. (2022). https://doi.org/10.1101/2022.02.18.461833.

[13] B. Kobe, The leucine-rich repeat as a protein recognition motif, Current Opinion in Structural Biology. 11 (2001) 725–732. https://doi.org/10.1016/S0959-440X(01)00266-4.

[14] B. Kobe, J. Deisenhofer, The leucine-rich repeat: a versatile binding motif, Trends in Biochemical Sciences. 19 (1994) 415–421. https://doi.org/10.1016/0968-0004(94)90090-6.

[15] J. Bella, K.L. Hindle, P.A. McEwan, S.C. Lovell, The leucine-rich repeat structure, Cell Mol Life Sci. 65 (2008) 2307–2333. https://doi.org/10.1007/s00018-008-8019-0.

[16] K.K. Nightingale, K. Windham, K.E. Martin, M. Yeung, M. Wiedmann, 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. 71 (2005) 8764–8772. https://doi.org/10.1128/AEM.71.12.8764-8772.2005.

[17] A. Holch, H. Ingmer, T.R. Licht, L. Gram, *Listeria monocytogenes* strains encoding premature stop codons in inlA invade mice and guinea pig fetuses in orally dosed dams, J Med Microbiol. 62 (2013) 1799–1806. https://doi.org/10.1099/jmm.0.057505-0.

[18] K. Sobyanin, E. Sysolyatina, M. Krivozubov, Y. Chalenko, A. Karyagina, S. Ermolaeva, Naturally occurring InlB variants that support intragastric *Listeria monocytogenes* infection in mice, FEMS Microbiology Letters. (2017) fnx011. https://doi.org/10.1093/femsle/fnx011.

[19] F. Krieger, A. Möglich, T. Kiefhaber, Effect of Proline and Glycine Residues on Dynamics and Barriers of Loop Formation in Polypeptide Chains, J. Am. Chem. Soc. 127 (2005) 3346–3352. https://doi.org/10.1021/ja042798i.

[20] T. Chakraborty, M. Leimeister-Wächter, E. Domann, M. Hartl, W. Goebel, T. Nichterlein, S. Notermans, Coordinate regulation of virulence genes in *Listeria monocytogenes* requires the product of the prfA gene, J Bacteriol. 174 (1992) 568–574. https://doi.org/10.1128/jb.174.2.568-574.1992.

[21] A. Lingnau, E. Domann, M. Hudel, M. Bock, T. Nichterlein, J. Wehland, T. Chakraborty, Expression of the *Listeria monocytogenes* EGD inlA and inlB genes, whose products mediate bacterial entry into tissue culture cell lines, by PrfA-dependent and -independent mechanisms, Infect Immun. 63 (1995) 3896–3903. https://doi.org/10.1128/iai.63.10.3896-3903.1995.

[22] Q. Luo, M. Rauch, A.K. Marr, S. Müller-Altrock, W. Goebel, In vitro transcription of the *Listeria monocytogenes* virulence genes inlC and mpl reveals overlapping PrfA-dependent and -independent promoters that are differentially activated by GTP, Mol Microbiol. 52 (2004) 39–52. https://doi.org/10.1111/j.1365-2958.2003.03960.