How did *Listeria monocytogenes* become pathogenic?

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades der Bayerischen Julius-Maximilians-Universität Würzburg

vorgelegt von

Eva Yee Wah Ng

aus

Hong Kong

Würzburg, 2001

Die vorliegende Arbeit wurde von mir selbständig unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt.

This work was completed independently by myself using the described literature and materials.

Eva Y. W. Ng

Eingereicht am: .....

Gutachter der Dissertation:

1. Gutachter: Prof. Dr. Werner Goebel

2. Gutachter: Prof. Dr. Jörg Hacker

Doktorurkunde ausgehändigt am: .....

Cottleston, Cottleston, Cottleston Pie
Why does a chicken, I don't know why.
Ask me a riddle and I reply:
Cottleston, Cottleston, Cottleston Pie.

-Pooh
A. A. Milne,
Winnie-The-Pooh, 1926.

CONTENTS

Acknowledgements

To ALL and each in your own way, for your presence and your help in the completion of this work, I thank you.

In particular:

Prof. Werner Goebel who enabled and supported this research.

Jürgen Kreft and Ursula Rdest for their aid in solving myriad problems, large and small.

Michael Schmid and Michael Wagner at the Technische Universität München for making phylogenetic trees.

All who helped out in the *Listeria* sequencing project for bearing with me.

Volker Thiel and Prof. Stuart Siddell at Virology, Dag Harmsen and staff at Hygiene, Susanne Michel and Prof. Kaldenhoff at Botanik for graciously lending their help and their automated sequencers for the sequencing of the *Listeria* genome.

Lionel Franguel, Philippe Glaser and Frank Kunst at the Institute Pasteur, who welcomed me to their lab during the *Listeria* Genome Project.

Silke Gertz, Susanne Engelmann, and Prof. Michel Hecker of the University of Greifswald for hosting and teaching me to make beautiful 2D gels.

Fritz Titgemeyer at the University of Erlangen for the gift of the pWH1509 plasmid series.

Andrea Spory, Brigitte Kimmel and Dagmar Baier for helping me translate the abstract into Deutsch. They are responsible for the jokes.

Justin Daniels and Fredi Engelbrecht, Silke Hom, Birgit Bergmann, Michael Herler, Marcus Rauch, Steffi Altrock, Silke Weber, Carmen Dickneite, Jutte Bohne, Regine Böckmann, Mio Lalic, Marcus Beck, Jörg Schlaghuis, Andreas Bubert, Prof. Szalay and Prof. Roy Gross for discussions, fun and just being there.

My uncle G. P. Poon, without whose gift I would not have the computer needed for the thesis, the DNA analyses, the genome sequencing project and a whole lot of more.

and especially, to...

Marian Castro, Richie Knopp and Holm Knopp, for their care, joy and friendship.

How did *Listeria monocytogenes* become pathogenic?

Abstract

Listeriae are Gram positive, facultative, saprophytic bacteria capable of causing opportunistic infections in humans and animals. This thesis presents three separate lines of inquiries that can lead to the eventual convergence of a global view of *Listeri*a as pathogen in the light of evolution, genomics, and function.

First, we undertook to resolve the phylogeny of the genus *Listeria* with the goal of ascertaining insights into the evolution of pathogenic capability of its members. The phylogeny of *Listeriae* had not yet been clearly resolved due to a scarcity of phylogenetically informative characters within the 16S and 23S rRNA molecules. The genus *Listeria* contains six species: *L. monocytogenes, L. ivanovii, L. innocua, L. seeligeri, L. welshimeri*, and *L. grayi;* of these, *L. monocytogenes* and *L. ivanovii* are pathogenic. Pathogenicity is enabled by a 10-15Kb virulence gene cluster found in *L. seeligeri, L. monocytogenes* and *L. ivanovii*. The genetic contents of the virulence gene cluster *loci*, as well as some virulence-associated internalin *loci* were compared among the six species. Phylogenetic analysis based on a data set of nucleic acid sequences from *prs, ldh, vclA, vclB, iap,* 16S and 23S rRNA genes identified *L. grayi* as the ancestral branch of the genus. This is consistent with previous 16S and 23S rRNA findings. The remainder 5 species formed two groupings. One lineage represents *L. monocytogenes* and *L. innocua*, while the other contains *L. welshimeri, L. ivanovii* and *L. seeligeri*, with *L. welshimeri* forming the deepest branch within this group. Deletion breakpoints of the virulence gene cluster within *L. innocua* and *L. welshimeri* support the proposed tree. This implies that the virulence gene cluster was present in the common ancestor of *L. monocytogenes, L. innocua, L. ivanovii, L. seeligeri* and *L. welshimeri;* and that pathogenic capability has been lost in two separate events represented by *L. innocua* and *L. welshimeri*.

Second, we attempted to reconstitute *L. innocua* of its deleted virulence gene cluster, in its original chromosomal location, from the *L. monocytogenes* 12 Kb virulence gene cluster. This turned out particularly difficult because of the limits of genetic tools presently available for the organism. The reconstitution was partially successful. The methods and approaches are presented, and all the components necessary to complete the constructs are at hand for both *L. innocua* and the parallel, positive control of *L. monocytogenes* mutant deleted of its virulence gene cluster.

Third, the sequencing of the entire genome of *L. monocytogenes* EGDe was undertaken as part of an EU Consortium. Our lab was responsible for 10% of the labor intensive gap-closure and annotation efforts, which I helped coordinate. General information and comparisons with sister species *L. innocua* and a close Gram positive relative *Bacillus subtilis* are presented in context. The areas I personally investigated, namely, sigma factors and stationary phase functions, are also presented. *L. monocytogenes* and *L. innocua* both possess surprisingly few sigma factors: SigA, SigB, SigH, SigL, and an extra-cytoplasmic function type sigma factor (SigECF). The stationary phase genes of *L. monocytogenes* is compared to the well-studied, complex, stationary phase networks of *B. subtilis*. This showed that while genetic competence functions may be operative in unknown circumstances, non-sporulating *Listeria* opted for very different approaches of regulation from *B. subtilis*. There is virtually no overlap of known, stationary phase genes between *Listeria* and Gram negative model organism *E. coli*.

Zusammenfassung

Listerien sind Gram-positive, fakultativ intrazelluläre, saprophytische Bakterien, die in der Lage sind, bei Mensch und Tier opportunistische Infektionen hervorzurufen. Die vorliegende Arbeit veranschaulicht drei unterschiedliche Versuchsansätze, die schließlich hinsichtlich Evolution, Genom- und Funktionsanalysen zur Konvergenz der globalen Sichtweise von *Listeria* als pathogener Mikroorganismus führen können.

Zunächst wurden phylogenetische Analysen durchgeführt, die einen Einblick in die Evolution des pathogenen Potentials von Mitgliedern der Gattung *Listeria* geben sollten. Aufgrund mangelnder phylogenetischer Informationen bezüglich der 16S und 23S rRNAs war eine genaue phylogenetische Entschlüsselung der Gattung *Listeria* jedoch nicht möglich. Die Gattung *Listeria* umfaßt sechs verschiedene Spezies: *L. monocytogenes*, *L. ivanovii*, *L. innocua*, *L. seeligeri*, *L. welshimeri* und *L. grayi*, wobei *L. monocytogenes* und *L. ivanovii* zu den pathogenen Bakterien zählen. Die Pathogenität wird hierbei durch ein 10-15 kb großes Virulenzgencluster determiniert, das in *L. monocytogenes*, *L. ivanovii* sowie in *L. seeligeri* vorzufinden ist und neben einigen Virulenz-assoziierten Internalin-Genen zum genetischen Vergleich der sechs verschiedenen Spezies herangezogen wurde. Die im Rahmen der phylogenetischen Analysen untersuchten Nukleinsäuresequenzen der Gene *prs*, *ldh*, *vclA*, *vclB*, *iap* sowie die der 16S und 23S rRNA-Gene deuteten darauf hin, daß *L. grayi* dem gemeinsamen Vorläufer der Gattung *Listeria* am nächsten steht, was mit den bisher verfügbaren 16S und 23S rRNA-Daten übereinstimmt. Die verbleibenden fünf Spezies bilden zwei Gruppen, die sich zum einen aus *L. monocytogenes* und *L. innocua* und zum anderen aus *L. welshimeri*, *L. ivanovii* und *L. seeligeri* zusammensetzen. Die Positionen der im Virulenzgencluster von *L. innocua* und *L. welshimeri* identifizierten Deletionen bestätigen den hier vorgeschlagenen Stammbaum, was darauf hindeutet, daß im gemeinsamen Vorfahren von *L. monocytogenes*, *L. innocua*, *L. ivanovii*, *L. seeligeri* und *L. welshimeri* das

Virulenzgencluster vorhanden war und daß das pathogene Potential in *L. innocua* bzw. in *L. welshimeri* durch zwei unabhängige Ereignisse verloren ging.

Weiterhin wurde versucht, das 12 kb-Virulenzgencluster von *L. monocytogenes* in der ursprünglichen Lokalisation in das Chromosom der Spezies *L. innocua*, die eine Deletion des Virulenzgenclusters aufweist, zu integrieren. Dieser Ansatz erwies sich aufgrund der derzeit limitierten Methodik zur genetischen Manipultation dieses Organismus als sehr problematisch und führte bisher nur teilweise zum Erfolg. Die angewandten Strategien zur Klonierung der erforderlichen Konstrukte, die zur Erstellung der beschriebenen *L. innocua*-Mutante und einer *L. monocytogenes*-Mutante mit Deletion des Virulenzgenclusters erforderlich sind, werden vorgestellt.

Der dritte Schwerpunkt der Arbeit befaßte sich mit der vollständigen Sequenzierung des Genoms von *L. monocytogenes* EGDe, die im Rahmen eines EU-Konsortiums durchgeführt wurde. Unser Labor war zu 10 % an der Lückenschließung zwischen den Sequenz-Contigs sowie an der Annotierung beteiligt, für deren Koordinierung ich verantwortlich war. Vergleiche der Genomsequenzen von *L. monocytogenes*, *L. innocua* und dem verwandten Gram-positiven Bakterium *Bacillus subtilis* werden vorgestellt, wobei ein besonderer Schwerpunkt auf den Genen für Sigma-Faktoren und Stationärphase-Funktionen liegt. Sowohl *L. monocytogenes* und *L. innocua* enthalten mit SigA, SigB, SigH und einem extrazytoplasmatischen Sigma-Faktor, SigECF, überraschend wenige Sigma-Faktoren. Die Stationärphase-Gene von *L. monocytogenes* werden mit dem gut untersuchten, sehr komplexen Stationärphase-System von *B. subtilis* verglichen. Dies zeigte, daß, obwohl genetische Kompetenz unter nicht bekannten Umständen eine Rolle spielen könnten, in *Listeria* völlig unterschiedliche Mechanismen der Genregulation wirksam sind. Es ist keinerlei Überlappung mit den bekannten Stationärphase-Genen des Gram-negativen Modellorganismus *Escherichia coli* festzustellen.

Chapter 1

Introduction

Chapter 1. Introduction

Pathogenesis is an anthropomorphic concept. The only aim of 'pathogenic' bacteria, whether obligate or opportunistic, is to live. An opportunistic pathogen is an organism that occasionally or accidentally encounters another organism, whereby both parties incur collateral damage. The host simultaneously presents myriad threats as well as unexplored niches to the 'pathogen'. Virulence is a measure of the fitness cost to the host upon supporting and/or resisting the exogenous lodger. Therefore, virulence does not necessarily translate to bacterial success in conquering new and stable niches in a given host.

Two out of 6 bacterial species of the *Listeria* genus are such opportunistic pathogens. *L. monocytogenes* is a mammalian foodborne pathogen potentially lethal in humans and animals; it is a major concern for the food handling and processing industries. *L. ivanovii* often infects ungulates via silage, and is thus an economic concern for the livestock industries. The other four species of the genus are nonpathogenic. These *include L. innocua, L. seeligeri, L. welshimeri,* and *L. grayi*. All *Listeriae* are non-sporeforming, motile, rod shaped, facultative-anaerobic, Gram positive bacteria. The predominant reservoir of *Listeriae* is environmental and widespread. With the exception of clinical specimens, all *Listeriae* can be found in the same environments, namely soil, rotting vegetation, sewage, rivers, salt water estuaries, digestive tracts of healthy animals including humans and their feces (Jones 1992; Rocourt & Seeliger 1985).

Human exposure to *Listeria spp*. via ingestion is extremely common. Listeriosis is rare despite constant exposure, but occurrences tend to present as outbreaks traceable to manufactured foods. Their ability to grow between -0.4^{o}C to 50 ^{o}C increases their contamination risk in diary products, meats, seafood and other processed food products via selective enrichment during refrigeration. As in animals, listeriosis in humans, particularly neonates, the elderly and the immunocompromised, is an invasive disease characterized by bacteremia, including occasional central nervous system involvement leading to meningitis and death. The liver, central nervous system, and fetuses are the favored targets of these organisms, and this latter preference translates to high risk for involuntary abortions in pregnant carriers (Farber & Peterkin 1991; Lorber 1997). Thus, the immune status of the host largely determines the clinical outcome of the encounter with this pathogen. However, there is mounting evidence that when taken in high doses (10^{11} organisms in contaminated milk), *L. monocytogenes* can cause febrile gastroenteritis in healthy adults (Dalton et al 1997; Schlech 1997), which may lead to occasional cases of invasive disease. The extent to which these two clinical manifestations are due to strain differences or dosage effects is unknown. There is no documentation of the disease spreading from person to person.

Listerial pathogenesis is mainly attributed to the bacteria's ability to invade and grow intracellularly in mammalian hosts cells, including macrophages. Without any involvement of virulence factors, *Listeria spp.* can cross the intestinal barrier passively and reach deep tissue organs within 15 minutes in the rat ileal loop model. Enteritis associated intestinal lesions occur only with large doses of innoculum ($\geq 10^9$ CFU/loop) (Pron et al 1998). Uptake and replication in M cell associated macrophages (Siebers & Finlay 1996) and dendritic cells (Guzman et al 1995) may allow bacteria to travel via lymph or blood to deep tissue organs and establish systemic infection. The majority of bacteria concentrate and replicate in the liver where they are eventually killed by resident macrophages and migrating neutrophils (Cousens & Wing 2000; Ebe et al 1999). For those that evaded killing, they can grow exponentially and spread from cell to cell between macrophages and hepatocytes and among hepatocytes. Depending on the immune status of the host, bacterial proliferation may lead to further infection of other organs, especially fetal and nerve tissue, resulting in abortion or fatal disease (Cousens & Wing 2000).

Genes attributed to virulence phenotype

The virulence gene cluster genes:

Major virulence functions of pathogenic *L. monocytogenes* that allow its intracellular lifestyle in hosts have been localized to a cluster of six genes, 9-10 Kb in length, referred to as the virulence gene cluster. Three members of the *Listeria* genus possess some form of this virulence gene cluster. *L. ivanovii* has a very similar virulence gene cluster to that of *L. monocytogenes*, while nonpathogenic *L. seeligeri* contains a more elaborate cluster with five additional genes. The contents and evolution of the virulence cluster *loci* in each of the listerial species are discussed in detail in Chapter 2 (see Chp2. Fig.1).

The six characterized virulence genes of *L. monocytogenes* are primarily under the positive control of master regulator PrfA (see Figure 1.1). The *prfA* gene itself is included in this cluster and is therefore positively autoregulated via the promoter driving *plcA* and *prfA*. Disruption of *prfA* can lead to a 3 log decrease in virulence in mice (Camilli et al 1991; Camilli et al 1993). PrfA is a 27Kd DNA binding protein of the Crp-Fnr family (Leimeister-Wachter et al 1990) and (Lampidis, unpublished data) with specific affinity to a 14bp palindromic sequence (PrfA box, consensus: 5' TTAACANNTGTTAA 3') located at approximately -41bp relative to the transcriptional start site (Freitag & Portnoy 1994; Freitag et al 1993; Williams et al 2000). Presence of PrfA is necessary but insufficient to activate global virulence gene expression (Renzoni et al 1997, Klarsfeld, 1994 #45). The virulence phenotype appears to be co-regulated by other pathways such as catabolite repression (Milenbachs et al 1997) and stress (Sokolovic et al 1993). Environmental signals that influence virulence gene expression include medium composition (Bohne et al 1996; Ripio et al 1996), temperature (Leimeister-

Wachter et al 1992), stress conditions (Sokolovic et al 1993), carbon sources (Milenbachs et al 1997; Park & Kroll 1993), iron (Bockmann et al 1996), and contact with intact mammalian cells or mammalian cell extracts, or bacterial internalization into host cells (Renzoni et al 1999). Nothing is known about the environmental sensors and their signal relays that eventually activate PrfA. Although the exact factors involved are unknown, *in vitro* observations suggests that co-regulated PrfA-dependent transcription is mediated through an alternative component of RNA polymerase--possibly by an alternative sigma factor or a modifier of PrfA (Bockmann et al 2000; Dickneite et al 1998).

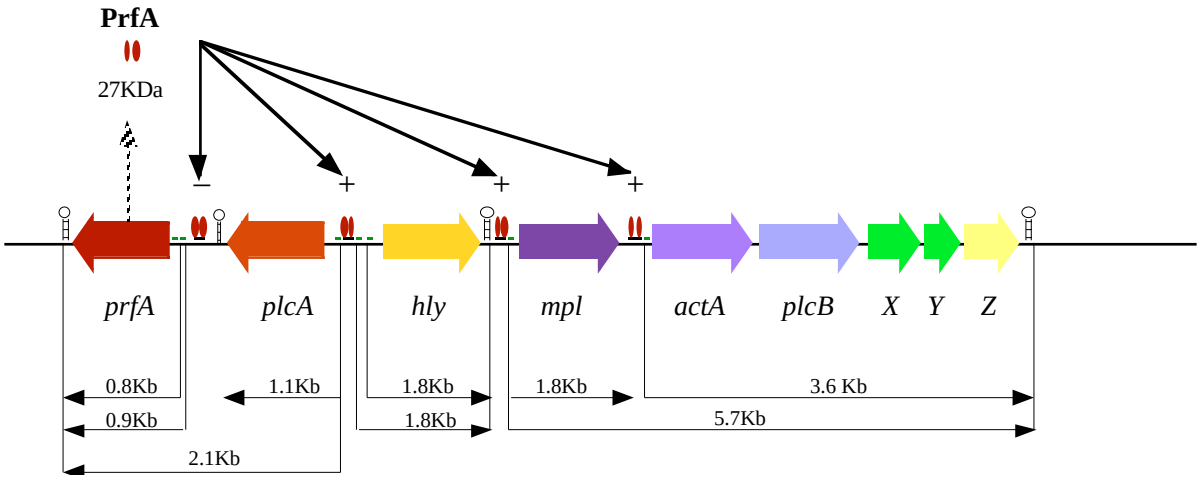


Figure 1.1. The virulence gene cluster of *L. monocytogenes* is under the control of PrfA. PrfA homodimers (red ovals) bind preferentially to the PrfA boxes located in the promoter regions of the virulence gene operons. Note the transcription start sites (green boxes), PrfA boxes (black boxes), transcriptional terminators (hairpins), and the resulting transcripts observed. The functions of open reading frames X, Y and Z have not been characterized. (Schematic diagram not to scale, after (Engelbrecht 1999)).

The gene product absolutely essential for the virulence phenotype is the sulfhydryl (SH)-activated, pore-forming hemolysin, or listeriolysin (LLO) encoded by *hly*. LLO facilitates bacterial escape from the primary phagosomes of host cells into the host cytosol before they are killed in the phagolysosomes (Bielecki et al 1990; Cossart et al 1989; Davies 1983). In addition, LLO mediates bacterial escape from the double membrane vacuole upon entry into the neighboring cell during cell to cell spread (Gedde et al 2000). Inactivation of *hly* results in a 5 log reduction of virulence in mice (Gaillard et al 1986). *Hly*$^-$ strains are nonpathogenic.

Together with LLO, two phospholipases facilitate the breaching of host cell membranes in the processes described above. *plcA* encodes a phosphatidylinositol-specific phospholipase C (Mengaud et al 1991a). *plcB* encodes a phosphatidylcholine phospholipase C with a broad spectrum of substrates (Goldfine et al 1993). PlcB requires a metalloprotease

encoded by *mpl* for proper maturation and activity (Raveneau et al 1992). While the disruption of either phospholipase only caused very slight reductions of virulence, *plcA-plcB* double mutants showed a 2.5 log increase in the 50% lethal dose (LD_{50}) in mice infected intravenously. Moreover, they were defective for primary phagosomal escape and impaired in cell to cell spread (Smith et al 1995).

ActA endows *Listeriae* with actin polymerization capability to drive intracellular movement within host cells. *ActA* deletion mutant showed a 3 log reduction of virulence in mice (Brundage et al 1993). Using actin polymerization at one pole of the bacterium, intracellular bacteria form extrusions from the infected cell to the neighboring cell. These projections are phagocytized by the neighbors, resulting in bacteria contained in first cell's membrane that are engulfed by the phagosome of the target cell (Tilney & Portnoy 1989). The dissolution of this double membrane requires the products of *hly, plcB, mpl* and to some extent, *plcA* (Gedde et al 2000; Marquis et al 1995; Smith et al 1995). Therefore, the gene products of the entire virulence gene cluster are required for the full efficiency of cell to cell spread.

The internalins:

Besides the virulence gene cluster, the internalins constitute a major virulence gene family whose members are found scattered in the genomes of *L. monocytogenes* (Engelbrecht et al 1996; Gaillard et al 1991; Lingnau et al 1995; Raffelsbauer et al 1998), and *L. ivanovii* (Engelbrecht et al 1998a; Engelbrecht et al 1998b). Some but not all internalins identified to date contribute to *Listeria*'s host cell tropism. The internalin genes encode extracellular proteins containing varying numbers (6-16) of characteristic, 22 amino acid long, Leucine rich repeats (LRRs). LRR proteins are wide spread in eukaryotes and prokaryotes and they mediate protein-protein interaction for diverse purposes. Bacterial LRR proteins are found in *Yersinia, Shigella, Salmonella,* and *Burkholderia,* but they differ slightly from the listerial specific motif (reviewed in (Kobe & Deisenhofer 1995; Marino et al 2000)).

The LRR regions alone, of either InlA or InlB, are necessary and sufficient to confer their respective invasive activities towards mammalian cells (Braun et al 1999; Lecuit et al 1997). The 1.86 A resolution X-ray crystal structure of the LRR region of InlB has been determined (Marino et al 1999), see Figure 2.2. Each LRR (xLxxLxLxxNxLxDIxxLxxLx) corresponds to a structural unit consisting of a β-strand and an α-helix. Conserved positions (L or N) form the core of the 22 amino acid repeat unit. These units stack onto each other in a right-handed manner. Asparagine residues at position 10 mediate the interaction with the above LRR; thus, conservation of the Asn10 in the first LRR is not necessary. In the first LRR, position 20 is always a hydrophobic residue and is thought to interact with the N-terminal sequence, which stabilizes it. The highly conserved, inter-repeat sequence is

expected to form an analogous cap to stabilize the last LRR. The outward facing positions (Fig 2.2B), particularly over the concave β-strand surface, are available for external protein-protein interactions (Marino et al 2000).

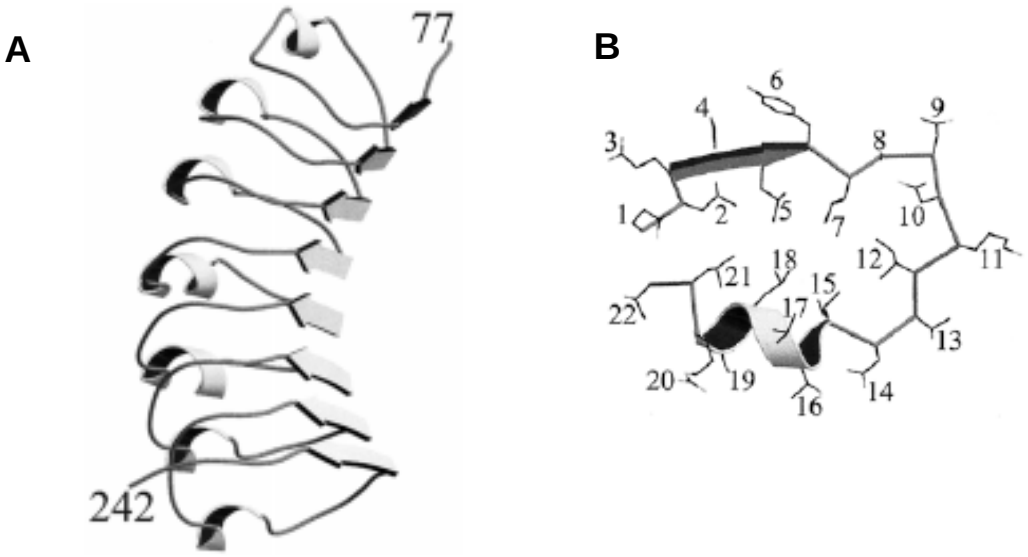


Figure 1.2. A. Structure of the InlB LRR region (residues 77-242). The right-handed coil of the LRR alternates between β-strands and 3_{10}-helixes. The β-strand forms the concave face of the molecule and have a superhelical twist not observed in other (non-listerial) LRR proteins. The β-strand region is predicted to form the interface for potential protein-protein interactions. **B.** Structure of a single InlB LRR. Internal positions 2, 5, 7, 10, 12, 14, 15, 17, 18, 21 are conserved for structural reasons. The remaining positions are solvent exposed and variable. From (Marino et al 2000).

All internalins of *L. ivanovii* reported to date (i-InlC, i-InlD, i-InlE, i-InlF) are small, secreted, and controlled by global virulence regulator PrfA. In contrast, only one such internalin, InlC, is reported in *L. monocytogenes.* Small internalins usually consist of six LRRs and no C-terminal domains. In intravenously infected mice, InlC deficient *L. monocytogenes* mutants caused a 1.5 log increase in LD_{50} (Engelbrecht et al 1996), while i-InlE-InlF deficient *L. ivanovii* mutants caused ≥1.5 log increase in LD_{50} (Engelbrecht et al 1998b). Their eukaryotic targets are not known.

Many cell surface associated,'large' internalins have been identified in *L. monocytogenes*; these include InlA, InlB, $InlC_2$, InlD, InlE, InlG, InlF and InlH. All except InlB share the cell wall anchor signal commonly found in Gram positive bacteria, which destines them to be covalently bound to the cell wall petidoglycan by the enzyme sortase. This signal consist of the LPXTG motif, a C-terminal hydrophobic domain and a charged tail (Dhar et al 2000; Navarre & Schneewind 1999; Schneewind et al 1993). Of these 'large' internalin

genes, only *inlA* and *inlB* are controlled, partially, by global virulence regulator PrfA. InlA targets eukaryotic E-cadherin containing Proline in residue 16, and this specificity defines InlA's affinity to human, chicken or guinea-pig epithelial cells, but not to the mouse or rat counterparts (Lecuit et al 1999; Mengaud et al 1996). InlB's eukaryotic target is mammalian complement receptor gC1q-R (Braun et al 2000). InlB confers invasiveness towards a large range of cell types (Braun et al 1998; Parida et al 1998), but this effect depended largely on the multiplicity of infection (m.o.i.) of bacteria to host. Unexpectedly, InlB may be important for intracellular replication in mouse liver beyond permitting initial invasion *in vivo* (Gregory et al 1997). *InlAB* double mutants showed a one log increase in LD_{50} in mice infected intravenously (Dramsi et al 1997). The phenotypes of the remaining large internalins of *L. monocytogenes* have not been clearly deciphered. Individual mutants of *inlC$_2$DE, inlC$_2$, inlD, inl E, inlF,* are not defective in invading the usual battery of cultured cell lines (Dramsi et al 1997). Even more curiously, deletions of *inlG, inlH, inlE,* and *inlGHE* caused increases of invasiveness in all cell types tested except in macrophages, which remained the same as wildtype (Bergmann and Raffelsbauer, personal communications). Invasiveness may not be the exclusive function of internalins. The concerted action of internalins with each other or with other gene products may have additive, synergistic, antagonistic, or competitive effects in different environments. Possibly, some non-PrfA controlled internalins may not be involved in pathogenesis at all.

Although certain virulence functions are well defined *in vitro*, their roles in the actual pathology and persistence of infection *in vivo* are not well understood. There are clearly multiple pathways for *Listeria* to end up replicating in the liver, albeit with unequal efficiencies. Mutants deficient in cell to cell spread (ActA or PlcB), or invasion of hepatocytes, epithelial and endothelial cells (InlA and InlB), were all able to invade and proliferate in neutropenic mouse livers. This finding led to the hypothesis that there are two separate pathways for *Listeria* to persist in the liver: cell to cell spread and external invasion via the internalins and p60--see below (Appelberg & Leal 2000).

Other virulence associated functions:

Additional proteins are involved in virulence functions. The "invasion associated protein" (*iap)* locus encodes a murein hydrolase, and has two separate functions. IAP, also termed p60 for its molecular size of 60KDa, is necessary for proper cell division and for the invasive phenotype in mouse fibroblasts, hepatocytes and macrophages (Bubert et al 1992b; Hess et al 1995; Kuhn & Goebel 1989; Wuenscher et al 1993). A handful of yet to be characterized factors are implicated in adhesion functions. Some of these bind eukaryotic fibronectin, a glycoprotein prevalent in the extracellular matrix and cell surfaces (Gilot et al

1999). The Ami protein, an autolysin with amidase activity that is not essential for septation, is also implicated in virulence (Milohanic et al 2000).

Stress proteins are essential for listerial intracellular survival in professional phagocytes. In addition to the importance of immune evasion, these host cells are routes of dissemination to other tissues from the gut. Sigma B driven, (Class II) general stress response genes appear not to be required for virulence (Wiedmann M 1998). However, listerial homologues of *B. subtilis* Class III heat shock response proteins (ClpC, ClpE, ClpP) are crucial for virulence as well as stress tolerance in *L. monocytogenes*. ClpP is a serine protease, while ClpC and ClpE are regulatory ATPases of the HSP100 class of chaperones. These Clp proteins, when complexed with other ATP binding Clp proteins that dictate substrate specificity (e.g. ClpA, ClpX, ClpY), conduct ATP-dependent degradation of abnormal or short-lived regulatory proteins and disassembly of protein complexes and aggregates. Clp proteins therefore ensure, not only proper protein folding, but also regulate transcription of important cellular functions such as competence (Schirmer et al 1996). Listerial *clpC* mutants have reduced heat and salt tolerance, phagosomal escape, adhesion and invasion of mouse hepatocytes, and impaired *inlA, inlB* and *actA* transcription. The *clpE* mutant is defective in stationary phase survival, and the *clpCE* double mutant is salt and heat sensitive as well as disabled in cell-division. When virulence is measured in mice infected intravenously, the LD_{50} of *clpC* and *clpE* mutants increased 2 to 3 logs and 2 logs respectively, but the *clpCE* double mutant increased 4 logs as compared to the wildtype (Nair et al 1999, Nair et al 2000; Rouquette et al 1998; Rouquette et al 1996). Deletion of *clpP* resulted in diminished availability of active LLO, and in the bacteria's decreased readiness to escape phagosomal vacuoles of macrophages. The LD_{50} of *clpP* mutants showed a 3.5 log increase against the wildtype (Gaillot et al 2000).

The regulatory interface between stress response functions and virulence functions is not well understood. How ClpC modulates *inlA, inlB* and *actA* gene transcription is unknown. Under conditions that normally induces virulence genes during exponential growth, the activation of PrfA suppresses ClpC expression; but entry into stationary phase overrides this suppression (Ripio et al 1998). This finding agrees with observations in cell culture. In 2D gel analyses, stress response protein profiles are not observed in replicating *Listeria* in the cytosol (Hanawa et al 1995). The Clp regulon is negatively controlled by the transcriptional regulator CtsR (Nair et al 2000a). Whether PrfA suppression of ClpC is exerted via the activation of repressor CtsR, or by an indirect route, is not known.

A two-component signal transduction system encoded by *lisRK*, is reported to influence virulence (Cotter et al 1999). Mutants deleted of *lisK*, the histidine kinase sensor, showed 1 to 2 log reduced virulence in mice infected intraperitoneally. In addition, they have

increase ethanol and stationary phase associated acid stress tolerance, but profound sensitivity to acid stress in growth phase. The role of LisRK in virulence regulation is unclear.

Intracellular replication also requires specific adaptations such as the activation of transporters to sample environmental signals from different host cellular compartments or take up nutrients from the host cytosol. Nonpathogenic *L. innocua* is unable to replicate intracellularly when they are introduced by microinjection into mammalian cells (Götz and Goebel, personal communications). Since glucose-1-phosphate (G-1-P) is the primary degradation product of glycogen, it is readily available in liver cells. Utilization of glucose-1-phosphate is strictly under positive PrfA control in *L. monocytogenes*. Moreover, growth on G-1-P does not appear to trigger catabolite-repression of PrfA, unlike growth on glucose, cellobiose and various other carbon sources (Ripio et al 1997). The decoupling of G-1-P utilization from catabolite repression is in agreement with the demand that carbon catabolism occur concurrently with other virulence functions, this ensures successful intracellular replication and cell to cell spread. The G-1-P transporter is encoded by an *hpu* homologue in *L. monocytogenes* and possesses a PrfA box in its promoter (Beck et al., personal communications). The OppABCDF system is essential for the uptake of oligopeptides greater than 3 amino acids in *L. monocytogenes.* Although the LD_{50} is not affected by the deletion of *oppA* in comparison to the wildtype, the mutant shows impaired phagosomal escape in macrophages, and delayed bacterial growth in the livers and spleens of infected mice (Borezee et al 2000b).

Genetic diversity and exchange in *Listeria spp*.

There is great heterogeneity in the natural population of *L. monocytogenes*. DNA-DNA optical comparisons of different isolates revealed a large spectrum of intraspecific relatedness with possibly some overlap with *L. innocua* (Hartford & Sneath 1993). Virulence gene polymorphisms revealed three distinct lineages within *L. monocytogenes*. Two of these are derived from human and animal isolates: I (serotype 1/2b and 4b) and II (serotype 1/2a, 1/2c and 3a). Lineage I is linked to human epidemics. Lineage III (serotype 4a) is non-human animal associated with a different ribotype (Wiedmann et al 1997). All *L. ivanovii* isolates collected so far belong to serotype group 5 and appears to be much less diverse than *L. monocytogenes* as a group as measured by DNA-DNA relatedness. There are less samplings and less data on the intraspecific relatedness in the nonpathogenic strains, although DNA-DNA hybridization studies also indicate diversity within these species (Rocourt 1982).

Lateral gene transfer is a powerful phenomenon accounting for numerous cases of pathogenic trait acquisition (James B. Kaper 1999). Two telltale signs indicate the occurrence of lateral transfer events. One is the intrinsic differences between the segment of DNA in

question and its resident genome: overall GC content difference, changes in codon usage patterns, frequencies of di- or trinucleotides, and associated remnants of mobile genetic elements such as flanking sequence repeats, integrases and transposes. The second is the similarity discrepancies of genes or blocks of genes not seen in other members of the group, but show close homology to phylogenetically distant organisms. As in the laboratory, bacteria can acquire foreign DNA via bacteriophages, natural transformation, and conjugation. However, the fact that bacteria maintain small, limited sized genomes indicate that assimilation and maintenance of any non-essential DNA over time is highly restrictive, and depend highly on the selective value they conferred. There are no demonstrated cases of virulence genes acquired via lateral gene transfer in *Listeria* to date.

The prevalence of genetic exchange in *Listeria* in nature is unknown. Bacteriophages are common in *Listeria spp*., but they tend to have extremely narrow host ranges that fall within small numbers of serovars of the same species (Hodgson 2000). They also tend to exhibit the same GC content profiles as their host (Loessner et al 2000; Loessner et al 1994). Thus, the known listerial phages are not likely to account for the acquisition of novel traits from distantly related organisms, and the transfer they effect between closely related organisms are likely undetectable. Natural competence has never been demonstrated in *Listeria spp*. in the laboratory, but this cannot be ruled out in nature. The efficiency of recombining non-mobile foreign DNA into the genome falls exponentially with DNA sequence divergence due to mismatch correction systems (Vulic et al 1997, Majewski & Cohen 1998). Sexual isolation is further exacerbated by the rarity of nearly identical stretches of sequences between donor and recipient DNAs that allow strand invasion events to initiate recombination (Majewski & Cohan 1999). However, as demonstrated by naturally acquired antibiotic resistances, genetic exchange between *Listeria* and diverse organisms are facilitated by rampantly promiscuous conjugative transposons such as Tn916 and Tn1545 (Clewell et al 1995; Poyart-Salmeron et al 1992), and broad host range plasmids (Poyart-Salmeron et al 1990). These routes of genetic exchange require that donors and recipients be present at the same time and in the same place, while transduction and transformation do not.

Aims:

The various studies presented here evolve around the central question: What makes a pathogen? Given a lineage of organisms containing both pathogens and non-pathogens, what engender these different lifestyle choices? What pre-adaptations are shared to allow the pathogen-host interaction? What new components are acquired and old components modified, disabled or discarded to optimize flexible adaptation to multiple niches? These questions motivate the three studies presented here. The first approach presented in Chapter 2 examines the historical relationship between the members of this lineage, and inquires the basis from

which they diversified. The second approach presented in Chapter 3 is reconstitution, which can eventually identify the changes needed to convert a non-pathogen to its pathogenic relative. The third approach presented in Chapter 4 catalogs the genome contents of pathogenic *L. monocytogenes* for comparison with its nonpathogenic relative *L. innocua*. Since the last two projects are ongoing and too wide in scope for one individual lifetime, I am merely presenting my participation in the launching of these studies.

Chapter 2

Evolution of the genus *Listeria*

Solving the phylogeny of genus *Listeria*

In order to understand in the nature of the differences in pathogenic vs. non-pathogenic pairs of closely related organisms, it is crucial to know their genealogical relationship to each other. Moreover, since the genomes of *L. monocytogenes* and *L. innocua* are currently being sequenced, an evolutionary context is crucial to provide a framework with which to view each organism's genomic contents, their differences, and the development of their individual traits.

The records of evolutionary histories reside in the organisms' DNA. 16S or 23S ribosomal RNA have been the favored molecules for phylogenetic studies because of their universality, constancy of function, resistance to horizontal gene transfer, phylogenetically informative content covering varying spans of evolutionary distances. Using 16S rRNA, the genus *Listeria* was finally correctly established as a distinct taxon in *the Clostridium-Lactobacillus-Bacillus* branch of the eubacterial tree (Ludwig 1985; Stackebrandt 1981). Within this branch, it is most closely related to genus *Brochothrix* (Ludwig 1984). However, because of the great conservation of rRNA function and thus its great selective constraint, it is an unsuitable chronometer for measuring relatively recent events. Till now, establishing the phylogenetic relationships within the genus *Listeria,* except for *L. grayi,* have been problematic because of the highly similar listerial 16S rRNA or 23S rRNA molecules (Collins et al 1991; Rocourt 1988; Sallen et al 1996).

Unlike *L. gray (murrayi)* which were sufficiently divergent for 16S and 23S rRNAs to distinguish them as a separate lineage, the remaining mix of listerial-like organisms were often loosely classified as '*L. monocytogenes'*. These included bacteria exhibiting varying hemolytic and pathogenic properties, the latter was defined as by their role as disease culprits or the ability to infect mice in the laboratory. This loosely defined bunch was discovered to contain 5 distinct groups using DNA-DNA hybridization (Rocourt 1982). The groups were later defined as *L. monocytogenes* contain members that are generally hemolytic and pathogenic, the non-hemolytic and non-pathogenic group *L. innocua,* hyperhemolytic and pathogenic (particularly to sheep) *L. ivanovii,* non-hemolytic and non-pathogenic *L. welshimeri,* and mildly hemolytic but non-pathogenic *L. seeligeri*. For typing schemes, refer to (Jones 1992).

Because of the high level of homogeneity among all the species of *Listeria* except *L. grayi,* several schemes of relatedness among the five species have been forwarded. Using multilocus enzyme electrophoresis (MLEE) on 18 enzyme loci, Boerlin et al. proposed that *L. seeligeri* and *L. ivanovii* form one group while *L. innocua, L. welshimeri* and *L. monocytogenes* form another. In this scheme, *L. welshimeri* and *L. innocua* are the most

recently derived sister species of that group (Boerlin et al 1991). However, the 16S rRNA phylogeny reported by Vaneechoutee et al. placed *L. monocytogenes* and *L innocua* as sister branches, with *L. welshimeri* forming a deeper branching within the group (Vaneechoutte et al 1998). Clearly, *L. innocua* and *L. welshimeri* are closely related, but *L. monocytogenes* and *L. innocua* are equally closely related, and these two species are often physiologically indistinguishable, although one is potentially pathogenic and the other benign.

This study attempts to trace the evolution of the pathogenic lifestyle within the *Listeria* genus using selected sequence data from all six species. For each of the six species, DNA sequences flanking and including the virulence cluster *locus* were either obtained for this study or assembled from existing sources. Aside from the virulence gene cluster, sequence information from other chromosomal regions known to encode virulence associated internalins in some species were obtained for others. Along with house-keeping flanking genes of the virulence gene cluster, 16S and 23S rRNA, *iap*, the *locus* encoding the murine hydrolase of dual cell-division and virulence roles was also recruited to augment the data set for phylogenetic reconstruction. This represents the entire available set of DNA sequence information common to the 6 listerial species at this time. The phylogenetic studies were undertaken with the collaborative efforts of Michael Wagner and Michael Schmid at the *Lehrstuhl für Mikrobiologie* of *Technische Universität München* which furnished the software and facilities for the phylogenetic computations.

The genetic contents of each of the species, in the context of the further resolved phylogenetic relationships of the genus *Listeria* also presented here, leads us to postulate the following: 1. The main virulence gene cluster represents an assembly of genes constituting a "virulence cassette" once shared by the common ancestor of *Listeria monocytogenes, L. innocua, L. ivanovii, L. seeligeri* and *L. welshimeri*. 2. Were this cassette ever acquired from a foreign source, it no longer shows exogenous signs or traces of acquisition. 3. Within this genus, the virulence gene cluster and hence the pathogenic lifestyle, has been lost in at least two separate and independent events represented by the species *L. innocua* and *L. welshimeri*. 4. Today's 'internalins' are frequently duplicated and rapidly evolving genes with listerial-specific motifs. Genes encoding internalin-like, LRR-containing sequences seem to be present genus-wide as secreted or cell-wall anchored proteins; their diversification to specific pathogenic functions likely reflects recent species-specific or even strain-specific adaptations.

RESULTS: Virulence gene cluster *loci* of *Listeria spp*.

The complete virulence gene cluster *loci* of the 6 listerial species.

In order to obtain the complete genomic organization of the virulence gene cluster *locus* in all the listerial species, all existing virulence cluster sequences were assembled from public data bases. The data set was then completed using borders defined by conserved sequences within the genes *prs* and *ldh*, housekeeping genes respectively flanking upstream and downstream of the virulence cluster *locus*. Conserved primers were designed complementary to signature regions of these housekeeping genes and the subsequently obtained PCR products were cloned and sequenced. The chromosomal sequences of the virulence cluster *loci* of *L. innocua*, *L. welshimeri*, *L. grayi*, the 5' and 3' fragments of *L. seeligeri,* and the 5' and 3' fragments of *L. ivanovii* are reported here. Table 2.1 lists the optimized PCR conditions, PCR product information, and the accession number for each of the sequenced fragments. Table 5-B.4 (Chapter 5) lists the primers used.

To avoid redundancy and confusion in the nomenclature found among the previously published descriptions of various sequences, it is necessary to rename the previously designated "*orf_*" here as virulence cluster *locus* (*vcl_*). The previously known *orfA* and *orfB* of *L. monocytogenes* (accession number M82881(Vazquez-Boland et al 1992)) is designated here as *vclA* and *B*, while the previously designated *orfA, B* and *C* of *L. seeligeri* (X97014 (Lampidis 2000)), which are unrelated to the *orfA* and *orfB* mentioned above, are renamed here as *vclC, vclD* and *vclE* respectively. Homologues of *vclB* are present in all *Listeriae* while homologues of *vclA* are present in all *Listeriae* except *L. grayi*.

Figure 2.1 shows the comparative genomic organizations of this chromosomal region in the six species of *Listeria*. The sequences derived from this study are represented by letter codes, while information assembled from existing sources are noted by numbers. Table 2.2 shows the functional assignments of these open reading frames. Not surprisingly, the virulence genes were not present *in L. innocua, L. welshimeri* and *L. grayi*. Of note is that *vclB*, putatively coding for an 110 amino acid, conserved protein of unknown function is invariably present in all listerial species and is extremely conserved across the genus (70-92 % nucleic acid sequence similarity; 79-100% amino acid identity). Within this chromosomal region, the invariable elements are *prs, vclB* and *ldh*, while the regions between them appear rather plastic, accommodating the virulence gene cluster and/or other genes in the different species.

The map of *L. grayi* is most unlike that of the other listerial species in that *vclB* is adjacent to *prs*, while 5 other coding sequences (CDS): *vcl*C, *vclJ, vclF1, vclG1, vclG2,* and *vclF2* not shared in other species precede *ldh*. v*clA,* whose function is unknown, is present

Virulence Gene Cluster Loci: Region between *prs* and *ldh* in *Listeria spp*.

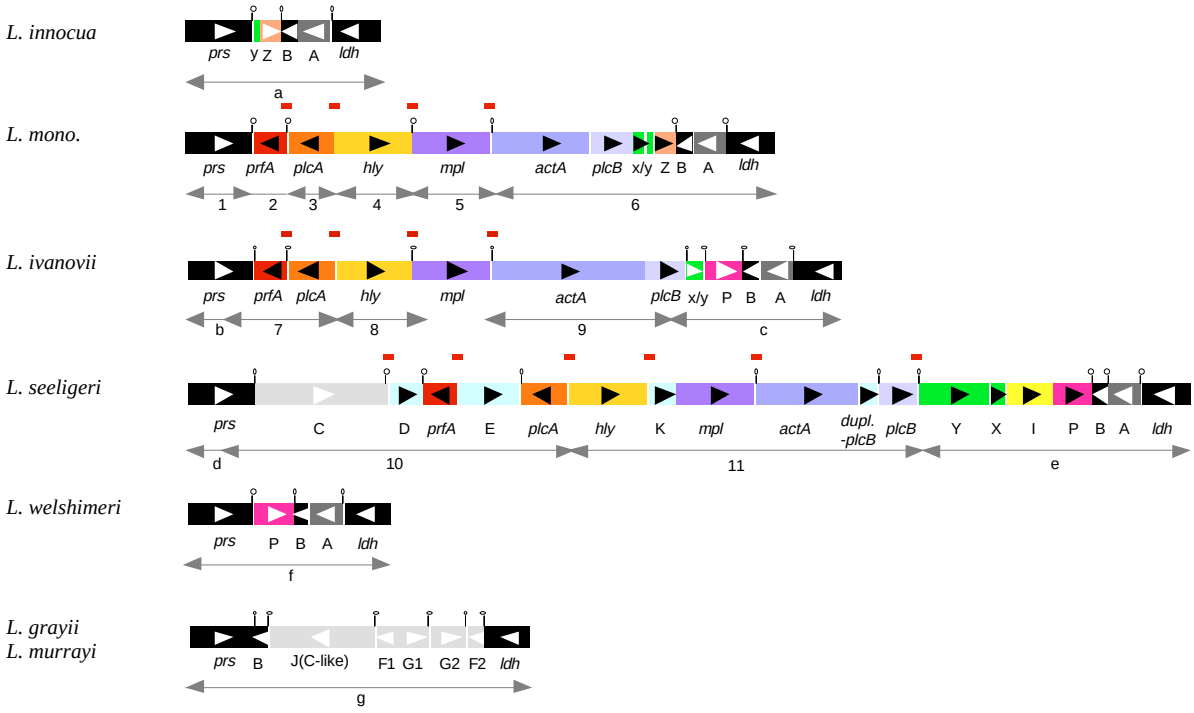


Figure 2.1. A schematic diagram of the virulence cluster *locus* (*vcl*) flanked by the housekeeping genes (black boxes) *prs*, *vclB* and *ldh* in the six member species of the genus *Listeria*. Genes that are proven, or are potentially controlled by master virulence regulator PrfA are shown with black arrows. Red boxes above denote PrfA boxes, and stemloops are predicted transcriptional termination signals.

vclA is present in all species except *L. grayi*. *vclP* is present in *L. welshimeri*, *L. seeligeri* and *L. ivanovii*. *vclZ* is present in *L. monocytogenes* and *L. innocua*. *vclY* and *vclX* are inverted in *L. seeligeri* with respect to all other *loci* carrying *vclXY*. Species-specific genes (stippled boxes) not under PrfA control include *vclJ*, *vclF1*, *vclG1*, *vcl G2*, *vcl F2* of *L. grayi;* and *vclC* of *L. seeligeri.*

Sources: Sequences presented in this study are defined as (letter code, accession number).
a, AJ249804; b, AJ249806; c, AJ249805; d, AJ249807; e, AJ249738; f, AJ249808; g, AJ249739. Sequences assembled from public sources are defined as (number code, accession number (reference)). 1, M92842 (Gouin et al 1994); 2, M55160 (Leimeister-Wachter M 1990); 3, X54618 (Leimeister-Wachter et al 1991; Mengaud et al 1991a); 4, X15127 (Domann & Chakraborty 1989); 5, X54619 (Domann et al 1991); 6, M82881 (Vazquez-Boland et al 1992); 7, X72685 (Lampidis et al 1994); 8, X60462 (Haas et al 1992); 9, U19035 (Gouin et al 1995; Kreft et al 1995); 10, X97014 (Lampidis 2000); 11, X97014, pending update (Lampidis 2000).

Table 2.1. Extending Listerial chromosomal information: Optimal PCR conditions, products obtained, and accession numbers of sequenced products.

Chromosomal region amplified	Organism	Product obtained	Product contains additional genes	Accession Number (Designation in Fig1)	Primer Pair	DNA polymerase	Optimal Annealing Temp (°C)
Virulence gene cluster	*L. monocytogenes* EGD	11 Kb	yes	NA	prs1> <ldh1	rTth, XL	50
	L. innocua Sv6b	2.8 Kb	yes	AJ249804 (a)	prs1> <ldh1	Deep Vent	54
	L. ivanovii ATCC 19119	3.4 Kb	yes	AJ249805 (c)	ivan-plcb1> <ldh1	rTth, XL	53
	L. seeligeri SLCC 3954	4.5 Kb	yes	AJ249738 (e)	see-vclH1> <ldh1	rTth, XL	50
	L. seeligeri SLCC 3954	0.8 Kb	no	AJ249738 (e)	see-plcb1> <see-vclH3	Deep Vent	54
	L. welshimeri SLCC 5334	2.9 Kb	yes	AJ249808 (f)	prs1> <ldh1	rTth, XL	53
	L. grayi	6.2 Kb	yes	AJ249739 (g)	prs1> <ldh1	rTth, XL	53
	L. murrayi	6.2 Kb	yes	NA	prs1> <ldh1	rTth, XL	53
Extending 5' of *prs*	*L. ivanovii* ATCC 19119	0.8 Kb	NA	AJ249806 (b)	con-prs1> <ivan-prs1	Taq	52
	L. seeligeri SLCC 3954	0.8 Kb	NA	AJ249807 (d)	con-prs1> <ivan-prs1	Taq	52
inlC locus	*L. monocytogenes* EGD	2.4 Kb	NA	NA	rpls2> <infC1	Taq	54
	L. innocua Sv6b	2.7 Kb	yes	AJ249401 (i)	rpls2> <infC1	Taq	54
	L. ivanovii ATCC 19119	2.1 Kb	yes	AJ249400 (h)	rpls2> <infC1	Taq	54
	L. seeligeri SLCC 3954	non-specific	NA	NA	rpls2> <infC1	Taq	none
	L. welshimeri SLCC 5334	2.4 Kb	yes	AJ249399 (j)	rpls2> <infC1	Taq	53
	L. grayi	non-specific	NA	NA	rpls2> <infC1	Taq	none
	L. murrayi	non-specific	NA	NA	rpls2> <infC1	Taq	none
i-inlDC locus	*L. monocytogenes* EGD	1.2 Kb	no	AJ010599 (k)	li-inlD4> <li-emr5	Taq	45
	L. innocua Sv6b	2.8 Kb	yes	AJ249398 (l)	li-inlD4> <li-emr5	Taq	45
	L. ivanovii ATCC 19119	3.9 Kb	NA	NA	li-inlD4> <li-emr5	Taq	50
3' of *i-inlDC locus*	*L. monocytogenes* EGD	2 Kb	no	AJ1010600	li-emr6> <li-emr1	Taq	45
	L. innocua Sv6b	2 Kb	no	NA	li-emr6> <li-emr1	Taq	45
	L. ivanovii ATCC 19119	2 Kb	NA	NA	li-emr6> <li-emr1	Taq	45
inlGHE locus (inlC$_2$DE)	*L. monocytogenes* EGD	6.7 Kb	NA	NA	pGluco2> <desuc1	rTth, XL	54
	L. innocua Sv6b	1.4 Kb	no	AJ249403 (n)	pGluco1> <desuc1	Taq	45
	L. ivanovii ATCC 19119	1.4 Kb	no	AJ249402 (m)	pGluco1> <desuc1	Taq	54
	L. seeligeri SLCC 3954	1.5 Kb	no	NA	pGluco2> <desuc1	Taq	47
	L. welshimeri SLCC 5334	1.6 Kb	no	NA	pGluco2> <desuc1	Taq	none
	L. grayi	0.8 Kb ?	no	NA	pGluco2> <desuc1	Taq	none
	L. murrayi	0.8 Kb ?	no	NA	pGluco2> <desuc1	Taq	none

NA = not applicable.
For details o f PCR reactions under each enzyme, please see section 2.4 Materials and Methods.

Table 2.2. Coding sequences (CDS) identified in the listerial chromosomal regions of the virulence gene cluster, between *rpls* and *infC*, and between the *rrn* and drug efflux pump operons.

Location	CDS	Present in	Size	Features	Greatest similarity	Organism with best matches	Blast score	P(N)
Virulence gene cluster	*vclA*	all except *grayi*	223-224aa	--	--	--	--	--
	vclB	all	110aa	--	conserved hyp. protein	*V. cholerae* (pir:G82395)	BlastP 333	3E-31
	vclX	*monocytogenes* *ivanovii* *seeligeri*	107aa 32+77aa 115aa	SP	--	--	--	--
	vclY†	*monocytogenes* *innocua* *seeligeri*	59aa 15aa 469aa	SP, membrane anchor. PrfA box (*see*)	numerous similar proteins in *L. monocytogenes* (not similar to VclC)	*L. mono EGDe (894.1)**	BlastP 458	1.2E-47
	vclZ	*monocytogenes* *innocua*	153aa 51aa	SP	hyp. lipoprotein	*E. coli* (sp:P33354)	BlastP 255	2.4E-35
	vclP	*welshimeri* *ivanovii* *seeligeri*	264-287aa	--	conserved hyp. protein, probable phosphoesterase?	*Thermotoga maritima* (pir:G72200)	BlastP 240	1E-19
	vclC	*seeligeri*	902aa	SP, 5x 78-80aa repeats, membrane anchor	C alpha antigen precursor (Bca)	*Strep. agalactiae* (sp:Q02192)	BlastP1298	E-144
	vclD	*seeligeri*	228-248aa	PrfA box	probable sugar isomerase. (AraD)	*E. coli* (prf:1303258C)	BlastP 364	3E-68
	vclE	*seeligeri*	423aa	PrfA box	--	--	--	--
	vclK	*seeligeri*	187aa	--	--	*L. seeligeri*	--	--
	vclI	*seeligeri*	304aa-284aa	SP, 5 LRR˜, no anchor.	Internalin (InlB) also VclC and VclJ	*L. mono* (M67471) *L. seeligeri and L. grayi*	BlastP 311	5E-28
	vclJ	*grayi*	716aa	SP, 1x LRR˜, 3x 78-79aa repeats, membrane anchor.	*L. mono* EGDe (1490.1, 1738.1), also L. seeligeri VclC, C alpha antigen	*L. mono* EGDe (1490.1)* *L. mono* EGDe (1738.1)* *L. seeligeri* *Strep.agalactiae* (PD018579)	BlastP 1168 BlastP 1078 BlastP 147 BlastP 165	1.3E-120 7.2E-115 2.3E-18 E-11
	vclF1	*grayi*	125-127aa	--	transcription regulator of the MerR family	*H. influenzae* (sp:Y186_HAEIN)	BlastP100	8.50E-17
	vclF2	*grayi*	126aa	--	transcription regulator of the MerR family	*Archaeoglobus fulgidus* (pir:A69334)	BlastP 88	2.50E-19
	vclG1	*grayi*	241aa	--	hyp. Oxyacyl (acyl carrier protein) reductase	*Agrobacterium tumefaciens* (AtsC, gpu:U59485_33)	BlastP 279	4.70E-44
	vclG2	*grayi*	253aa	--	hyp. Oxyacyl (acyl carrier protein) reductase	*Agrobacterium tumefaciens* (AtsC, gpu:U59485_33)	BlastP 360	1.90E-44
Between *rplS* and *infC*	--	*innocua*	119aa	--	hyp. protein (YtcD)	*B. subtilis* (pir:B69989)	BlastP 334	3.00E-31
	--	*innocua*	177aa	--	putative NAD(P)H oxidoreductase (YdeQ)	*B. subtilis* (pir:C69779)	BlastP 420	4.00E-41
	--	*welshimeri*	142aa	--	hyp.transcriptional regulator	*B. firmus* (BFU89914_7)	BlastP 133	1.00E-07
	--	*welshimeri*	287aa	--	conserved hyp. protein (YesF)	*B. subtilis* (pir:H69795)	BlastP 237	2.00E-19
Between *rrn* and efflux pump (*emr*) regulator	*orf1*	*innocua*	109aa	--	--	--	--	--
	orf2	*innocua*	215aa	--	--	--	--	--

SP is signal peptide predicted

† vclY in *L. monocytogenes* and *L. innocua* is present only as a truncated sequence.

˜ LRR stands for Leucine Rich Repeat, a 22amino acid unit (xLxxLxLxxNxLxDIxxLxxLx), which is characteristically present, usually in tandem, in the internalin proteins of *L. monocytogenes* and *L. ivanovii*. See Figure 2.2b below.

*data obtained from the *Listeria monocytogenes* Genome Project. The number is the Individual Protein File (IPF) designation.

downstream and adjacent to *vclB* in all *Listeriae* except *L. grayi,* although remnants of *vclA* sequence appear between *vclF2* and *ldh* of *L. grayi*. *vclP*, coding for a putative phosphate transfer enzyme, is specific to *L. seeligeri, L. welshimeri* and *L. ivanovii*. However, traces of *vclP* intragenic sequences could be detected between *vclJ* and *vclF1* of *L. grayi*. *vclZ,* coding for a putative membrane protein with similarity to a hypothetical *E. coli* protein, is specific to *L. monocytogenes* and *L. innocua*. Database searches revealed no known proteins similar to VclX, which is present in *L. monocytogenes, L. ivanovii*, and *L. seeligeri*.

Three prime to the established virulence genes, the *L. seeligeri* virulence gene cluster contains a few intriguing coding sequences, *vclY, vclX*, *vclI* and *vclP*, which are potentially under PrfA control by virtue of the presence of a PrfA box in the promoter region 5' of *vclY.* Analysis of *vclY* revealed that this is the entire gene corresponding to the truncated sequences seen *in L. monocytogenes* and *L. innocua*, and potentially encodes a cell-wall anchored

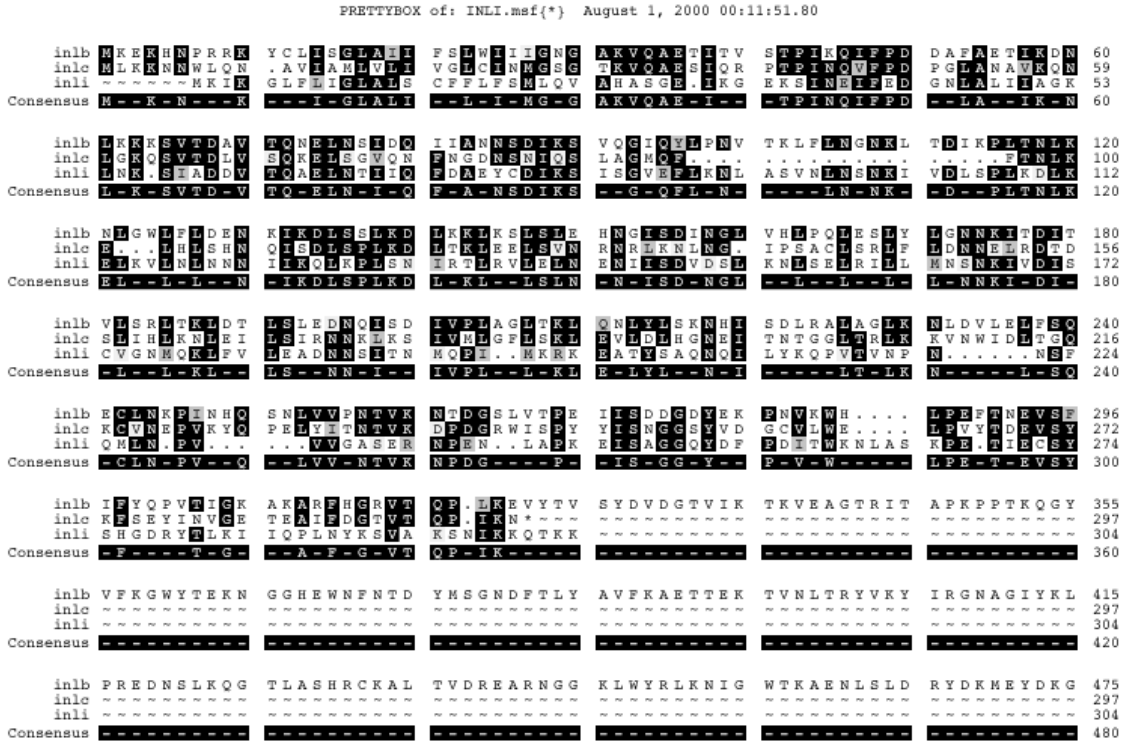


Figure 2.2a. (above) Alignment of *L. seeligeri* VclI to *L. monocytogenes* large internalin InlB and small internalin InlC. Black boxes denote amino acid identity, gray boxes denote similarity. Signal peptide cleavage sites for InlB and InlC are at amino acid 35 (VQA-ET/S); the predicted cleavage site for VclI is AHA-SG. VclI has almost 6 LRRs; and terminates at a similar place as the small internalins in the inter-repeat region. The remainder of the InlB sequence contains the B-repeats and the cell wall association region.

Figure 2.2b. (below) Alignments of the Listerial specific, 22 amino acid Leucine Rich Repeats from InlB, InlC and VclI according to the structural interpretation of Marino et. al., 2000. L = Leu/Ile conserved structural residues (Val, Met, Phe, Ala also tolerated); o = unconserved external facing residues thought to be important for protein-protein interaction, especially over the b strands which form the concave side of the protein. N = Asn residue (^) is important for stacking of the repeat unit to the one above. * denotes the hydrophobic, external facing residue of the first LRR thought to be important for contact with the N-terminal cap. Yellow = hydrophobic residues. Red = negatively charged residues. Blue = positively charged residues.

Alignments of LRRs from InlB, InlC and VclI.

	LRR		β strand		3_{10}Helix	
AA	pos.	oL	ooLoLo	oNoLooL ^	ooLoo *	Lo
B	77	SL	DQLLAN	NSDLKSV	QGLQY	LP
	99	NV	TKLFLN	GNKLTDL	KPLTN	LK
	121	NL	GWLFLD	ENKLKDL	SSLKD	LK
	143	KL	KSLSLE	HNGLSDL	NGLVH	LP
	165	QL	ESLYLG	NNKLTDL	TVLSR	LT
	187	KL	DTLSLE	DNQLSDL	VPLAG	LT
	209	KL	QNLYLS	KNHLSDL	RALAG	LK
	231	NL	DVLELF	SQ		
C	76	GV	QNFNGD	NSNLQSL	AGMQF	FT
	98	NL	KELHLS	HNQLSDL	SPLKD	LT
	120	KL	EELSVN	RNRlKNL	NGLPS	AC
	141	_L	SRlFLD	NNELRDT	DSLLH	LK
	163	NL	ELLSLR	NNKLKSL	VMLGF	LS
	185	KL	EVLDlH	GNELTNT	GGLTR	lK
	207	KV	NWLDLT	GQ		
I	69	TL	LQFDAE	YCDLKSL	SGVEF	LK
	91	NL	ASVNLN	SNKLVDL	SPLKD	LK
	113	EL	KVLNLN	NNLLKQL	KPLSN	LR
	135	TL	RVLELN	ENLLSDV	DSLKN	LS
	157	EL	RLLLMN	SNKLVDL	SCVGN	MQ
	179	KL	FVLEAD	NNSLTNM	QPL	

surface protein because of the presence of the Gram positive, cell-wall anchor signal sequence LPNTG (Schneewind et al 1993). Several proteins similar to VclY are present within the *L. monocytogenes* genome. There appears to have been a genetic inversion event in at least one of the virulence cluster carrying species or ancestral species as the order of *vclX* and *vclY* is inverted in *L. seeligeri* with respect to the *L. monocytogenes* arrangement.

The deduced VclI protein of *L. seeligeri* exhibits a remarkable amino acid sequence and structural resemblance to the internalins of *L. monocytogenes*, particularly to InlB. VclI contains a predicted signal sequence, the N-terminal cap region, 5-6 listerial specific LRRs, and terminates within the inter-repeat region at a place similar to the terminations of the small internalins; see Figure 2.2. VclI is the first internalin protein to be described in *L. seeligeri*.

(For a detailed discussion of the other *L. seeligeri* virulence genes, please see Lampidis et al. (Lampidis 2000)).

Since PrfA is the positive master-regulator of a majority of the known virulence genes, the presence of a PrfA box may implicate a potential virulence function of a gene in question. Figure 2.1 illustrates the locations of PrfA boxes. *L. seeligeri's* sequence predicts that a PrfA controlled promoter can drive the transcription of *vclY, vclX, vclI* and *vclP* on one polycistronic message; thus these genes are potentially important for virulence. In *L. monocytogenes*, *vclX* and *vclY*, along with *vclZ* can be transcribed via the PrfA controlled promoters of *mpl* and possibly *actA* (Mengaud et al 1991b; Vazquez-Boland et al 1992). Unlike in *L. seeligeri*, the *vclP* of *L. ivanovii* and *L. welshimeri*, and *vclX* of *L. ivanovii* are probably not controlled by PrfA as no PrfA boxes are present in the immediate upstream promoters of these configurations. For a discussion of PrfA boxes, see Brehm et al. (1996) and Williams et al. (2000) (Brehm et al 1996; Williams et al 2000).

Horizontal gene transfer?

The GC content of the virulence cluster *loci* genes was examined in order to detect significant differences that might indicate its acquisition via horizontal gene transfer. The average GC contents of the virulence genes of the cluster under PrfA control were 36%, 36% and 34% for *L. monocytogenes, L. ivanovii,* and *L. seeligeri* respectively. The average GC contents of the virulence cluster *loci* including all open reading frames between *prs* and *ldh* were 36%, 36% and 35% for *L. monocytogenes, L. ivanovii* and *L. seeligeri* respectively. There were no large discrepancies seen in the individual virulence genes from the reported, total genomic GC contents of 37-39% for *L. monocytogenes*, 37-38% for *L. ivanovii*, and 36% for *L. seeligeri* (Jones 1992). Therefore, there are no marked deviations in GC content between the virulence cluster *loci* from the total genome. Likewise, the GC contents of the virulence cluster *loci* of *L. innocua* (37%), *L. welshimeri* (37%), and *L. grayi* (41%) resemble their genome averages of 36-38%, 39% and 45% respectively. The only difference of note is *vclF2* of *L. grayi* with 34%GC vs. the genome content of 45%GC.

Furthermore, no insertion sequences (IS), obvious transponson, phage, or plasmid genes were detected. No direct repeats, '59-base elements' of integron gene capture systems, or partial identities there of flanking ORFs were identified.

Instead, the *L. innocua vcl locus* appears as *L. monocytogenes* minus the *prfA-plcA-hly-mpl-actA-plcB-vclX* genes. The borders of this apparent deletion were degenerate, such that the precise event(s) could not be reconstructed from sequence information. There were likely multiple deletion events that led up to *L. innocua's* present state as searches of the intergenic region of *prs* and *orfY* showed some short matches to the deleted virulence genes.

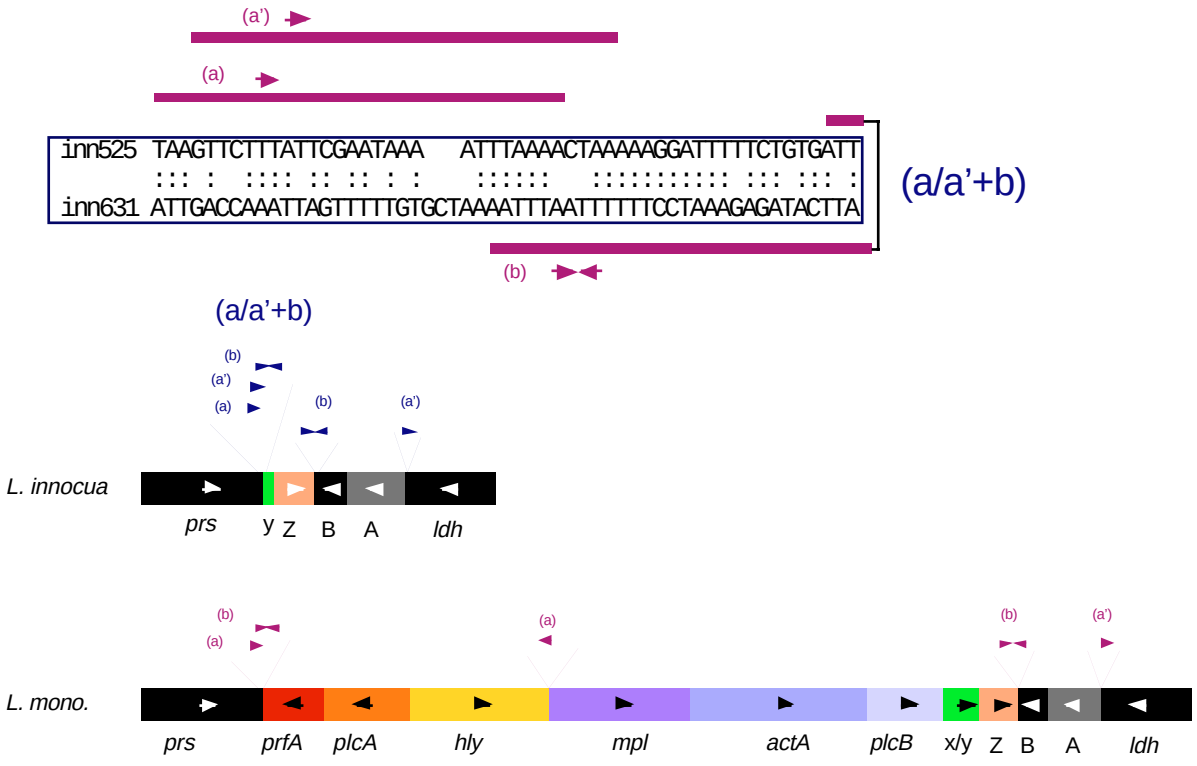


Figure 2.3. Region in *L. innocua* between and including the 3' of *prs* and *orfZ* forms a large inverted repeat (a/a'+b). This region contains motifs **a**, **a'** and **b** are found in other intergenic regions in both *L. innocua* and *L. monocytogenes*; most motif b's can form hairpin structures and often coincide with transcriptional terminators.

A peculiar structure was found between the *prs* gene and truncated *orfY* in *L. innocua*. See Figure 2.3. This intergenic sequence contains a large inverted repeat of approximately 110bp. This inverted repeat contains within it 3 smaller regions designated here as motifs **a**, **a'** and **b** which appear in the intergenic sequences throughout the *vcl loci* of both *L. monocytogenes* and *L. innocua*. Motif **b** is itself an inverted repeat which coincide with several transcription termination signals. Although the value of these observations are dubious, they became important later for practical reasons during the cloning of the *L. monocytogenes* virulence gene cluster in Chapter 3.

Results: Various internalin *loci* of *Listeria spp*.

Besides the main virulence cluster, an ever growing family of internalin genes are discovered to contribute to virulence in *L. monoctogenes* and *L. ivanovii*. These genes are known to be scattered in different sections of the genomes of both species. Using PCR with

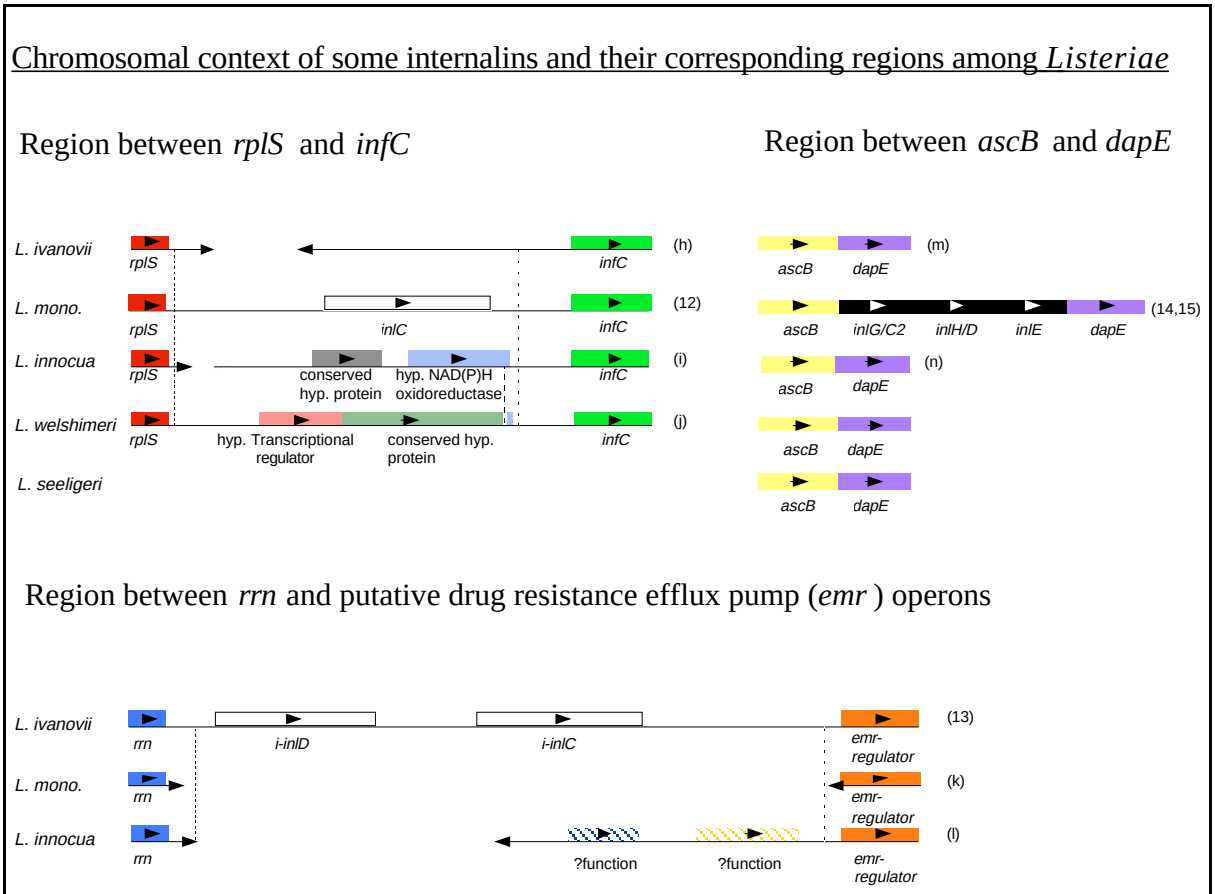


Figure 2.4. A schematic diagram of some chromosomal regions carrying internalin genes and the corresponding *loci* in other *Listeriae*. *inlC, i-inlD, i-inlC* (white boxes) are under strict PrfA control, while *inlGHE* (black boxes) are not. Regions of sequence divergence are defined by vertical dashed-lines spanning the alignments. Note the diversity of genetic content bordered by invariable housekeeping genes.

Sources: Sequences presented in this study are defined as (letter code, accession number) h, AJ249400; i, AJ249401; j, AJ249399; k, AJ010599; l, AJ249398; m, AJ249402; n, AJ249403. Sequences assembled from public sources are defined as (number code, accession number (reference)). 12, Y07640 (Engelbrecht et al 1996; Engelbrecht et al 1998a); 13, Y07639 (Engelbrecht et al 1998a); 14, AJ007319 (Raffelsbauer et al 1998); 15, U77368 (Dramsi et al 1997).

primers targeting house-keeping genes flanking some of these *loci*, we attempted to examine the corresponding chromosomal regions in the other *Listeriae*. Although many primer pair combinations were tried under different annealing temperatures for all six species, only some of these yielded specific products. Primers that amplified successfully are listed in Table 5-B.4 (Chapter 5), and the applied PCR conditions and obtained results are summarized in Table 2.1. This approach allowed the identification of the corresponding *inlC* (*L. monocytogenes*) genomic location in *L. ivanovii, L. welshimeri and L. innocua*; the *inlGHE* (*L. monocytogenes)* region in *L. innocua, L. ivanovii, L. welshimeri, and L. seeligeri,* and the

i-inlDC locus (*L. ivanovii)* in *L. monocytogenes*, and *L. innocua*. Figure 2.4 is a schematic illustration of these regions, Table 2.2 contains the functional assignment of the new open reading frames. The *inlGHE* genes of *L. monocytogenes* are located between the *ascB* (or *bglH*) gene encoding 6 phospho-beta-glucosidase, and the *dapE* (or *msgB*) gene encoding succinyl-diaminopimelate desuccinylase. The fragment size of the products resulting from PCR using primers targeting regions within the *ascB* and *dapE* genes showed that both genes are located directly adjacent to each other within the genome of *L. innocua, L. ivanovii, L. welshimeri* and *L. seeligeri*, indicating that *inlGHE* or *inlC$_2$DE* (Dramsi et al 1997; Raffelsbauer et al 1998) are unique insertion(s) in *L. monocytogenes*. This finding was confirmed by sequence analysis of the *L. ivanovii* and *L. innocua* fragments.

The *i-inlDC* genes of *L. ivanovii* are located between a ribosomal RNA (*rrn)* operon and a multi-drug efflux pump (*emr*) operon (Engelbrecht et al 1998a). Using specific primers that hybridize within this *rrn* operon and within the putative transcriptional regulator gene of the efflux pump operon, we identified two potential open reading frames encoding proteins of unknown function in *L. innocua*, but found no additional genetic content between the *rrn* and efflux pump operons in *L. monocytogenes*.

Although the *inlC* gene of *L. monocytogenes* encodes the closest homologue of *i-inlC* of *L. ivanovii*, it is situated in a different chromosomal location: between the *rplS* and the *infC* genes which respectively encodes ribosomal protein L19 and translation initiation factor IF3 (Engelbrecht et al 1998a). The cloned and sequenced PCR products from PCR reactions using primers to *rpls* and *infC* revealed, as in the case of *i-inlDC* described above, remarkable heterogeneity of genetic content between these extremely conserved house-keeping genes. In *L. innocua*, two genes exist in place of *inlC*, one resembles an NAD(P)H oxidoreductase and the other encodes a conserved hypothetical protein. *L. welshimeri* contains two genes entirely different from the above, a potential transcriptional regulator and another conserved hypothetical protein. Interestingly, downstream of the latter gene remains a fragment of sequence resembling the 3' end of the NAD(P)H oxidoreductase found in *L. innocua*, indicating that the *L. innocua* genetic arrangement is the ancestral state and the *L. welshimeri* version is a more recent replacement. *L. ivanovii*, on the other hand, contains a stretch of DNA that has no apparent coding sequences.

Results: Phylogeny of genus *Listeria*

16S and 23S rRNA had been used in earlier attempts to decipher the phylogenetic relationships of the members of the genus *Listeria*, which form a monophyletic grouping

Figure 2.5. Polymorphic sites within the genes used for listerial phylogenetic inference. **A** (previous page): *prs*, *ldh* and *vclB* genes. **B** (below): *iap* genes. The vertical numbers denote nucleotide positions from the start site of each gene. Numbers 1, 2 or 3 below the sequence alignments denote the codon position of each variable site. The underlined nucleotides signify changes that induce amino acid replacements. The alignments are done by Michael Schmid and Michael Wagner at the *Lehrstuhl für Mikrobiologie* of *Technische Universität München.*

A. ***prs* Polymorphic Sites**

```
           5555555666666666666666666666666666777777777777777777777777777777777788888888888888888888888888888888999999999999999999
           7777899000001112333445666777888999000011122222222333445566677778889999000011112222334445667889999900000111222334444 5
           4567547013692581369284679258147369256814701345679258270302812470360258124702360358140465479281245780369258137392568 1
consensus  GCGCTTATCTGTACAACCTTAATATTTTTTTTAACACTTAAATACGTGATCGAACATGATTATTAAAATTAGAAATCACGTTACTCTATTATAACACTCATAACACTAATATGTAT
  prs Lmo  .............GT...........C.........T...G.......A.T......CTA.GT.......G....T.....C....C.AT.GT...T...........T..C
  prs Lin  ......G........T....T...C.CC.CC....G.....................G.CT.C........T........A..C.........GT...T..G.T.......T..C
  prs Lwe  .......C..ACTT.GTTC..T.............................A....T........TG......TG.CTGTAC..T.........G.........TGT..GA......
  prs Liv  ....G..C..A.....TT..TG..A.......C..TTAA..GCT..........TT...AC...TG......G...T......T..CT.....G...G.......TAG.A......
  prs Lse  ....G.G..........T...TC.......C...T......G.........AA.....TT....G.....CG..C........T....C..GG......T..G.T.T.......C..
  prs Lgr  CGCG.C.CAA..GTC.T.CG..CCGCCCC..A.G..TAACC.CTAAAATC.TCT..A......A..GGACCC..G...T.G..T.TA.CC.GT...TG.TC.G.G.G.G.GC..G.
  cod.pos  1231333313333333333333313333333333331333331312313333133313331333313333133333333331313133333333233333133
```

***ldh* Polymorphic Sites**

```
           0000000000000000000000000001111111111111111111111111112222222222222222222222222222223333333333333333333333333333333333333
           0002233334445555667788999900122223444555556667778888991111222333444455666677778888899000011122222233344555556666677778888999
           1231703695680179460547036928103698147034692582470169020369258147036728014703692568947036928912347934989014790346725681478345679
consensus  GAATATTTTTATAATAGTCTCTTCATTTAACCTCTTTCAGTCATTCTCTACTCATAATCTTATATTTACTACCCAAAATCCTCCTTTTCACTAAACCAAAAAAACAAATACTAAATACCAAACAAAA
  ldh Lmo  ...CTCCC.......G....T.........TT...AC.C..T.....TA.A....G..T....C.C.....TAT...C.......C..T......T.........C........TC..........
  ldh Lin  .......C.C.......C..T....C....T....AC.C..TG....TA.G....G..T....CC..............C.A.CC.................CG.......TC..........
  ldh Lwe  .....C...C..GG......ACA...AC....C.......G...A.....T..C..............T.....T...........C.....G..............G.....CT.T.......T
  ldh Lse  ACG..CC....C.G.G.C.CA...G.G.T..T.T..C....AT...GTC...TG..TAG..G........T....TC.............T.G...T.......A.T...TG.G....T..G....T
  ldh Liv  .......AC..A.......C.....C......C........AC.........T.......C........A....TG..CT...A..C.A.T.....T.......A.T..........T........C
  ldh Lgr  ........G.TC..C.A.AA.CCTCCC.CG...GGCCTGA..TA.AA..CGC..CG...A..C...CTA...T.T..C.TAAA.G..CTGGCCTCGTCGCGTGCTT.CGGTGCG..GTTTTGAGCCG
  cod.pos. 1233333333132332131333333333333333333331333331333133133333333333331332333333333131333333331312332313312333233131331333331312313
```

***vclB* Polymorphic sites**

```
           0000000000000000000000000000000011111111111111111111111111111111111112222222222222222222222222222222222333333333333
           0000111223344455666777888899999000111112222333444445556666677778888889990000011111222333344445666778888999900001111222
           6789258373625814039358145701346258145780369258123470690256814790367892890147823469289146703468147360289134703692358901 2
consensus  TGTAACAATCATCATACTCGATTCGTCTTAGCTAAACTAACCCTATACAATTTACATACTTGAAATAGAAACAACGAATATCACGTTGTTTATATAATCTATATCATCATTTACAGCGC
  vclb Lmo  .....T.............CC.........T.GT...C.T..GC.G...C....T..T.C.....C.T.....T...........C......G..TCAC.C.....T...C.......
  vclb Lin  ............G......CC....T.....C.C.....T..GC.G...C....T....C.....C.C............T........C..GCT.........TCTTC...T.....
  vclb Lwe  A........A...GC.T.A...........T.GG......TT.T.C....CC..CC.TA.A.........TCG.T.......T..A.CAC.C.CC.CT....A...G.A.A....TT.
  vclb Lse  ....T....T.........A..CT...........T..............C....A..GC...T..A.......AC.C.C....G..CA...........C....C....C.....A.
  vclb Liv  ....T.....G.......A....T..T.......TC.C...T...A......CGT....A....T..A.G...G.C....CA...A...G..C...G.T....A....T....TG..T.
  vclb Lgr  CACGTACCC..CT..G.AT.CC.AACTGATA..T.TACCC.T.C..CGGCC.A..CCCACC.ACGGCGA..GA.CTAGCAC..CGACCC...GC.GT...CGAT.TTCAC.G.GG.TAAA
cod. pos.  3123333233333333331333313313133333313133333333123333313313333233312333123331231333313323331333333313313233333331331231
```

B. *iap* Polymorphic sites

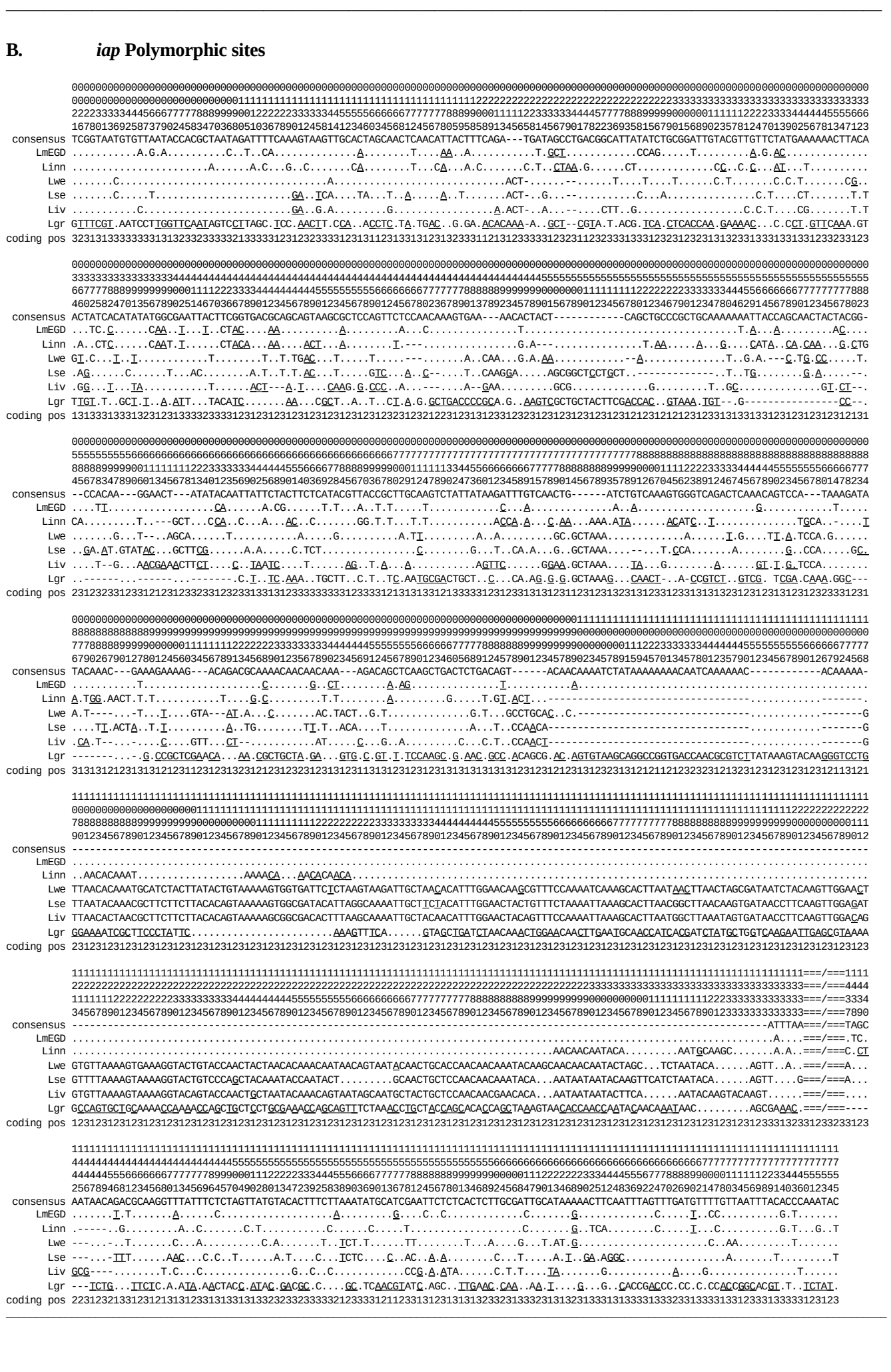


Table 2.3. Pairwise comparisons of *prs, vclB, ldh, iap*, 16S and 23S rRNA among *Listeria* species. This table is derived from alignments done by Michael Schmid and Michael Wagner at the *Lehrstuhl für Mikrobiologie* of *Technische Universität München.*

Similarity nucleic acid [%] / Identity amino acid [%]		*L. monocytogenes* EGD						*L. innocua* Sv6b						*L. welshimeri* SLCC 5334						*L. ivanovii* ATCC 19119 or SLCC 2379						*L. seeligeri* SLCC 3954						*L. grayi*					
		ldh	vclB	prs	iap	16S	23S	ldh	vclB	prs	iap	16S	23S	ldh	vclB	prs	iap	16S	23S	ldh	vclB	prs	iap	16S	23S	ldh	vclB	prs	iap	16S	23S	ldh	vclB	prs	iap	16S	23S
L. monocytogenes EGD	ldh							94						88						87						86						76					
	vclB								92						84						83						87						73				
	prs									93						88						87						88						81			
	iap										87						85						84						83						65		
	16S						99.6						99.4						98.7						99.3						96.6					
	23S						99.7						99.7						99.6						99.7						97.4					
L. innocua Sv6b	ldh	99												90						90						86						77					
	vclB		100												84						85						86						73				
	prs			99												87						87						89						79			
	iap				89												84						80						81						65		
	16S					n.a.							99.4						98.7						99.4						96.3					
	23S						n.a.						99.8						99.5						99.6						97.5					
L. welshimeri SLCC 5334	ldh	98						98												90						87						78					
	vclB		99						99												84						84						70				
	prs			97						98												91						89						79			
	iap				87						85												85						85						62		
	16S					n.a.						n.a.												99.1						99.5						96.7		
	23S						n.a.						n.a.												99.6						99.7						97.5
L. ivanovii ATCC 19119 or SLCC 2379	ldh	98						99						98												88						78					
	vclB		99						99						98												88						72				
	prs			98						97						96												87						79			
	iap				83						80						86												86						61		
	16S					n.a.						n.a.						n.a.												99.6						96.7		
	23S						n.a.						n.a.						n.a.												99.7						97.2
L. seeligeri SLCC 2379	ldh	98						97						97						98												76					
	vclB		97						97						96						98												73				
	prs			98						98						97						99												78			
	iap				84						81						89						86												61		
	16S					n.a.						n.a.						n.a.						n.a.												96.8		
	23S						n.a.						n.a.						n.a.						n.a.												97.3	
L. grayi	ldh	84						85						84						85						83											
	vclB		80						80						80						80						79										
	prs			90						90						90						90						90										
	iap				56						58						54						52						54									
	16S					n.a.						n.a.						n.a.						n.a.						n.a.								
	23S						n.a.						n.a.						n.a.						n.a.						n.a.								

n.a. Not applicable

The similarity values were calculated based on regions for which sequence information is available for all listerial species.

within the Gram-positive bacteria with a low DNA G+C content (Collins et al 1991; Sallen et al 1996; Vaneechoutte et al 1998). *L. grayi* had consistently appeared as the most ancestral branch of the genus in these studies. However, due to the high sequence similarities of 16S (98.7-99.6%) and 23S rRNA genes (99.5-99.7%) between the other members of the genus *Listeria*, the divergence of these species could not be resolved using these molecules. In order to enhance the resolution capacity of phylogenetic analysis for the genus *Listeria*, all available genetic information for the 6 listerial species were combined in a concatenated data set. This data set included in addition to the 16S and 23S rRNA molecules, the housekeeping genes *prs*, *vclB*, and *ldh* flanking the virulence cluster, and the *iap* gene sequences located elsewhere in the listerial genome.

The data were aligned and analysed using various phylogeny software packages available at the *Lehrstuhl für Mikrobiologie* of *Technische Universität München* by our collaborators Michael Wagner and Michael Schmid from the above institution.

Table 2.3 lists their pairwise nucleic acid similarities and amino acid identities of these genes among the 6 listerial species. Due to the high similarity of the deduced amino acid sequences of Prs, VclB and Ldh, most probably reflecting a very recent radiation among the members of the genus *Listeria*, phylogenetic inference was based on comparative analysis of nucleic acid sequences of *ldh, prs, vclB, iap*, 16S and 23S rRNAs. Figure 2.5 illustrates the character states of the polymorphic positions for *ldh*, *prs*, *vclB* and *iap* nucleic acid sequences used for treeing. Please refer to Collins et al. 1991 (Collins et al 1991) for the 16S rRNA data set, and Sallen et al. 1996 (Sallen et al 1996) for the 23S rRNA data set. Phylogenetic analyses were performed for each of the genes included in the alignments as well as for different concatenated combinations.

The data set.

Ldh codes for lactate dehydrogenase (LDH) which is generally about 310 amino acids in length. The region encompassing the last 134 amino acids was available for all 6 listerial species. This region represents the last 2 of the 6 conserved amino acid blocks of LDH as determined by the 'Blocks' database. For phylogenetic analysis the *ldh* genes of *Bacillus caldolyticus, Bacillus caldotenax, Bacillus stearothermophilus, Bifidobacterium longum, Deinococcus radiodurans, Lactobacillus casei, Lactobacillus sakei, Lactococcus lactis, Mycoplasma genitalium, Mycoplasma hyopneumonia, Streptococcus mutans, Streptococcus pneumonia, Thermotoga maritima, Thermus aquaticus* were obtained from GenBank, aligned and used as outgroups. Phylogenetic trees for the *ldh* genes were estimated from the nucleotide data set by distance, parsimony, and maximum likelihood methods. For all methods, the listerial *ldh* genes formed a monophyletic cluster with *L. grayi* as the deepest branch. The branch lengths of the other listerial species were extremely short and with the

exception of a consistent grouping of *L. monocytogenes and L. innocua*, no common branching patterns could be observed using the different treeing methods (data not shown).

Prs encodes ribose-phosphate pyrophosphokinase, a 318 amino acid protein in *L. monocytogenes*. The *prs* data set used for the 6 listerial species encompasses sequence information from residues 191 to 318, which contains the last 2 of 5 conserved amino acid blocks of PRS as determined by the 'Blocks' database. Phylogenetic trees for the *prs* genes were estimated from the nucleotide data set by distance, parsimony, and maximum likelihood methods. In distance and maximum likelihood methods *L. grayi* had the longest branch length. Considering *L. grayi* as outgroup, two stable groupings, (i) *L. monocytogenes* and *L. innocua* and (ii) *L. ivanovii, L. welshimeri* and *L. seeligeri*, were supported by all methods (data not shown). Independent from the treeing method used *L. seeligeri* always represented the deepest branch within the *L. ivanovii, L. welshimeri* and *L. seeligeri* cluster.

VclB is a conserved protein of unknown function found in all 6 listerial species, *Vibrio cholerae*--closest known homologue, other Gram positive bacteria and Archaea. Phylogenetic trees for the *vclB* genes were estimated from the nucleotide data set by distance, parsimony, and maximum likelihood methods. Results obtained were almost identical to those for the *prs* genes. However, the branching order within the *L. ivanovii, L. welshimeri* and *L. seeligeri* grouping differed depending on the treeing method used (data not shown). If the respective *vclB* homologues of *E. coli* and *B. anthracis* were included in the analysis, *L. grayi* always represented the deepest branch within the monophyletic *Listeria* cluster.

Iap has been described earlier. Both 16S rRNA and *iap* (gene and mRNA) have been exploited as target molecules for the detection and identification of *Listeriae* (Bubert et al 1992a; Greisen et al 1994; Wagner et al 1998; Wang et al 1992; Wang et al 1993). The P60 sequences of *Listeriae* consist of conserved N-terminal and C-terminal domains flanking a variable region containing varying numbers of TN repeats. Phylogenetic analysis of the *iap* genes using the different treeing methods supported the *L. monocytogenes* / *L. innocua* grouping if maximum likelihood or maximum parsimony methods were applied. However, both species were not monophyletic in neighbor joining trees. All methods consistently suggested a clustering of *L. ivanovii, L. welshimeri* and *L. seeligeri* (data not shown). In this cluster *L. welshimeri* always represented the deepest branch.

Trees derived from combined data sets.

In order to combine the phylogenetic information existing in the different genes, a composite tree was calculated based on the nucleic acid sequences of the concatenated 16SrRNA-23SrRNA-*iap-vclB-prs-ldh* data set (Figure 2.6). Independent of the treeing method applied, *L. grayi* always had the longest branch. The remaining 5 species consistently

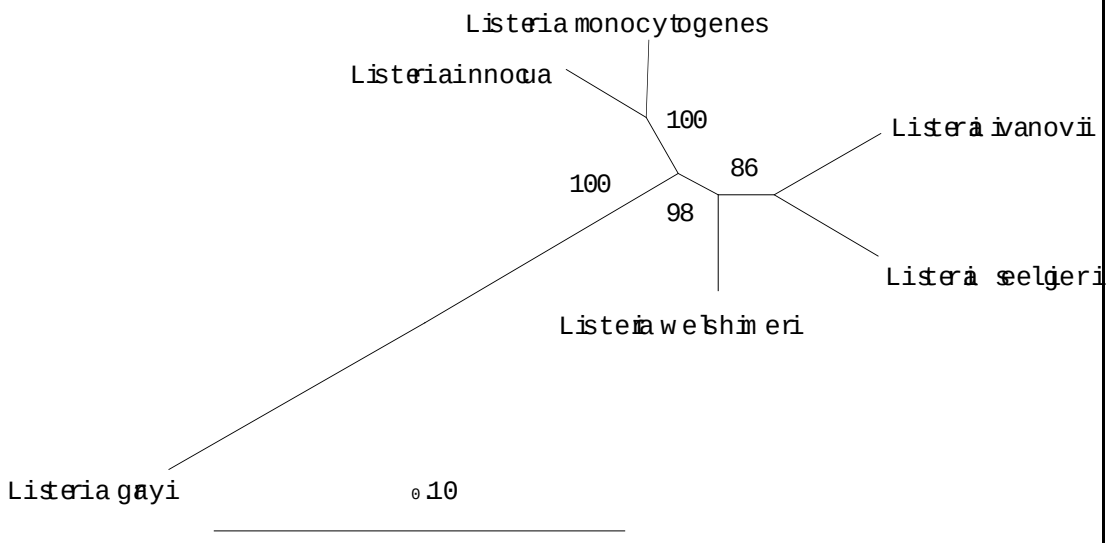


Figure 2.6. The composite phylogenetic tree of the genus *Listeria* based on comparative analysis of the concatenated nucleic acid sequences of 16S and 23S rRNA, *iap* (with the variable TN repeat region omitted), *prs*, *vclB*, and *ldh*. Tree topology and branch lengths were obtained from maximum likelihood analysis. Bootstrap values for branches were calculated using maximum parsimony (1000 resamplings). The bar indicates 10% estimated sequence divergence. Michael Wagner and Michael Schmid produced the tree at the *Lehrstuhl für Mikrobiologie* of *Technische Universität.*

fell into two distinct groups. One grouping represents *L. monocytogenes* and *L. innocua*, while the other contains *L. welshimeri, L. ivanovii* and *L. seeligeri*, with *L. welshimeri* forming the deepest branch within this group. Since the *iap* genes contain the majority of the informative sites in the concatenated data set, we were concerned that this molecule carried undue weight in the composite tree. Thus we applied filters which included only selected regions of the *iap* gene for phylogenetic analysis. Filters covering the 5' conserved block (alignment positions 1-1342), the 3' conserved block (pos. 1429-1755), and the 5' and 3' conserved blocks together were applied. In an additional analysis, the entire *iap* genes were omitted from the concatenated data set. None of the above mentioned permutations significantly affected the composite tree topology.

While the different phylogenetic analyses provided mostly consistent results, the exact branching order within the *L. welshimeri, L. ivanovii* and *L. seeligeri* could not be resolved unambiguously (see above). To obtain additional phylogenetic information, we also performed comparative sequence analysis of the *vclA* genes, which are present in all *Listeriae* except *L. grayi*. Unlike *prs*, *ldh*, *vclB* and *iap*, the nucleic acid similarity (73.7 – 90.5%) and amino acid identity values (75.8 – 94.5%) of *vclA* are significantly lower between the 5 listerial species (Table 2.3) thus allowing the use of both nucleic acid and amino acid based

phylogenetic analyses. The results confirmed the phylogenetic position of *L. welshimeri* in the composite tree described above. Using both amino acid and nucleic acid data, *L. welshimeri* appeared almost equi-distant to both *L. innocua / L. monocytogenes* and the *L. ivanovii / L. seeligeri* branches (data not shown). Further supporting this topology, the nucleic acid and deduced amino acid sequences of *vclP* genes (present only in *L. welshimeri, L. ivanovii* and *L. seeligeri*) show that the *L. ivanovii* and *L. seeligeri* sequences are more similar to each other than either are to the *L. welshimeri* counterpart (nucleic acid similarity / amino acid identity between *L. ivanovii* and *L. seeligeri* is 82% / 86%, 77% / 80% between *L. welshimeri* and *L. seeligeri*, and 73% / 74% between *L. welshimeri* and *L. ivanovii*).

DISCUSSION

Evolution of the virulence gene cluster.

Phylogeny of the genus *Listeria* and the loss of virulence capability in *L. innocua and L. welshimeri.*

We used a variety of *loci* including genes for 16S and 23S rRNA, "invasion associated protein", lactate dehydrogenase, ribose-phosphate pyrophosphokinase, and *vclB*, a conserved protein of unknown function, to determine the phylogeny of the genus *Listeria*. The resulting, composite, nucleic-acid tree (Figure 2.6) was further corroborated by phylogenetic analyses of the *vclA* or *vclP* genes, and VclA proteins. *L. grayi* represents the oldest branch of the genus while the remaining five species radiated recently into two lineages from a common ancestor. One lineage contains *L. monocytogenes* and *L. innocua*, while the other contains *L. welshimeri, L. ivanovii and L. seeligeri*. In the latter group, *L. welshimeri* occupies the deepest branch of this group.

This bifurcation of the *L. monocytogenes / L. innocua* lineage from the *L. welshimeri / L. seeligeri / L. ivanovii* group predicted by the phylogenetic tree is independently supported by the deletion pattern seen around the virulence gene cluster, and the presence or absence of the genomic markers *vclP* and *vclZ* in the respective organisms. *L. welshimeri, L. ivanovii* and *L. seeligeri* form one group containing the intact *vclP* gene at the 3' end of the virulence gene cluster, which encodes a putative phosphate transfer enzyme. Neither *L. monocytogenes* nor *L. innocua*, which form the other group, contain any *vclP* sequence. Likewise, *vclZ*, encoding a hypothetical lipoprotein, is present only in the *L. monocytogenes / L. innocua* lineage. Both *vclP* and *vclA* appear ancestral in the genus even though they no longer exist as coding sequences in present day *L. grayi*.

The phylogenetic tree we propose here implies that the virulence cluster genes of *Listeria* have been shared by the progenitor species of all *Listeriae* except *L. grayi* (if *vclP* is disregarded). Whether the progenitor of *L. grayi*, and indeed of all *Listeriae*, had already carried these virulence functions cannot be inferred to date. However, we can conclude that the virulence gene cluster has been lost in two independent events, presently seen as *L. innocua* from one group and as *L. welshimeri* from the other group of the two recently divergent listerial lineages.

Our findings based on nucleic acid sequences disagree with an earlier interpretation offered by Boerlin et al. based on MLEE (Boerlin et al 1991). While the position of *L. grayi* is in agreement, the members forming the two bifurcating groups differed. The MLEE analysis placed *L. innocua* and *L. welshimeri* as sister species in the *L. monocytogenes* group. *L. seeligeri* and *L. ivanovii* formed the other bifurcating group. This interpretation implies that the loss of the virulence gene cluster could have occurred once in the progenitor of *L. innocua* and *L. welshimeri*. None of our phylogenetic trees derived from two independent chromosomal locations, nor the chromosomal deletion breakpoints of the virulence gene clusters loci of *L. innocua* and *L. welshimeri* support the Boerlin scenario.

The discrepancies may be due to the different nature of the data from which the scenarios are derived. Recombination (lateral gene transfer) and evolutionary convergence of enzymatic functions may have obscured the MLEE data, giving misleading phylogenies if undetected. In MLEE analyses generally, the genes, their chromosomal location, the extent and nature of changes corresponding to the enzymes examined are unknown, and therefore cannot be assessed for occurrences of lateral gene transfer. Likewise, the enzymes examined cannot be verified to correspond to the same DNA sequences and not derived from paralogous gene products exhibiting similar enzyme activities between the species.

Our proposed tree also differ from that offered by Vaneechoutte et.al.(1998), which was based on 16S rRNA, using the neighbor joining method alone. This placed *L. welshimeri* as a deeper branch of the *L. monocytogenes / L. innocua* group instead of the *L. ivanovii / L. seeligeri* group. We failed to get consistent tree topologies using Maximum likelihood, neighbor joining, and distance methods with 16S rRNA data. Sallen et al. (1996) and Collins et al. (1991) also failed to definitively resolve these species using 16S rRNA. This is obviously due to the shortage of phylogenetically informative sites on the molecule. The genetic organizations of the various virulence gene cluster loci do not oppose the Vaneechoutte tree, since *vclP* appears ancestral to the genus, and is likely replaced by *vclZ* in the progenitor of *L. monocytogenes* and *L. innocua*.

While the exact position of *L. welshimeri* may still be debatable, the relationship of *L. monocytogenes* and *L. innocua* as sister species is clear from the collective data. When genome-wide data is available for several species, then the extent of lateral gene replacements

can be assessed, and perhaps, the uncontestable, definitive position of *L. welshimeri* can emerge. The problem of lateral gene transfer in the present data set is discussed below.

No evidence that virulence gene cluster arise recently via horizontal gene transfer.

The unremarkable variance of GC content observed in the virulence genes versus the house-keeping genes does not fit the, by now, classic model of a "pathogenicity island" acquired *en bloc* from a foreign source. However, this does not completely exclude horizontal gene transfer as a method of acquisition of these virulence traits; since any trace of GC content difference would not be detectable now if the GC content of the original source did not differ from that of *Listeriae* or the progenitor of *Listeriae*, or if the differences has been obliterated over time. In the latter case, the acquisition occurred long enough ago for the genetic content of this fragment to have fully adapted to the constraints of the listerial genomes.

The natural ecological niches of all *Listeriae* overlap, and are predominantly environmental. All six species are commonly found in soil, rotting vegetation, sewage, contaminated waters of rivers, canals, and estuaries. All six species also can be found in the intestinal tract of healthy animal carriers, and *L. monocytogenes, L. ivanovii*, and *L. innocua* were often found in the stools of healthy humans (Farber & Peterkin 1991; Rocourt & Seeliger 1985). Some of the environmental distribution may be attributed to fecal contamination, while the majority may indicate a saprophytic lifestyle of most *Listeriae* (Jones 1992). Since *Listeriae* often share the same habitats, genetic exchange among the species is theoretically possible, although the frequencies of these events are unknown.

Notwithstanding the fact that the virulence gene clusters reside in the identical chromosomal location in all the carrying species, one could theoretically invoke the possibility of horizontal gene transfer within the genus to account for its presence in both the *L. monocytogenes* and the *L ivanovii / L. seeligeri* / *L. welshimeri* lineages. Given that the host range and the virulence cluster genetic organizations of *L. monocytogenes* and *L. ivanovii* are more similar than with that of *L. seeligeri*, one could postulate that gene transfer would more likely have occurred between the former two species. However, the phylogenetic analysis performed demonstrated a close evolutionary relationship of the genes of the virulence cluster between *L. seeligeri and L. ivanovii*, thus supporting the notion that no such lateral gene transfer took place between *L. monocytogenes* and *L. ivanovii*. Moreover, the phylogenetic relationships of these species based on the concatenated *prs, ldh, vclB* genes located in the virulence cluster *loci* are consistent with the phylogeny derived from the *iap* genes, which are located in a different chromosomal site. For the phylogeny of these *loci* to simultaneously agree, two lateral gene transfer events between different *Listeriae* would have to be invoked, rendering this possibility even more unlikely.

The ancestral virulence gene cluster.

Although we cannot know what exactly constituted the "ancestral" virulence gene cluster from their present day manifestations, it must have included at least *prfA, plcA, hly, mpl, actA, plcB, vclX,* and *vclY*; the latter present in full in *L. seeligeri* but only as relic sequences *in L. monocytogenes* and *L. ivanovii*. Since *vclP* is linked to PrfA control only in *L. seeligeri*, we cannot discern if *vclP* is "ancestral" to the virulence gene cassette or a specific adaptation in *L. seeligeri*. Were *vclP* a part of the ancestral virulence cassette, this could imply that *L. grayi* has had pathogenic capability as well. *VclI*, the first internalin-like gene to be reported in *L. seeligeri*, appears potentially PrfA regulated. This linkage of PrfA control with an internalin-like gene to the virulence gene cluster may represent an ancestral arrangement that gave rise to the PrfA-controlled, internalin genes found widely dispersed in the present day genomes of *L. monocytogenes* and *L. ivanovii*. Whether potentially PrfA controlled *vclD and vclE,* and the non-PrfA-controlled *vclC,* represent recent insertions into *L. seeligeri* or deletions from the "ancestral" cluster in *L. ivanovii* and *L. monocytogenes* cannot be determined. The content differences between these virulence gene clusters, the genetic inversion event(s) of *vclX* and *vclY,* and the presence of the partially duplicated *plcB* gene in *L. seeligeri* are testaments to the dynamic history of these *loci* as they underwent adaptations in their resident species.

Evolution of the internalin-like proteins.

The internalin genes represent a different scenario among *Listeriae*. The internalins of *L. monocytogenes*, with the exception of InlC, are larger than *L. ivanovii's* and are cell surface bound or associated (Navarre & Schneewind 1999). This cell wall anchoring is enabled by their additional C-terminal domains, which are missing from all *L. ivanovii* internalins reported to date. The smaller, secreted internalins of *L. ivanovii* known to date are all under strict PrfA control (Engelbrecht et al 1996; Engelbrecht et al 1998a; Engelbrecht et al 1998b), whereas only some of the larger internalins of *L. monocytogenes* are partially controlled by PrfA (Dramsi et al 1993; Lingnau et al 1995; Raffelsbauer et al 1998). Most of the known internalin genes of *L. monocytogenes* and *L. ivanovii* reside in numerous and diverse locations in their respective genomes. Many of these genes are present in multiple, divergent, tandem copies: *inlAB* (Gaillard et al 1991), *inlGHE* (Raffelsbauer et al 1998) or *inlC$_2$DE* (Dramsi et al 1997), *i-inlDC* (Engelbrecht et al 1998a), *i-inlFE* (Engelbrecht et al 1998b). Some of these insertion sites are shown here to be unique for the species described as their corresponding chromosomal locations in the other species invariably contain either nothing or something else bordered by the same highly conserved housekeeping genes, which no doubt mark genomic locations less tolerant of change. In addition to frequent duplications,

illegitimate recombination is evidently the mechanism that generated *inlH* of our EGD strain (Raffelsbauer et al 1998) from *inlC_2* and *inlD* of the EGD strain used by Cossart's group (Dramsi et al 1997). The *inlC_2DE locus* likely contains *inlGC_2DE* as the 5' reported sequence of *inlC_2DE* matches the partial corresponding sequence of *inlG* exactly (Raffelsbauer, personal communications). This gene conversion event thus resulted in a new internalin (*inlH*) and a net reduction of one internalin unit in this tandem gene array.

As suggested by Engelbrecht et al. (1998), interspecific gene transfer may have played a role in the dispersal of internalins. Because *i-inlDC* reside next to interrupted *tRNA* genes, which are frequently targets of integration sites by exogenous mobile genetic elements (Ochman et al 2000), Engelbrecht noted that this insertion might have been mediated by a lysogenic phage carrying the *i-inlD, i-inlC* and the multi-drug efflux pump (emr) operon ((Engelbrecht et al 1998a), personal communication). It is not apparent what mechanism introduced the internalin genes into this *locus*. There are no integrases, recognizable IS or phage elements in the vicinity (using the *L. monocytogenes* genome). The *emr* genes are present in both *L. monocytogenes* and *L. innocua*, but neither species carry internalins in this locus, therefore it is likely that the existence of this *emr* operon is historically independent of the *i-inlDC* genes. Nevertheless, in the context of *L. monocytogenes*, the small secreted PrfA-controlled *inlC* could have originated from the *L. ivanovii* homologue *i-inlC*. The mechanism(s) to account for the apparent mobility of these genes within and between genomes need further definition.

The leucine-rich repeat (LRR) motif, which is especially important in defining the biological activities of internalins (Braun et al 1999; Mengaud et al 1996), is found here to be present also in the non-pathogenic members of the genus. As mentioned earlier, *vclI* is an *L. seeligeri* internalin gene. Were it expressed, VclI would likely be secreted as it possesses a predicted signal peptide and lacks the C-terminal cell-wall anchor sequences. One LRR motif was also observed in *L. grayi*'s deduced VclJ (data not shown), which represents a large, anchored, surface protein with similarity to VclC of *L. seeligeri*. Both VclC and VclJ contain multiple 78-79 amino acid repeat units very similar to those observed in the C-alpha antigen encoded by the *bca* gene of *Streptococcus agalactiae;* 5 units are present in VclC while 3 units are found in VclJ. Preliminary information from the yet incomplete *L. innocua* genome also indicates the presence of multiple LRR proteins. Thus, it is likely that the listerial-specific LRR motifs are rampant entities in the genus. It might exist in a large variety of rapidly evolving surface molecules, each characterized by varying numbers of divergent LRR units, variable N-terminal and C-terminal amino acid sequence contexts as is exemplified by the small internalins, the large internalins, and now VclJ with its C-alpha antigen repeats. These different combinations presumably perform different functions while sharing the listerial LRR's mode of action.

Evolution of *Listeriae* as Pathogens.

L. monocytogenes and *L. ivanovii* are facultative, intracellular, mammalian pathogens. The general mode of transmission is foodborne. Although frequently found in the intestinal tracts of healthy human and diverse animal carriers, *L. monocytogenes* in humans (and animals) and *L. ivanovii* in animals (and rarely in humans) are also associated with severe infection involving septicemia and meningoencephalitis in fetuses, infants, the elderly and the immunocompromised with a high mortality rate (Jones 1992). *L. seeligeri* presents an enigma for this genus. The function of *L. seeligeri*'s PrfA controlled and potentially virulence genes is still a mystery since this species is considered non-pathogenic. This non-pathogenic status is defined by the experimental mouse model, and by the fact that *L seeligeri* had not been known to cause disease except in one documented case (Rocourt et al 1986). Experimentally, *L. seeligeri* was not able to express its virulence genes under laboratory conditions which were found to induce the *L. monocytogenes* counterparts. The *L. seeligeri* listeriolysin gene (*lso)* was however shown to be functional when the *L. monocytogenes prfA* gene was introduced into *L. seeligeri* and expressed under conditions determined for *L. monocytogenes* (Karunasagar et al 1997). These findings point to our lack of understanding at present of how and when *L. seeligeri* utilizes its virulence mechanisms, and indeed, to whom this pathogenic potential is primarily targeted. In other words, the entire host range of *Listeriae* most likely has not yet been determined.

Recently, PEST sequences have been identified in the Hly hemolysin of *L. monocytogenes* and were demonstrated as essential for pathogenecity and prolonged intracellular survival in mouse macrophages (Decatur & Portnoy 2000). These PEST motifs are eukaryotic in origin and target proteins for phosphorylation and/or degradation. By targetting Hly for host cytosolic inactivation, Listeria can effectively escape from the host phagosomes while preventing the host cell membrane from becoming compromised by residual Hly activity. This particular modification allows *Listeria* to maximally use the eukaryotic host cell as a haven against nutritional deprivation and immune attack. These PEST motifs are also present in the *hly* genes of *L. ivanovii* and *L. seeligeri*, indicating that all three species have been adapted for prolonged survival within eukaryotic cells.

The similar content and organization of the virulence gene clusters of *L. monocytogenes* and *L. ivanovii* might reflect their adaptation to mammalian hosts, while the *L. seeligeri* cluster with its additional genes might reflect adaptations to unknown host(s). One group reported in laboratory conditions, the ability of *L. monocytogenes* and *L. seeligeri* to survive and escape *Acanthamoeba castellanii* and *Tetrahymena pyriformis* ingestion by

rupturing the predatory protozoan (Ly & Muller 1990a; Ly & Muller 1990b). This survival mechanism is reminiscent of *L. monocytogenes* and *L. ivanovii's* pathogenic capability of breaching host cell compartments such as escaping the phago-lysosome of professional phagocytes and invading neighboring cells in their mammalian hosts. *Listeria*'s probable environmental association with protozoan opens the possibility that the virulence gene cluster could have evolved as a listerial defense mechanism against becoming protozoan food.

The recent estimate of the infectious dose calculated from outbreaks to induce listerosis in humans is 10^7 to 10^9 organisms in normal adults (Farber et al 1996). This is in contrast with infectious doses of 10 to 100 organisms sufficient to cause shigellosis in volunteer studies for *Shigella flexneri* (DuPont et al 1989), and the less than 10^3 organisms of *Salmonella spp*. calculated for Salmonellosis outbreaks (Blaser & Newman 1982). Compared to these more dedicated, pathogenic bacteria that also attack the gastrointestinal tract of humans and other mammals, the most virulent *Listeriae*, i.e. *L. monocytogenes*, still represents a relatively ill-adapted human pathogen. The high infectious dose, the finding that two listerial species exist in similar natural and mammalian environments after losing their pathogenic potential, and the fact that the environment contains the predominant distribution of *L. monocytogenes* and other *Listeriae*, indicate that their dominant lifestyle is not parasitic. *Listeriae*'s pathogenicity to humans is thus better characterized, in the present window of co-evolution, as ambivalent but opportunistic, rather than the classic parasite-host relationships represented by *Shigellae* and *Salmonellae*.

Chapter 3

Towards reconstituting a pathogenic *Listeria*

Reconstituting the virulence gene cluster in nonpathogenic *L. innocua*

The number of genetic components responsible for a specific co*mpl*ex trait can be bewildering. While mutagenesis is crucial for finding individual components involved in a particular phenotype, it cannot assess the contribution of yet unidentified genes in conjunction with the known ones. Positive reconstruction offers perhaps the most direct way of defining the genetic and environmental components that are necessary and sufficient to yield a specific phenotype. Reconstruction of course is of little use until enough components are already defined that contribute to the trait in question. The sum of the known components can then be used as a platform to test the role of other suspect components.

A number of *L. monocytogenes* virulence functions are well defined, particularly the genes of the virulence gene cluster and internalins A and B. The question now is how many more genes and how much do they contribute to form the virulence phenotype. For this question, the pathogenic-nonpathogenic pair of *L. monocytogenes* (strain EGD) and *L. innocua* (strain Sv6a) present a unique testing ground for virulence function accounting. Though they are very closely related (see Chapter 2), *L. innocua* has none of the virulence associated phenotypes observed *in L. monocytogenes*. The anticipated completion of the sequencing and the subsequent comparative analysis of the genomes of these two organisms will undoubtedly generate major new insights into new, suspect, virulence genes. Therefore at this moment, we are poised to benefit from a reconstruction model which allows multiple factor analyses and testing of *in silico* predictions in a modified *L. innocua* containing defined virulence genes of *L. monocytogenes*.

The first step of such a reconstitution is to re-create an *L. innocua* with the full co*mpl*ement of the virulence gene cluster of *L. monocytogenes*. Efforts to do this are presented here. This construct, when completed, will shed light on the magnitude of differences in adaptations towards pathogenesis not accountable at the sequence level between the two species, and can be used for measuring the contribution of known virulence genes as well as test suspect ones for specific phenotypes. Unknown functions of interest include the components of the signaling and regulatory components of *PrfA* activation, the metabolic adaptations to allow intracellular growth, and additional factors that confer or define host cell tropism. Aside from these identification and accounting purposes, progressive reconstitution can also be viewed as a pseudo-reverse evolution experiment to understand the divergence of these two species.

RESULTS

The virulence gene cluster locus spans 9 to 12 Kb in *L. monocytogenes* (strain EGD) depending on how the borders are defined. Since the subsequent studies desire a construct of *L. innocua* (strain Sv6b used) that is as similar to *L. monocytogenes* as possible to avoid potential polar effects and copy number induced artifacts, the goal was to insert the contents of the virulence gene cluster into the identical location in the *L. innocua* chromosome, which had discarded its version (see Chapter 2).

Genetic tools in *Listeria* spp. are severely limiting because of the lack of natural competence and the inefficiency of transformation by artificial means. Allelic exchange in *Listeria* requires cloning the desired fragment into a Gram positive/Gram-negative conditional shuttle vector whereby the plasmid DNA can be propagated and harvested from the *E. coli* host in sufficient amounts (2-5ug) to transform *Listeria* via protoplasting or electroporation. Because no positive selection system are available for the detection of double cross-over chromosomal integrands as in *Listeria*, the chromosomal integration procedure is time and labor intensive, routinely involving 20 or more continuous passages of a singly integrated population in non-selective media at 30^{o}C to encourage recombination, followed by subsequent PCR screens to discriminate the double cross-over events. For these reasons, the best strategy is one that requires the least number of integration events. However, this is hindered by a biological limitation of *E. coli*, which does not tolerate large plasmids, or large sequences of Gram positive DNA with very different GC compositions from its own. This difficulty is only mildly alleviated by using *E. coli* host strains DH10B or HB101, which carries the *doeR* mutation, rendering them more forgiving of larger plasmids.

The shuttle vector pWH1509E, whose replication in Gram positive bacteria is temperature sensitive, was used because it offered an additional antibiotic marker (Ampicllin or Tetracycline) for the screening of insert presence during cloning in *E. coli*. The Ampicllin (Amp) site was the chosen cloning site because its inactivation relieved further safety concerns over the introduction of Ampicillin resistance into *Listeria*, since ampicllin is currently an antibiotic of choice for treating Listeriosis. Erythromycin (Erm) was the selective agent in *Listeria*, and Tetracycline (Tet) was used in *E. coli*. See Figure 3.1.

Because of the anticipated difficulties in this exercise, all cloning procedures undertaken here included the following controls and tests in parallel: positive control for ligase activity, background controls for vector that failed to be properly digested, the ligatability of the fragments in question, the amounts of each fragment and vector used, the transformational efficiencies of the competent cells, the antibiotic plates, the specificity of the screening primers, and the activity of the polymerase for PCR. These controls at every step ensured that any failure to obtain proper transformants could not be caused by trivial technical reasons.

The following results describe the development of a workable strategy for integrating the virulence gene cluster into *L. innocua*, and in parallel, into the positive control, *L. monocytogenes* mutant, PKP1, whose virulence cluster genes are deleted except for *prfA* (Engelbrecht et al 1996). The ability of the genes carried on each successive construct to recomplement PKP1 provide a proof of function for the genes introduced into the *L. innocua* background, whose display of the phenotype in question may be dependent on additional, unknown components.

First strategy for integrating the *L. monocytogenes* virulence gene cluster into *L. innocua*

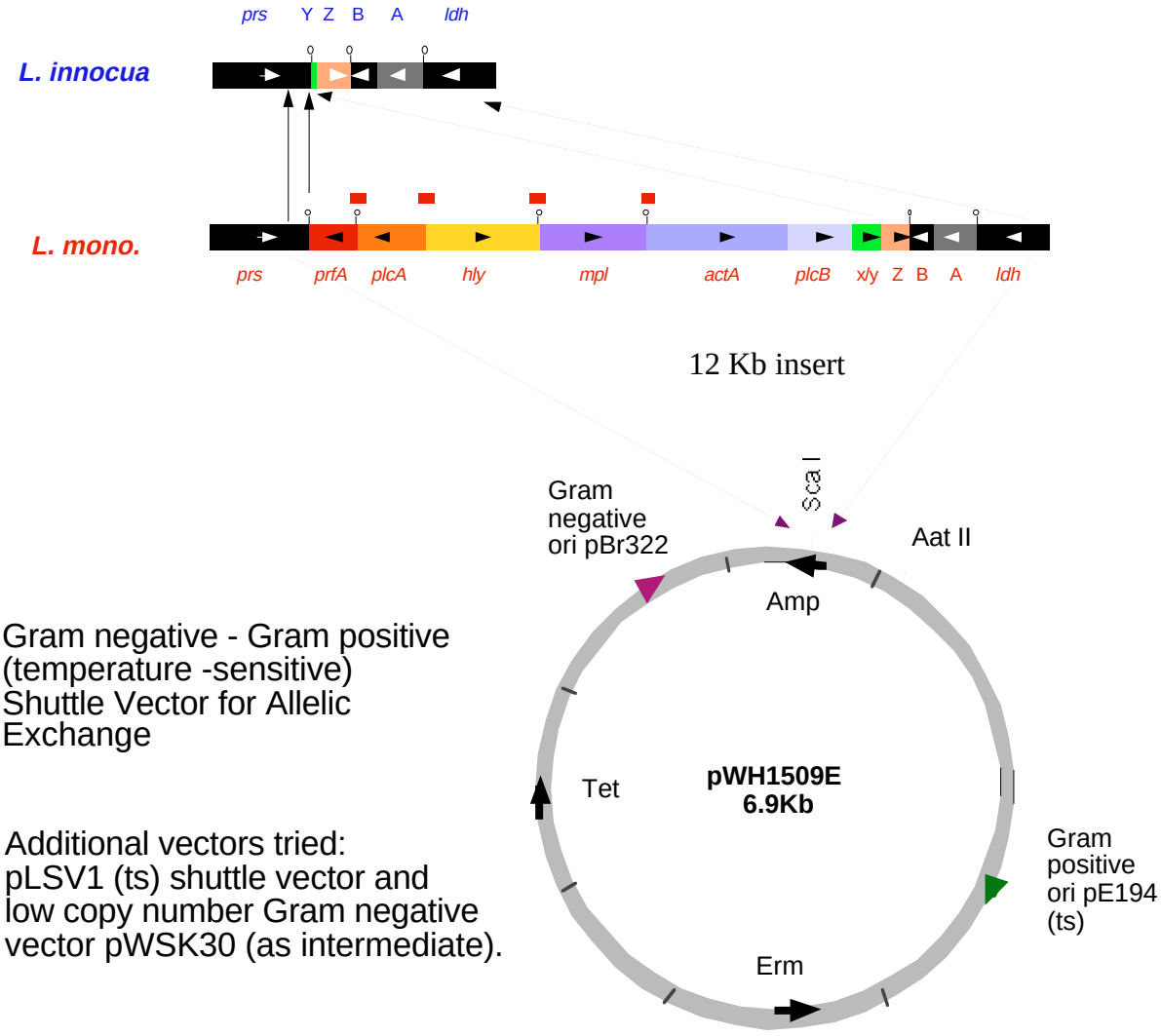


Figure 3.1. Direct cloning of the entire virulence gene cluster into a shuttle vector or an intermediate *E. coli* vector. Chromosomal integration would require cross-species tolerance at both 5' and 3' homologous recombination sites (5' = *prs-prfA*; 3' = *orfX*/Y, Z, B, A, *ldh*).

Cloning the entire 12 Kb fragment with homologous but non-identical sequences for integration into *Listeria.*

The chromosomal maps of the virulence gene cluster locus of *L. innocua* and *L. monocytogenes* revealed the known sequences available for recombination between these two species (Figure 3.1). A 12 Kb, blunt ended, PCR product was obtained from *L. monocytogenes* chromosomal DNA by PCR using long range polymerase rTth (Perkin Elmer) which had been spiked with Deep Vent polymerase with editing function to reduce error rate, and primers *prs*2(Sal)> <*ldh*5(Eco). This fragment was either directly ligated to vectors linearized with restriction enzymes producing blunt ends (for pWH1509E, the Sca I site within the Ampicillin gene), and transformed into *E. coli* DH10B cells, or phosphorylated with T4 polynucleotide kinase and then ligated to linearized vectors that were dephosphorylated with shrimp alkaline phosphatase (SAP), and processed as above. The shuttle vector pLSV1, and *E. coli* low copy number plasmid, pWSK30, were used in parallel as alternative cloning vectors. Sixteen tetracycline resistant (Tet^R)- Ampicillin sensitive (Amp^S) pWH1509E based transformants (both phosphorylated or not), 36 white pWSK30 (with dephosphorylation) based transformants (standard blue/white β–galactosidase activity screen), and 200 pLSV1 based transformants (phosphorylated) were screened by PCR to detect the correct insert. Positive clones yielded a 312bp product using primers *hly*2> and <*mpl*2 targeting the *hly* and *mpl* genes.

Seven pWH1509E (dephosphorylated) Tet^R/Amp^S clones, 1 white pWSK30 based clone, and a handful of pooled pLSV1 based clones were positive on the first day of screening. Upon propagation either by plating or in broth, all of these clones lost their inserts when subsequently checked by PCR with the same diagnostic primers above. This phenomenon was seen repeatedly, in each of the different vectors used, and the PCR signal decreased proportional to the degree of propagation. Therefore, *E. coli* proved unable to stably replicate 12Kb of Gram positive DNA in any of the vectors used. Remarkably, all the clones lost the entire inserts and none were seen with varying sized inserts. Upon closer examination of the 12Kb sequence, several very similar sequences representing the transcriptional termination signal regions of *prs, orfZ, B* and *ldh* were detected (see Fig 2.3 in Chapter 2). Since these locations represented the extreme ends of the 12Kb fragments, it is conceivable that their presence promoted deletions of what looked like the entire insert as viewed on agarose gels.

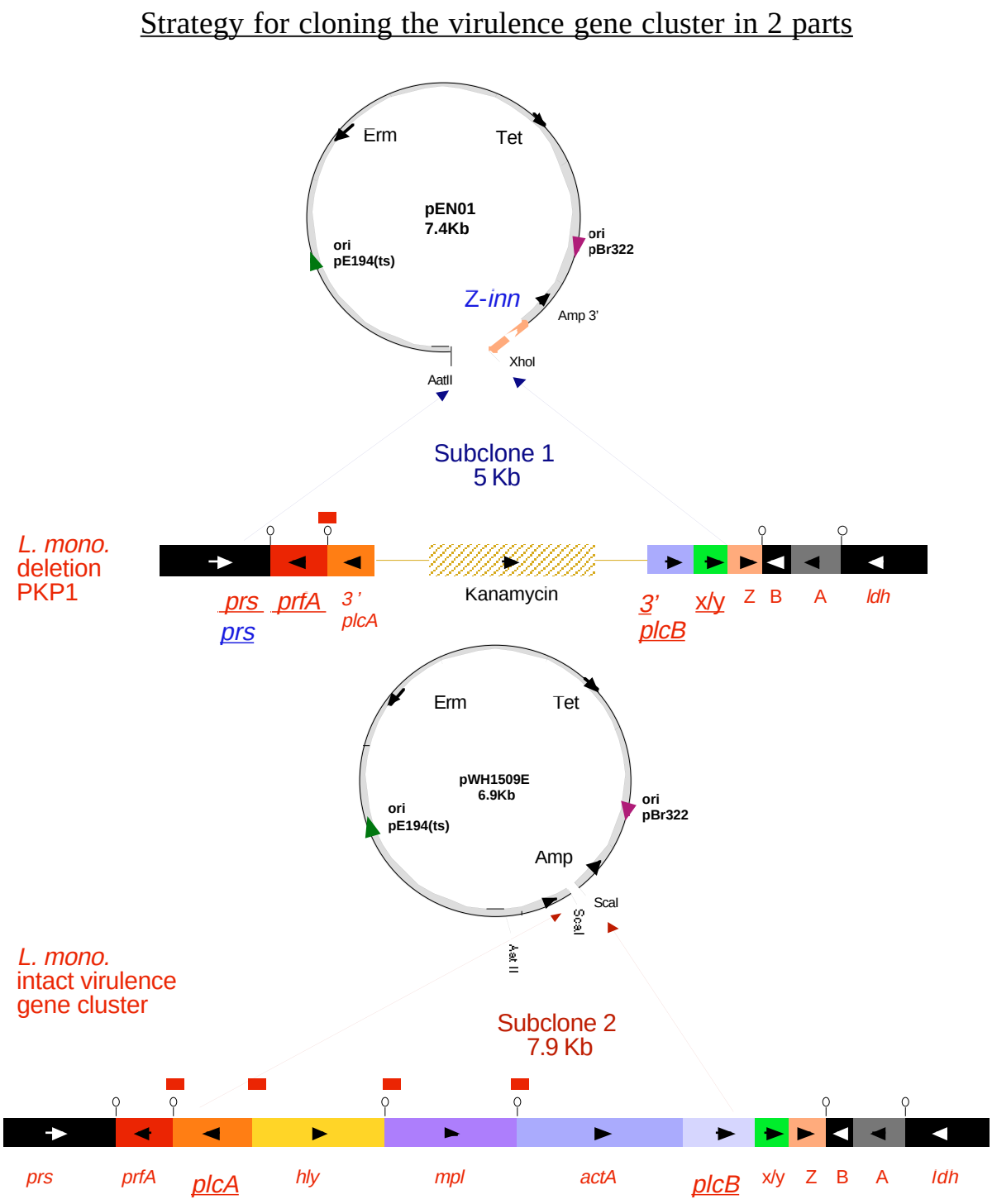


Figure 3.2. Attempt in cloning the virulence gene cluster in 2 parts, with 2 subsequent chromosomal integration steps. The first step employed subcloning of the virulence gene cluster deleted strain PKP1 (subclone 1), which offered a reduced size fragment but contained the appropriate 5' and 3' integration sequences. Its subsequent integration provide a Kanamycin marker, whose loss could indicate the co*mpl*ete integration of subclone 2. (Sequences for chromosomal recombination are underlined. Red *indicates L. monocytogenes* sequence. Blue indicates *L. innocua* sequences. Black indicates sequences introduced in the prior established integrand).

Cloning of the virulence gene cluster in 2 parts.

The second strategy was designed with the goals in mind to decrease the fragment size for cloning while avoiding the presence of multiple repeat sequences, and maximizing the ease of performing the subsequent two chromosomal integrations. This would employ the mutant PKP1 that carried *prfA*, partial *plcA*/*plcB*, interrupted in between them by a Kanamycin resistance marker (Kn.). The first integration clone would carry sequences containing *orfZ* of *L. innocua*, which would permit homologous recombination in the 3' border, fused to the 5Kb fragment from PKP1 representing *prs* to *orfX/Y*. The second integration clone would carry the virulence genes deleted from PKP1: *plcA, hly, mpl, actA,* and *plcB.* Its successful integration could be identified by the loss of the Kanamycin resistance of the first integrand. Both constructs would be made in pWH1509E. See Figure 3.2.

The preliminary cloning of *L. innocua orfZ*, a 500bp PCR fragment generated with Deep Vent polymerase (primers: *orfZ*8-inn(XhoI)> <*orfZ*9-inn), ligated into the ScaI site of pWH1509E, occurred with ease. The resultant plasmids were named pEN01 and pEN02 depending on the insert orientation. pEN01 and 2 were linearized with XhoI, a site introduced into the 5' end of the *L. innocua orfZ* fragment. These were used to clone the 5Kb PCR product generated by rTth/DeepVent polymerase from PKP1 chromosomal DNA, using XhoI sites engineered into the PCR primers (*prs*1(XhoI)> and <*orfX*1(XhoI)). The 8 Kb fragment containing *plcA, hly, mpl, actA, plcB* was similarly obtained by PCR from *L. monocytogenes* EGD chromosomal DNA (primers *plcA*1> <*plcB*1), phosphorylated with T4 PNK, and ligated into ScaI linearized, SAP dephosphorylated, pWH1509E.

Neither clonings were successful. No Kn^R transformants were select, and PCR screens of 65 pEN01 based clones and 65 pEN02 based clones proved empty (diagnostic primers: *plcB*2> <*orfZ*9-inn). Two different sized Tet^R colonies were observed for the 8Kb cloning, many of the small colonies failed to propagate (not satellites), the 72 propagatable Tet^R/Amp^S transformants were PCR screened (primers: *hly*2> <*mpl*2), but they too were empty. Using MAX Efficiency DH10B competent cells (BRL: transformation efficiency at 10^9 transformants/ug pUC19 DNA) as opposed to homemade electrocompetent DH10B (transformation efficiency at 10^7 transformants/ug pUC19 DNA, and $5x10^6$ for pWH1509E DNA) also failed.

At this stage, switching the intermediate cloning from *E. coli* to a low GC content, Gram positive bacterium as an intermediate host was considered. *Bacillus megatherium* and *Staphylococcus carnosus* were examined for feasibility. *Bacillus* protoplast transformations were plagued with numerous technical problems, while *S. carnosis* protoplasting only yielded 10^5 transformants/ug pWH1509K DNA (Kn. selection, data not shown), this efficiency was not promising for difficult, direct transformation of ligation products. This route was abandoned.

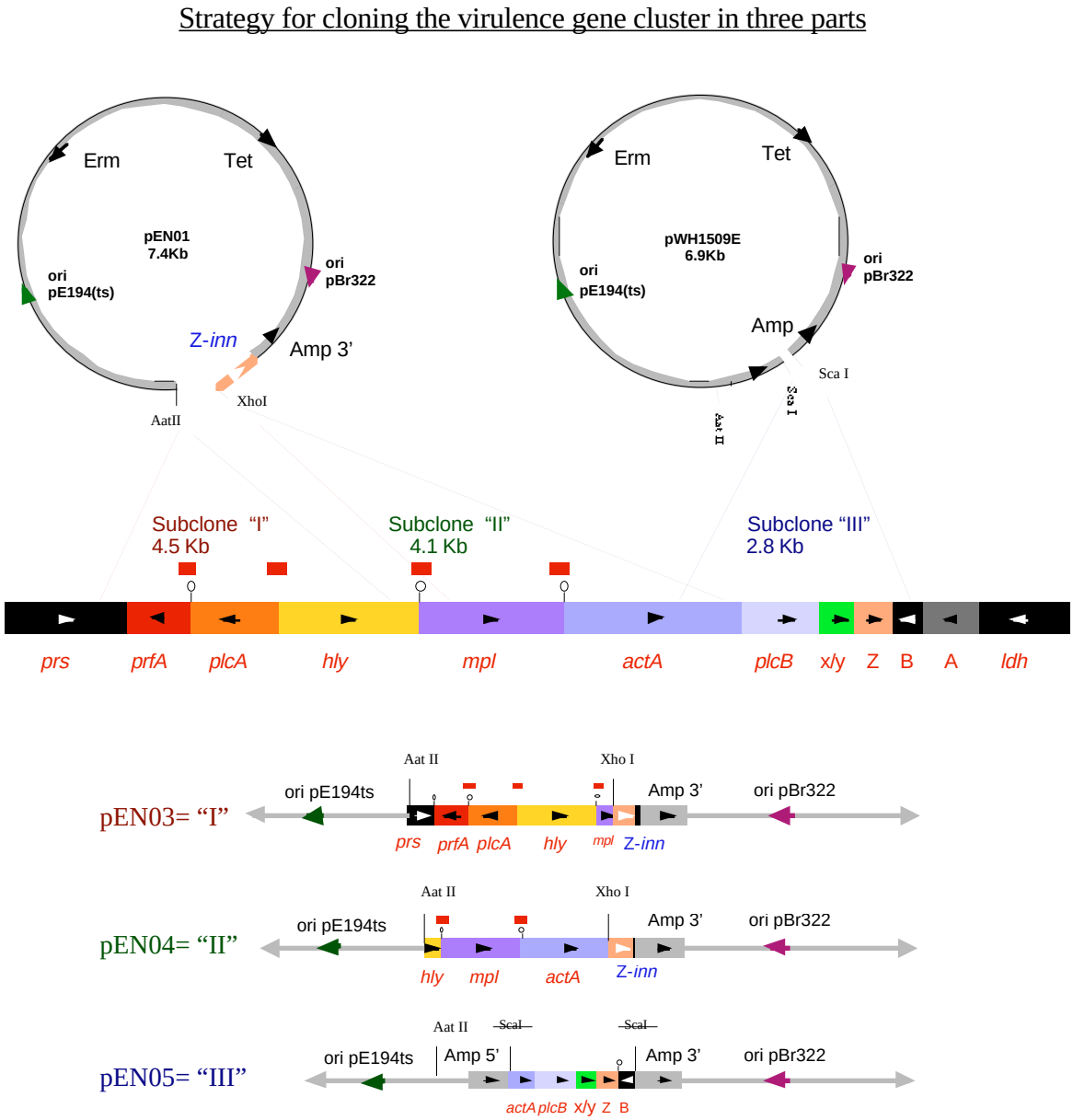


Figure 3.3. Three overlapping fragments containing the *L. monocytogenes* virulence gene cluster were cloned separately into the shuttle vector. The first two clones pEN03 and pEN04 employed pEN01 as a base which provided the *L. innocua orfZ* sequence the 3' border for chromosomal integration.

Cloning the virulence gene cluster in 3 parts.

It is apparent that the insert sizes need to be further reduced in the cloning of listerial DNA into *E. coli*. A third strategy was devised using three subclones. This is illustrated in Figure 3.3. Three PCR fragments were generated with rTth/Deep Vent polymerase. Fragment "I", carrying *prs*, *prfA*, *hly*, and 5' of *mpl*, is 4498bp (primers: *prs*5(Aat2)> <*mpl*2(XhoI)), fragment "II", caring *mpl* and *actA*-incom*pl*ete, is 4093bp (primers: *hly*3(Aat2)> <*ActA*1(XhoI)), and fragment "III", spanning the 3' end of *actA*, *plcB*, *orfXYZ*

to *orfB*-inco*mpl*ete, is 2815bp (primers: *ActA*2> <orfB3). At this point, two methods were tried. The first attempted cloning these directly into the shuttle vector, and the second involved cloning these fragments via an *E. coli* intermediate vector using the topoisomerase based TOPO TA cloning kit (Invitrogen). The second method yielded positive clones for all three fragments as confirmed by "cracking"to examine supercoiled plasmid size, and insert-plasmid junction sequencing. TOPO plasmids carrying fragments "I" and "II" were digested with AatII and XhoI, and these inserts were gel isolated and ligated to pEN01, which had been similarly prepared. Fragment "III" was released from the TOPO vector using the two EcoR1 sites provided by the vector flanking the insert. The sticky ends of this fragment were removed using DeepVent polymerase, and the resultant blunt-ended product was ligated to ScaI linearized pWH1509E. All were electroporated into *E. coli* DH10B cells.

Three out of 30 transformants were positive for "I" when PCR screened (diagnostic primers: *hly*2> <*orfZ*9-inn), 1 out of 62 transformants screened was positive for "II" (primers: *ActA*2> <*orfZ*9-inn), 3 "III" transformants were isolated from the 8 Amp^S/Tet^R out of 100 Tet^R colonies examined (diagnostic primers for both orientations: orfB2(Eco)> <3'AmpPst and orfB2(Eco)> <orfB3). The resultant plasmids were named pEN03 for "I", pEN04 for "II", and pEN05 for "III".

Chromosomal integrations into *L. innocua* and PKP1 required further modification of pEN03 and pEN04

To test the experimental feasibility of chromosomal integration using cross-species, homologous but non-identical sequences, pEN03 was electroporated into both *L. innocua* Sv6b and *L. monocytogenes* PKP1. The first crossover events were selected for in non-permissive temperatures under Erm selection. Temperatures operable for *L. monocytogenes* (41 to 43°C) were found to be lethal for *L. innocua*, but single crossover events were successfully isolated in *L. innocua* at 40°C. However, all crossover events observed took place at species-specific sequences. All *L. monocytogenes* primary integrations occurred using *L. monocytogenes prs-prfA*, while all *L. innocua* primary integrations took place at the *L. innocua orfZ*. The second integration event was promoted through continuous passage by diluting 0.5 to 1ml of culture in 20ml BHI, and grown at 30°C with no drug, which represented about 10 doublings per passage. A simple screen for the loss of Kanamycin resistance in the PKP1 primary integrand was used to identify second crossover events. The Kn. screens were done periodically between passages 20 to 50, but none were positive. At passage 50 or after 500 doublings, zero of the 2500 clones screened were Kn^S. This indicated that recombination between similar, but non-identical sequences (nucleic acid identity at >90% for *prs*, and 82% for *orfZ* over 500bp), occur at prohibitive frequencies in *Listeria*.

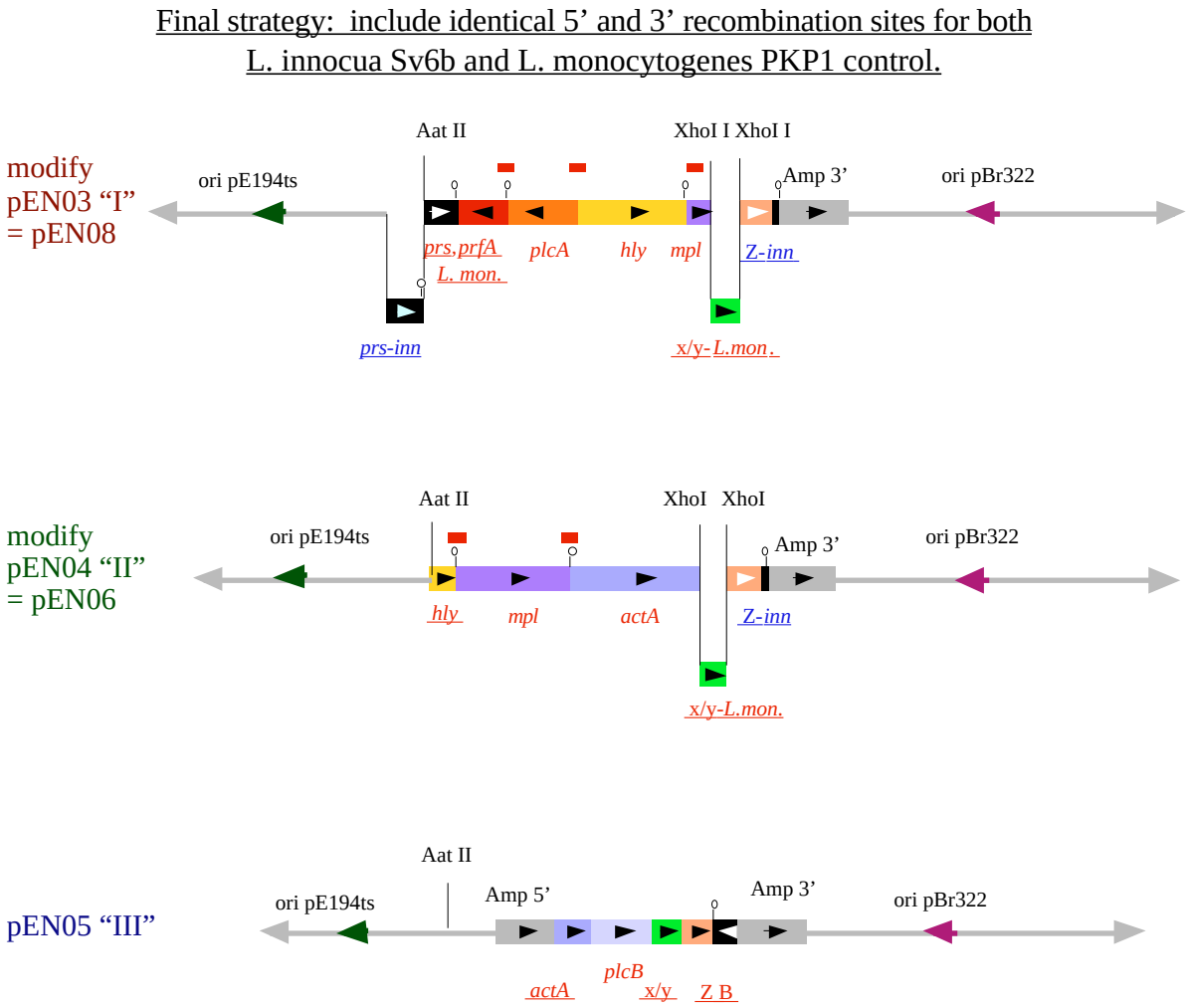


Figure 3.4. Final constructs carried on temperature sensitive shuttle plasmid pWH1509E contained all the required sequences for integration into the respective *Listeria spp*. Three sequential chromosomal double crossovers ("I" pEN08, "II" pEN06, "III" pEN05) will place the entire *L. monocytogenes* virulence cluster into *L. innocua* Sv6b and the positive control *L. monocytogenes* virulence gene cluster deletion mutant PKP1. Integration of "I" proved successful in both *L. innocua* and PKP1. Integrations of "II" and "III" are pending. (Sequences for chromosomal recombination are underlined. Red *indicates L. monocytogenes* sequence. Blue indicates *L. innocua* sequences. Black indicates sequences introduced in the prior established integrand).

Since it is essential that the same construct be integrated into the test and control pair, the plasmids pEN03 and pEN04 were modified such that each carried 5' and 3' integration sequences specific for both species. pEN05 needed no modification. Figure 3.4 illustrates the modifications made. Both "I", pEN03 and "II" pEN04 were co-modified to carry the *L. monocytogenes orfX* to allow 3' integration into PKP1. The 540bp *orfX* insert was generated by Taq polymerase (primers: *plcB*2(XhoI)> <*orfX*2(XhoI), digested by XhoI, and ligated into

pEN03 XhoI site. PCR checks confirmed the orientations (primers for "I": *hly*2> <*orfX*2(XhoI), "II": *ActA*2> <*orfX*2(XhoI) and proper insert numbers (primers for "I": *hly*2> <*orfZ*9-inn, "II": *ActA*2> <*orfZ*9-inn). These constructs were named pEN07 for "I", and pEN06 for "II". pEN07 needed further modification to include *L. innocua* specific *prs* sequence to form its 5' recombination site. A 640bp *L. innocua prs* insert was generated by Deep Vent polymerase (primers: *prs*8-inn(Aat2)> <*prs*9-inn(Aat2)), digested with AatII, and ligated into the AatII site of pEN07. The insert orientation and number was confirmed by PCR (primers: *prs*5(Aat2)> <*prfA*3). This final construct of "I" was named pEN08.

Because pEN08 was constructed out of 4 consecutive clonings, and now encompasses a 6.1Kb insert of listerial sequences, one final PCR test was used to confirm that the arrangement and contents were stably maintained (primers: *prs*8-inn(Aat2)> <*prfA*3, *prs*5-inn(Aat2)> <*prfA*3, *hly*2> <*orfX*2, *hly*2> <*orfZ*9-inn, and *plcB*2> <*orfZ*9-inn). It was.

Chromosomal integration of pEN08 carrying *prfA, plcA, hly* into *L. innocua* and PKP1

pEN08 was electroporated into *L. innocua* Sv6b and *L. monocytogenes* PKP1. Primary integration was effected in non-permissive temperatures under Erm selection. Five prime recombination events in *L. innocua* Sv6b and *L. monocytogenes* PKP1 were identified by PCR from small scale chromosomal DNA sa*mpl*es with respective primers sets: *prs*4-inn> <*prfA*2 and con-*prs*2> <*mpl*2(XhoI). Three prime recombination events were detected likewise in both species by primers *hly*2> <orfB10. Both 5' and 3' integration events were detected in Sv6b primary integrands, while only one 5' and no 3' integrands were found in PKP1. Primary integrands from each species were passaged 18 times. Second crossover events were identified by the loss of Erm resistance and subsequent PCR screen using the same methods as for the first crossover. Species identity was confirmed using species-specific primers that discriminate the *i-inlDC* locus of each species (*L. innocua* : DC1-inn> <emr2-inn, *L. monocytogenes*: DC2-egd> <emr1-egd). See Chapter 2 for a description of the region between *rrn* and *emr* in these species.

The resultant *L. monocytogenes* strain was named "PKP1+1", and the resultant *L. innocua* Sv6b strain was named COR1 (for "corrupted1"). *L. monocytogenes* PKP1+1, now complemented with the *plcA* and *hly* genes, fully recovered the hemolytic phenotype, indicating that pEN08 carried at least a workable copy of *hly*. But surprisingly, *L. innocua* COR1, with the newly acquired *prfA, plcA* and *hly,* did not gain the hemolytic phenotype as anticipated (Rauch, personal communications). At the time of this writing, the second and third integrations are still pending. Figure 3.5 illustrates the integration schemes.

Chromosomal integrations of "I", "II" and "III" into
L. innocua Sv6b and *L. monocytogenes* PKP1 control

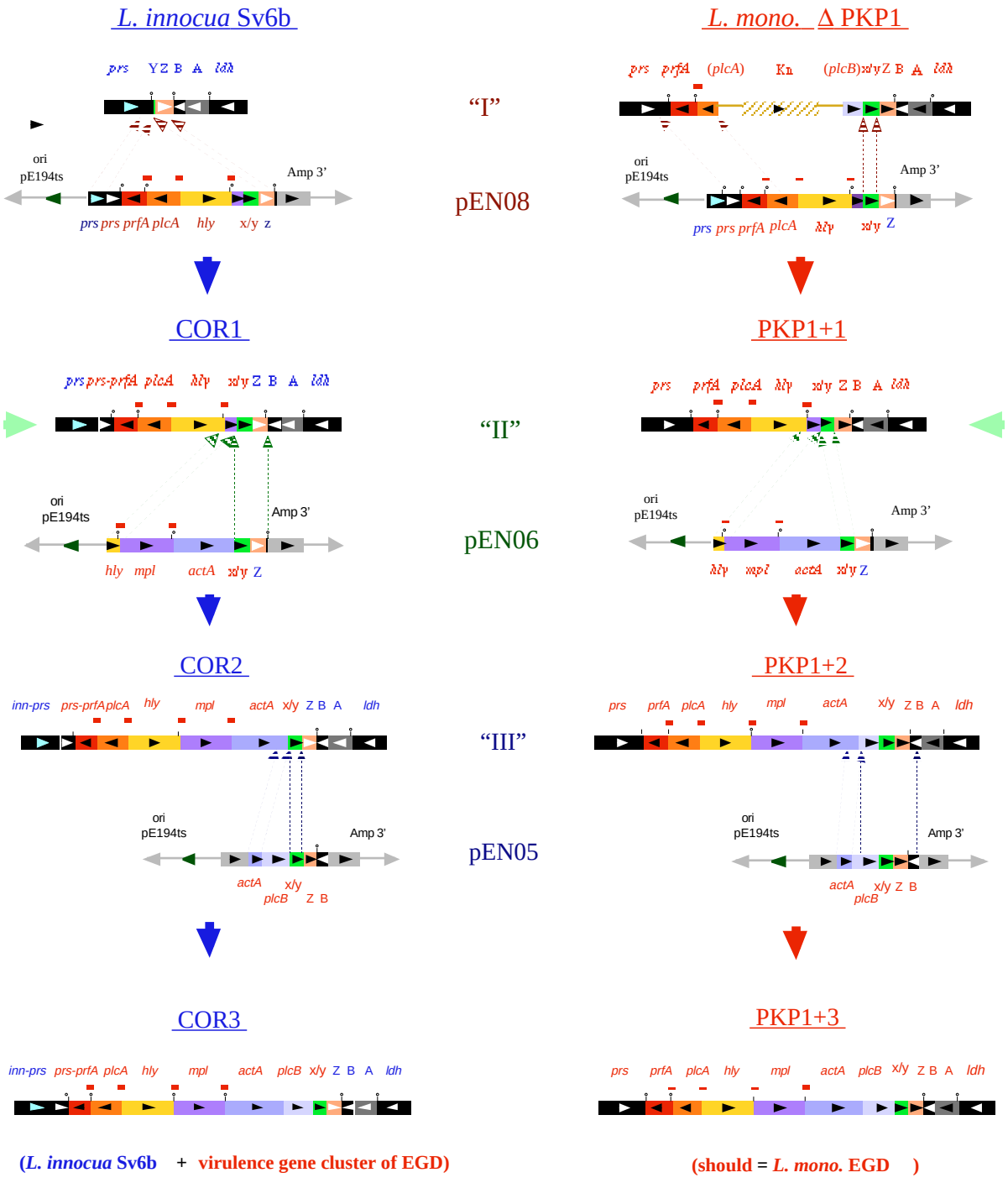


Figure 3.5. Successive chromosomal integrations to move *L. monocytogenes* EGD virulence gene cluster into *L. innocua* Sv6b, and parallel integrations into PKP1 as proof of function for the genes being moved. At this time, COR1 and PKP1+1 have been constructed. (Red alleles are of *L. monocytogenes* origin, blue alleles are of *L. innocua* origin).

DISCUSSION

After a seemingly sisyphian exercise, a working strategy for integrating the virulence gene cluster of *L. monocytogenes* into the *L. innocua* chromosome is finally at hand. All of the components for integration has been successfully cloned in *E. coli* intermediates. The first segment "I" conferring functions mediated by PrfA, PlcA, and Hly has been integrated into *L. innocua* (COR1), and the positive control, the *L. monocytogenes* deletion mutant, PKP1. The second segment "II" was also readily transformable into *L. innocua* (data not shown).

The difficulties encountered with cloning listerial DNA via *E. coli* intermediates is well known. Here, the *vcl* had to be divided into three overlapping pieces of less than 5Kb each for successful cloning. Although not unanticipated, the limits of cross-species recombination is unknown in *Listeria* simply because of a lack of precedence. Sequence divergences over >500bp between *L. monocytogenes* and *L. innocua* of 7% for *prs* and 18% for *orfZ* were beyond tolerance levels of the reciprocal recipients, and are expectedly so in the light of recent findings.

The reluctance of inter-species homologous recombination, or sexual isolation, is observed to be exponentially related to sequence divergence in all Gram negative and Gram positive organisms examined (Majewski & Cohan 1999; Majewski et al 2000). These are primarily due to mismatch repair systems (MutSL) and the scarcity of "minimum efficiently processed segments", which are short regions of near identities between donor and recipient DNAs required for successful strand invasion to initiate recombination. Surprisingly, the contribution of either mechanism to sexual isolation vary largely among organisms. For example, mismatch repair forms the predominant recombination barrier between different species of Gram negative bacteria such *Salmonella spp*. and *E. coli* (Vulic et al 1997), but is less important in *Streptococcus pneumoniae* (Majewski et al 2000), and is negligible in *Bacillus subtilis* (Majewski & Cohan 1998). On the other hand, *B. subtilis* requires identity between donor and recipient at both ends of the donor DNA while *E. coli* only requires identity at the 3' invading end (Majewski & Cohan 1999), but *B. subtilis* is tolerant of very divergent DNA as long as enough identical flanking sequences of a required stability (melting temperature) are available. The extent to which *Listeria* is similar to *B. subtilis* in mechanisms of sexual isolation is not known. Nonetheless, these studies were done using conjugation and transduction in Gram negative bacteria, and natural competence in the Gram positive bacteria cited. Since the foreign DNA substrates are presented in single strand form through these uptake mechanisms, they largely escape the activity of the host restriction-modification systems. Such mechanisms of DNA uptake are not observed in *Listeria,* thus sexual isolation is likely even more profound in *Listeria* than in its S. *pneumoniae* and *B. subtilis* relatives.

Potential Uses:

Although COR1, the *L. innocua* construct carrying *prfA, plcA* and *hly* genes, does not possess the full complement of the virulence gene cluster, it is already sufficient as a testing ground for functions related to the regulation of PrfA, with Hly acting as the reporter. A question still remains whether *prfA* is fully functional since COR1 is questionably hemolytic while the positive control PKP1+1 regained its hemolytic activity. The only difference between these construction processes is that PKP1 possessed its own *prfA* gene prior to integration of pEN08. This can be directly addressed by the sequencing of PCR products derived from the chromosomal fragment spanning the *plcA* promoter to the *prfA* structural gene of COR1. If no meaningful sequence changes are observed in the structural *prfA* gene or in the promoters of *prfA* and *plcA*, which drive *prfA* transcription, this non-hemolytic phenotype can be attributed to regulatory differences in *L. innocua*. Interestingly, *prfA-plcA-hly* constructs introduced into *L. innocua* on multiple copy number plasmids can express hemolytic activity and can escape primary phagosomal compartments (Goebel and Götz, personal communications). This implies that under those specific conditions, the number of PrfA molecules present in *L. innocua* was at sufficient levels in the 'activated' form to drive virulence gene expression. In addition, *L. innocua* transformed with plasmids containing the *prfA* can be triggered to turn on Prf*A* production upon exposure to eukaryotic factors (Renzoni et al 1999). This latter case indicate that the sensor for eukaryotic contact exists in *L. innocua* to signal virulence gene production. If the *prfA* gene of COR1 is not defective, then these observed differences between an integral copy versus multicopies reflect mechanisms governing the sensitivity of regulation of PrfA activity.

<u>Identifying regulators of virulence gene expression:</u>

Many pieces of evidences in *L. monocytogenes* show that the abundance of PrfA per se does not translate to PrfA activity (Klarsfeld et al 1994; Renzoni et al 1997). This apparent co-regulation via other means are under the influences of environmental factors including stress (Ripio et al 1998; Sokolovic et al 1993) and medium composition enhancement of virulence gene expression (Bohne et al 1996; Ripio et al 1996), catabolite repression of virulence genes under growth of certain carbon sources (Milenbachs et al 1997; Park & Kroll 1993), and activation of PrfA upon encountering eukaryotic cells (Renzoni et al 1999). Moreover, *in vitro* transcription assays using purified RNA polymerases isolated in different growth conditions show that this co-regulation may be mediated via alternative components of RNA polymerase (Lilac and Goebel, 2000, personal communications) and (Bockmann et al 2000). In this light, alternative sigma factor loading of RNA polymerase seems highly plausible since they form well documented, developmental switches in bacteria, governing life

cycle events such as entry into stationary phase (σS in E. coli and σH in B. subtilis) and the varying stages of sporulation in *B. subtilis* (σH, σF, σE, σG, σK); reviewed in (Wosten 1998). A likely candidate would be an alternative sigma factor linked to stress response, but unexpectedly, the general stress response sigma factor, σB, is probably not responsible for the expression of virulence genes in *Listeria* as its inactivation did not affect virulence in mice (Wiedmann M 1998).

An *in silico* search of the *L. monocytogenes* genome for alternative sigma factors (described in Chapter 4) yielded a total of 5 sigma factors. Beside the major σA and the general stress σB, homologues to σL, σH and an extracytoplasmic type factor (referred to here as σECF) was identified. σL regulates the degradative levanase operon in *B. subtilis* and *L. monocytogenes* and is under investigation by another group (Robichon et al 1997). We have chosen to further investigate the σH and σECF.

The prototype SigH, encoded by *rpoH* (*sigH*) in *B. subtilis*, is particularly active in transition phase gene expression although it is also produced during growth phase. SigH is crucial for the induction of the mutually exclusive pathways of competence development or sporulation, and also for DNA damage inducible response in *Bacillus*. Pathogenic adaptation of *Listeria* can perhaps be regarded as a transition function between different lifestyles, and may be subjected to similar regulation. *Bacillus* SigH is regulated via transcriptional suppression by AbrB, a global regulator suppressing post-exponential phase functions during active growth; and also by increased message stability and efficiency of translation upon entry into stationary phase (Haldenwang 1995). Several factors are known to control SigH post transcriptionally. These includes specific degradation upon acid stress by LonAB protease (Liu et al 1999), degradation upon exiting the initial stages of stationary phase by ClpC regulatory ATPase / chaperone (Nanamiya et al 1998), and direct activation of SigH dependent transcription via unknown means by ClpX regulatory ATPase / chaperone (Liu & Zuber 2000). ClpX is the first chaperone to be directly implicated in transcriptional activation in conjunction with a sigma factor. Since activation of PrfA controlled genes seems also to require an activator besides PrfA, a chaperone should also be considered in this role.

The second candidate resemble ECF sigma factors, which are named for their involvement in controlling extracytoplasmic functions ranging from heavy metal efflux, iron scavenging, virulence, to cell envelope maintenance in both Gram positive and negative bacteria. They are usually associated with membrane bound proteins which presumably act as sensors and as their anti-sigma factor. Upon receiving the appropriate signal, the membrane sensor releases the σECF for transcription activities (Missiakas & Raina 1998). This type of regulation also appear plausible for the control of PrfA. Signals such as MEM shift, charcoal, or host cell factors may cause release of active σECF which activate PrfA dependent transcription.

The construction of knockout mutants of σH and σECF in *L. monocytogenes* are underway to assess the role of these sigma factors on virulence gene expression (done by Michael Herler and Marcus Rauch). Complementing the knock out strategy is the cloning of σH and σECF under a constitutively expressed promoter (P_{SOD}) in a multi-copy shuttle vector (data not shown). The rationale of the latter approach is first to make either σH or σECF abundantly available, and to test if this availability cause activation of PrfA in conditions where PrfA is not normally activated. Alternatively, in-vitro assays of run-off transcription can be assessed in these extracts, with the additon of other factors, such as Clp chaperones, to directly assay transcriptional activity.

If COR1's lack of hemolytic activity is genuinely caused by regulatory phenomena, then it can be used to assay for hemolysin activation upon introduction of either of these σH and σECF constructs, or for that matter, any other positively acting candidate regulators. These putative regulatory factors can exert their influence indirectly on PrfA via modification of other factors or PrfA itself. But whether directly or indirectly, if a sigma factor exerts a positive effect on virulence gene expression, it would hopefully be detectable in COR1 or COR1 extracts.

What roles do internalins play in vivo?

How much known genes contribute to virulence *in vivo* are still enigmas because clear *in vitro* results such as lack of invasiveness of mammalian cells or inability to spread from cell to cell do not directly translate to the capacity for bacteria to survive *in vivo*. For example, mutants defective in gene products ActA and PlcB, important for intracellular spreading, and mutants defective in InlB and InlA, which confer invasiveness to extracellular bacteria were both found to be alive and replicating intracellularly in mouse livers (Appelberg & Leal 2000). Clearly, multiple components are required for a complex trait such as the ability to persist in the host liver. How much each component contribute to the overall phenotype can be tested when individual components are selectively added to the COR constructs.

Chapter 4

The Genome of *L. monocytogenes*

The genome of *L. monocytogenes*: Introduction

The designation *L. monocytogenes* represents a diverse population. Comprehensive surveys of *L. monocytogenes* isolates collected worldwide from clinical and environmental sources grouped the population into three distinct lineages, which encompasses all 13 known serogroups. Group I contains serotype b strains derived from human and animal sources. All food-borne epidemic isolates fell into this group. In particular, serotype 4b is associated with 64% of the epidemic outbreaks in humans, and serotype 1/2b and 4b combined accounts for 74% of all human isolates (McLauchlin 1990). Group II contains serotype a and c strains derived from human and animal sources not associated with epidemics. Group III contains only non-human derived isolates and is associated with serotype 4a. These groupings are consistent with all existing evidence of typing by flagella antigens, multilocus enzyme electrophoresis (MLEE), randomly amplified polymorphic DNA, restriction fragment linked polymorphism (RFLP) of *hly* and *inlA*, DNA sequence typing of *actA, hly, iap* and flagellin (*flaA*), and chromosomal pulsed-field gel electrophoresis (Piffaretti et al 1989; Rasmussen et al 1995; Wiedmann et al 1997). Cumulatively to date, all the studies showed strong linkage disequilibrium among the various *loci* examined which indicate a marked lack of horizontal gene transfer among the separate lineages, and that the population structure of *L. monocytogenes* is primarily clonal.

Strain EGDe of serotype 1/2a is being sequenced by the Consortium. This particular isolate was clinically derived and has undergone numerous passages in mice. It falls into lineage II of the above-described scheme. It is not known how different the present EGDe strain is from the original clinical isolate. This EGD isolate was chosen because passages in mice ensured the presence of genes required for the virulence phenotype.

Chromosomal maps of three *L. monocytogenes* strains are available. Lineage II strains EGDe (genome size 3.0 Mb (von Both et al 1999)) and Lo28 (serotype 1/2c, genome size 3.15 Mb (Michel & Cossart 1992)) exhibited the same chromosomal map. However, lineage I strain Scott A (serotype 4b, genome size 3.21 Mb (He & Luchansky 1997)) showed different mapping patterns from the other two. These gross differences in the chromosomal maps may be due to genome rearrangements or the presence of different lysogenized phages in addition to differences in genetic backgrounds.

The sequencing of the *L. monocytogenes* genome was the effort of the European *Listeria* sequencing project Consortium. The Consortium consisted of ten laboratories in France, Spain and Germany and was centrally coordinated at the Pasteur Institute in Paris by Pascal Cossart and Philippe Glaser. See the list of the Consortium participants on page 113. Our lab represented 10% of the effort in gap closing, sequence verification, and annotation. The project officially commenced in the Spring of 1998 with library construction in Paris. The

genomic DNA of strain EGDe was cloned as libraries varying from small (1-2Kb), medium (10-20Kb), to large bacterial artificial chromosome (BAC, 60-70Kb) clones in *E. coli* hosts. The LION and Charkraborty groups constructed the BAC libraries. The shotgun-sequencing phase employed primarily the small insert library, and began in the Fall of 1998 and was concluded in the Spring of 1999. Sequencing was done from both ends of each insert clone. The Phred-Phrap-Consed Software written by Phil Green at the University of Washington was used as the tool for sequence quality assessment and global contig assembly. At the beginning of closure phase, greater than 40,000 sequences representing 6.5X chromosomal coverage were obtained by the Consortium and these were assembled into approximately 250 contigs--depending on the stringency of the assembly parameters (Phred score requirement). Gaps were predicted by a mixture of methods including inference from medium size library clones and BAC clones, sequence comparison with known *B. subtilis* genome sequence, and by recombinatorial PCR of unmatched contig ends. These activities were handled by Philippe Glaser and Lionel Franguel at Pasteur. All consortium members took part in the annotation of the *L. monocytogenes* genome sequence. Thereafter, the Pasteur Institute used the annotated *L. monocytogenes* sequence to facilitate the assembly, annotation, and the comparative analysis of the *L. innocua* genome.

My role in this project was in the planning, implementation and coordination of our lab's effort in gap closure, sequencing, and the preparation of the annotation report of *L. monocytogenes*, particularly in the area of post-exponential phase functions. Most of this took place between May 1999 through July 2000. The manuscript on the genome comparison of *L. monocytogenes* and *L. innocua* is currently being prepared by Philippe Glaser and the Consortium. Once published, the respective sequences will be publicly accessible via the Pasteur Institute website (http://genolist.pasteur.fr/).

RESULTS:

Gap closure phase began with the successful assembly of approximately 200-250 contigs. Predicted gaps were tested by PCR from EGDe chromosomal DNA template using primers designed from the two predicted contig ends. When a unique PCR product was obtained, the fragment was sequenced in full to close the gap. Predicted gaps that failed to yield unique PCR products under various PCR conditions were erroneous predictions. Most gaps were smaller than 3Kb in size. Besides gaps closings, low quality sequence areas of assembled contigs were likewise verified by bridging the low quality area with PCR from chromosomal DNA template, and sequencing of the products.

We attempted two sequencing approaches, but only one was successful. Sequence output obtained from the ABI systems was compatible with the Phred-Phrap assembly software but output obtained from the Licor system was not optimally adapted, and the Licor

system was thus abandoned. In addition, we also had technical problems with primer quality for the Licor IRD-800 dye system. The sequence outputs from ABI systems were processed with locally installed Phred-Phrep-Consed software to ensure that the sequence quality met with the required standard of Phred Score >19. Sequences on both strands meeting the required Phred score were needed to close each gap. In total, we processed 22 gaps and 41 verifications, accounting for 285 ABI sequences.

Annotation proceeded prior the final closure of the genome. Automated bioinformatic searches for open reading frame (ORF) prediction, ORF sequence comparison and identification (e.g. Blast results) for all ORFs above 70 amino acids and known, smaller gene products were done by the Pasteur Institute and LION bioscience AG. Automated results were subsequently verified by all Consortium members between December 1999 through July 2000. Each group was assigned certain biological areas in which to finalize the annotation and extract information. My task was stationary phase functions. In addition, in light of our institute's interest in mechanisms of PrfA control, a sigma factor study was done.

The *Listeria monocytogenes* genome:

The genome was closed when the final remaining contig (chromosome) was circularized. According to the results of the Consortium, the total length of the circular *L. monocytogenes* EGDe genome is 2,944Kb. It contains one phage (129 Kb phage A118, inserted in *comK*) and three copies of an IS element. At a coding density of 90%, it encodes estimated 2932 genes with an average size of 912bp. There are 6 separately located ribosomal operons, 41 surface anchored (LPXTG) genes of which 19 were internalin-like, 61 other lipoproteins, and most remarkably, *L. monocytogenes* possess 42 phospho-transfer systems (PTS) accounting for 6% of the total genome. The circular map in Figure 4.4 illustrates the bi-directional origin of replication and various *loci*, including virulence genes.

Sigma Factor Search:

In order to identify all coding sequences resembling known sigma factors in *L. monocytogenes*, an exhaustive search was performed using each sigma factor sequence identified in *B. subtilis* against the *L. monocytogenes* genome. *B. subtilis* is the closest relative with a sequenced genome to *L. monocytogenes*, and possesses the most prolific and well-studied list of sigma factors. This procedure was repeated with all known *E. coli* sigma factors. The resulting *L. monocytogenes* genes identified as sigma factor-like were used to search against the general database to confirm its identity, and the *L. monocytogenes* bank to further identify any paralogs that might have escaped earlier detection.

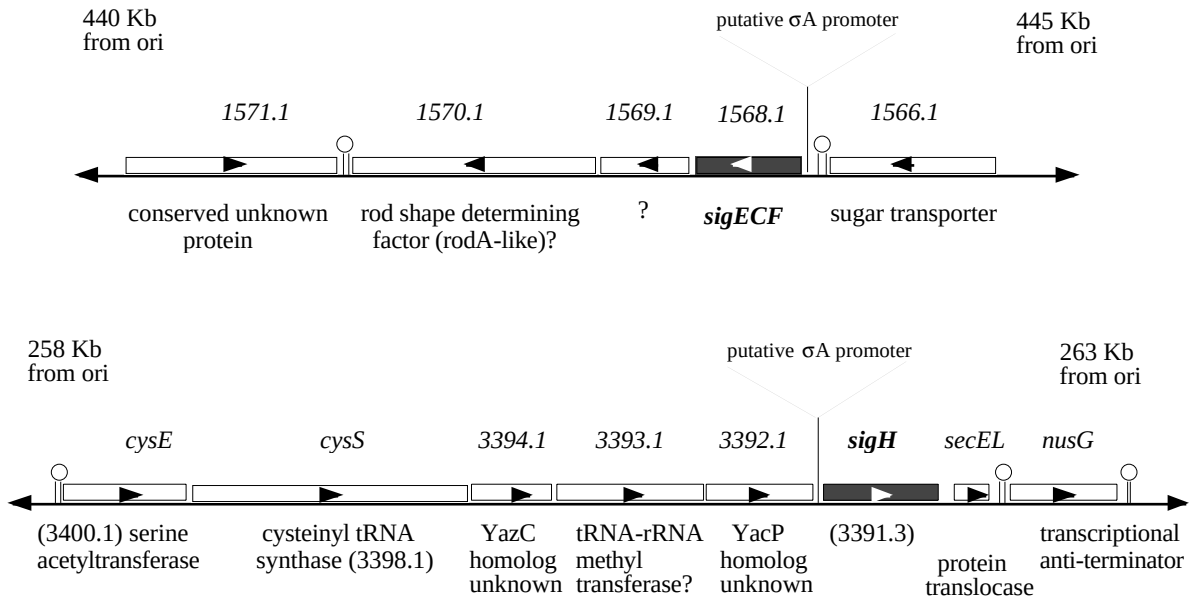
Surprisingly, only five sigma factors were identified in *L. monocytogenes*, compared to the 18 identified from the *B. subtilis* genome (http://genolist.pasteur.fr/SubtiList/) and the

seven known in *E. coli* (http://genolist.pasteur.fr/Colibri/). Three of these, SigA, SigB and SigL, have been previously identified in *L. monocytogenes*. The two unknowns fall into the σ70 class of sigma factors. They respectively resemble SigH and an extra-cytoplasmic factor (ECF)-type sigma factor similar to SigW or SigV of *B. subtilis*. Figure 4.1 shows alignments with the closest known, extracytoplasmic function (ECF) type, sigma factor homologs. Figure 4.2 shows alignments to the closest known SigH homologs. Table 4.1 shows the entire search findings. Figure 4.3 illustrates the chromosomal context of these genes.

Table 4.1. Sigma factors in *L. monocytogenes*.

Sigma factor	Alternative name	Size	Amino acid identity to *B. subtilis* ortholog	Previously reported
SigA	RpoD, σ43	374aa	81% (primary σ)	(Metzger et al 1994)
SigB	RpoF, σ37	259aa	67% (general stress σ)	(Becker et al 1998; Wiedmann M 1998)
SigL	RpoN, σ54	447aa	37% (alternative, levanase)	(Robichon et al 1997)
SigH	RpoH, σ30	201aa	45% (transition phase σ)	--
Sig-ECF type	(LiM0 1568.1)	166aa	29% to σV or σW	--

Figure 4.3. Chromosomal context and locations of *sigH* and *sigECF*. (Genes are designated by their individual protein file (IPF) numbers unless function can be definitively assigned. All other *L. monocytogenes* genes presently appearing on the Pasteur's ListiList website are preceded by LiM0, e.g. LiM0 1568.1).



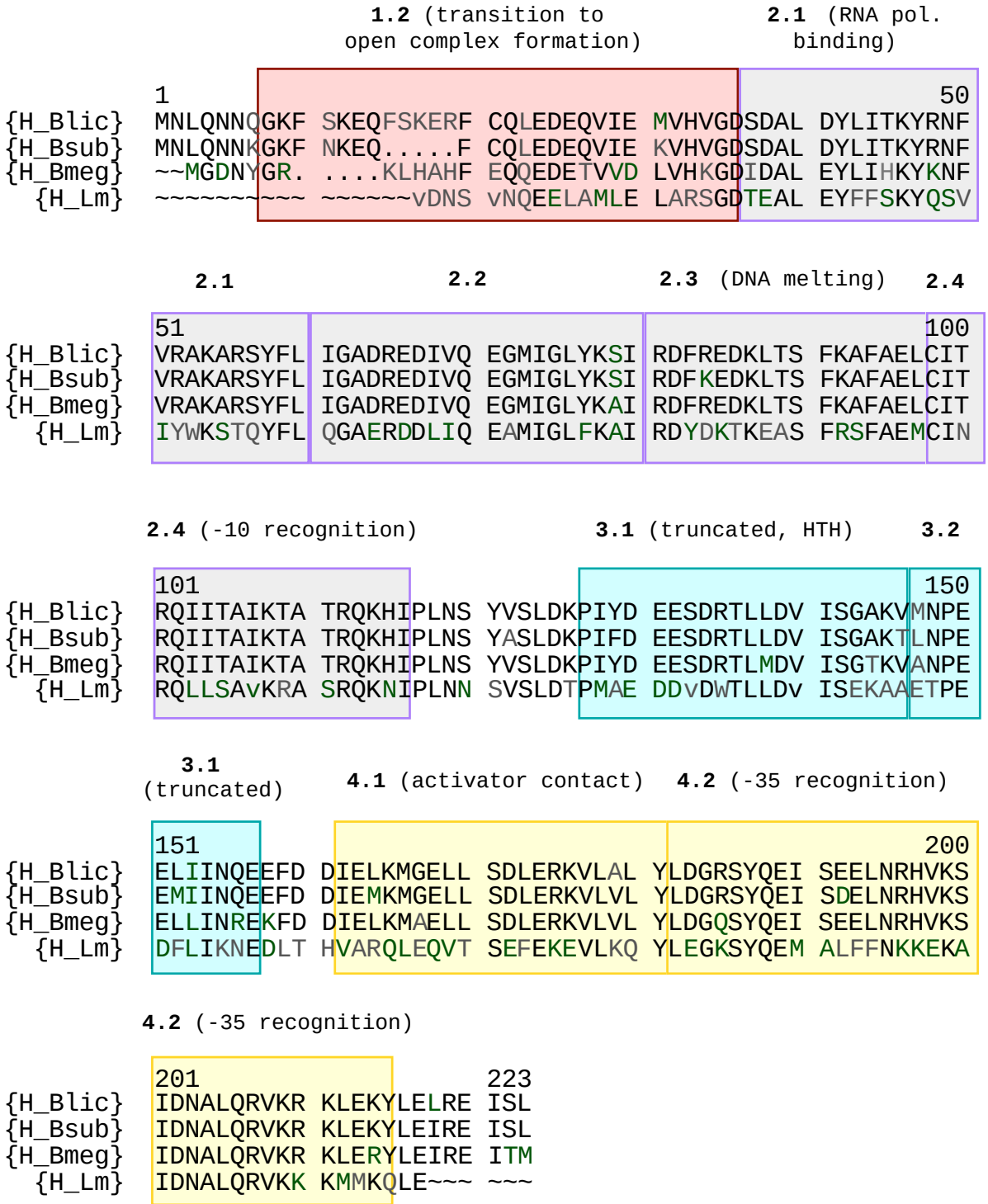


Figure 4.2. Alignment of sigma factor H of *L. monocytogenes* with σH of *Bacillus spp.* (*B. licheniformis* M29694, *B. subtilis* M29693, *B. megatherium* X59070). Functional domain designations are based on alignments with known σ70 structures (Lonetto et al 1994; Wosten 1998). Amino acid identities are shown in black, similarities in green.
HTH = helix-turn-helix, DNA binding motif.

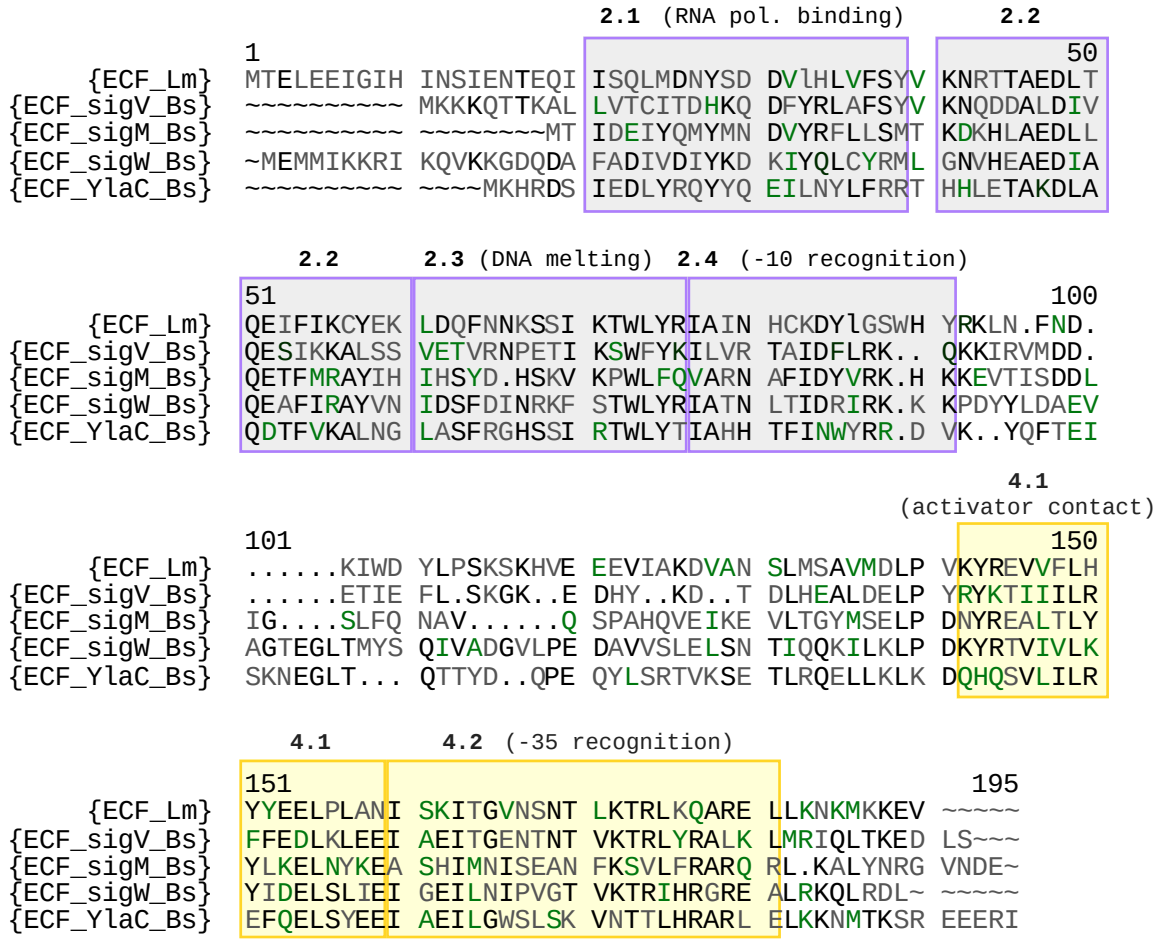


Figure 4.1. Alignment of ECF type sigma factor (gene designation LiM01568.1) of *L. monocytogenes* with known ECF σV, σM, σW, and ECFσ-like YlaC of *B. subtilis* (http://genolist.pasteur.fr/SubtiList/). Functional domain designations are based on alignments with known σ70 structures (Lonetto et al 1994; Sorokin et al 1997; Wosten 1998). Amino acid identities are shown in black, similarities in green.

Stationary phase genes in *L. monocytogenes* as compared with *B. subtilis* and *E. coli*:

For the purpose of this report, stationary phase is defined as any other time when the cell is not growing exponentially. This can be induced by gradual nutrient limitation or by sudden environmental changes such as temperature, chemical, or osmotic shocks. Because an organism's entry into stationary phase can be precipitated by multiple causes, leading to various differential pathways, this report contains but does not exhaustively define all the genes representing each of the further differentiation states. For example, during entry into stationary phase, *B. subtilis* can commit into several differentiation paths: competence

development and sporulation are mutually exclusive decisions, but the development of motility, the production and secretion of antibiotics and degradative enzymes are not exclusive functions. This report attempts to identify the presence or absence of the most important genes actively involved in the decision window, but does not try to list every gene of the competence machinery, or sporulation program, or osmotic shock response, etc. These further specific functions are presented by other individuals of the annotation team.

In order to identify genes that are potentially important for stationary phase functions in *L. monocytogenes*, *B. subtilis* and *E. coli* were used as reference organisms. Key word searches against the literature and the respective *B. subtilis* and *E. coli* genomes include: post-exponential, stationary phase, transition phase, diauxic shock, and other shock or stress responses. Genes that are known to be specific and/or important for post-exponential growth and survival formed a set of reference genes to be searched against the *L. monocytogenes* genome. In addition, the same key words were used to scan the first pass annotation results of the *L. monocytogenes* data. This allows the overlapping set of genes to confirm each other, as well as to identify potential stationary phase genes in *L. monocytogenes* not present in *B. subtilis* and *E. coli,* but are known in other organisms. Most of the genes presented here were identified directly from known stationary-phase genes. The first pass annotation of the *L. monocytogenes* genome offered scant information.

Table 4.2 catalogs the presence and absence of genes potentially involved in stationary--transition phase management in *L. monocytogenes*. The primary function and a brief description are provided. Potential paralogs within *L. monocytogenes* are also identified. *L. monocytogenes* genes are presented as their individual protein file (IPF) number. In addition to the annotation provided by Subtilist (*B. subtilis* database) and Colibri (*E. coli* database), references used to define the list include the following: (Antelmann et al 1997; Becker et al 1998; Braun et al 1996; Dons et al 1994; Flanary et al 1999; Fuge et al 1994; Galsworthy et al 1990; Ishihama 1997; Lazazzera 2000; Msadek 1999; Padilla et al 1998; Peel et al 1988; Sivasubramaniam et al 1995; Strauch 1993). References of *L. monocytogenes* homologs described in this list include the following: (Borezee et al 2000a; Borezee et al 2000b; Gaillot et al 2000; Nair et al 2000a; Nair et al 1999; Nair et al 2000b; Rouquette et al 1998).

Table 4.2. Stationary phase genes of *L. monocytogenes*

GENE	in *L. mono*? IPF no.	in *B.sub*?	FUNCTION	DESCRIPTION	Paralogs in *L. monocytogenes*
abrB	978_1	**y**	transition states regulator	Transcription regulator: regulation of transition state genes. Represses most post-exponential genes during exponential growth, activate others at transition phase	Unique in *L.mono*
ahpC	no	**y**	general stress, H_2O_2, stationary phase	*B.sub* class III stress response, alkyl hydroperoxide reductase. See Oxidative Stress Report.	Ortholog not in *L.mono*, paralog 2875.1 (YkuU)
aprE	no	**y**	stationary phase only	Serine alkaline protease (subtilisinE), extracellular	
cheA	652_1	**y**	chemotaxis	2-component histidine kinase sensor, regulates expression of chemotaxis genes, SigD regulated in *B.sub*. *L.mono*: no *sigD*, role in stationary phase unknown.	
cheY	651_1	**y**	chemotaxis	2component response regulator, regulates expression of chemotaxis genes, *sigD* regulated in *B.sub*. *L.mono*: no *sigD*, role in stationary phase unknown.	
clpC	3135_1	**y**	transition phase	class III stress response-related ATPase, chaperonin. In *L. mono*, ClpC is required for heat, salt, iron limitation, oxidative stress, and intracellular survival in macrophages, adhesion and invasion of mouse hepatocytes, affects transcription of virulence genes (inlAB,ActA tested to date). *clpC* mutant: 2log increase in LD50.	Paralogs in *L.mono*: 3635.2, 866aa=ClpB (*E. coli*), 2338.2, 748aa=ClpE
clpE	2338_2	**y**	heat shock	Chaperonin. In *B.sub*, *ctsR* heat shock regulon, not required for stress tolerance, induced by heat, puromycin, similar to mecB. In *L. mono*, regulates pleiotropic stress functions including heat stress response, cell division and virulence. *clpE* mutant: 2log increase in LD50	
clpP	754_1	**y**	stationary phase	Serine Protease. In *B.sub*: activity induced by heat, EtOH, salt. Important for stress, heat, starvation, regulation of competence, motility, degradative enzyme synthesis, sporulation. In *L. mono*: heat, salt stress, macrophage survival, virulence. *clpP* mutant: 3.5log increase in LD50	Paralog in *L.mono* 224.1 (220aa) 41%ID, 60%+.
clpQ	2901_1	**y**	heat shock	Protease. Possible beta-type subunit of the 20S proteasome, similar to the heat shock protein HslV of *E. coli*	

GENE	in *L. mono*? IPF no.	in *B.sub*?	FUNCTION	DESCRIPTION	Paralogs in *L. monocytogenes*
clpX	3722.1	**y**	stationary phase	ATP-dependent Clp protease ATP-binding subunit. In *B.sub*, needed for stress, starvation response, essential for sporulation and competence.	
comA	no	**y**	competence, sporulation regulation	2component DNA binding response regulator In *B.sub*, respond to *comP, comX, comQ* mediated quorum sensing.	
comF	3162_1	**y**	competence	DNA uptake, requires ComK for transcription; similar to the DEAD family of ATP-dependent RNA/DNA helicases; similar to PriA of *E. coli*	
comK	495_1	**y**	competence regulation	In *B. subtilis*, final inducer of competence, regulated by ClpC, ClpP, MecA, ComA, ComP, ComS, DegU, DegS, AbrB and itself.	Unique in *L.mono*, inactive. First 49aa truncated by prophage A118
comP	no	**y**	competence, sporulation regulation	2component histidine kinase sensor for comA in *B.sub*; respond to ComX, ComQ mediated quorum sensing.	
comQ	no	**y**	competence regulation	Quorum sensing in *B.sub*. Produces ComX pheromone, which stimulates competence via ComP sensor and ComA response regulator	
comS	no	**y**	competence regulation	In *B.sub*, 48aa peptide induced in response to quorum-sensing and nutritional stress; required for ComK synthesis; releases active ComK from complexed ComK/MecA/ClpC and allows comK to stimulate competence	
comX	no	**y**	pheromone, competence regulation	*B.sub* pheromone, 10aa (produced by comQ); stimulates sporulation and competence	
cspB	4189.1	**y**	stationary, cold shock	Cold shock response, In *B.sub*, induced in entry to stationary phase, deletion of *cspB/C* leads to lysis during stationary phase.	Paralogs in *L.mono*: 4189.1, 4186.1, 4158.1
cspC	4158.1	**y**	stationary, cold shock	Cold shock response, In *B.sub*, induced in entry to stationary phase, deletion of *cspB/C* leads to lysis during stationary phase. Designated *cspA* or *cspL* in *L.mono*	Paralogs in *L.mono*: 4189.1, 4186.1, 4158.1
cspD	4186.1	**y**	cold shock, not stationary phase	Cold shock response, but in *B.sub*, not used in stationary phase	Paralogs in *L.mono*: 4189.1, 4186.1, 4158.1

GENE	in *L. mono*? IPF no.	in *B.sub* ?	FUNCTION	DESCRIPTION	Paralogs in *L. monocytogenes*
ctsR	3129.1	**y**	transition phase, shock	DNA binding protein. In *B.sub* and *L.mono*: Negatively regulates *clpC, clpE, clpP* transcription. *L.mono ctsR* constitutive mutant: 2log increase in LD50 by i.v. infection; no change for oral infection; *ctsR* mutant: no effect on virulence	
dacC = pbp	no	**y**	stationary phase	Penicillin binding protein, expressed only early stationary phase via *sigH*	Not in *L.mono*
degQ	no	**y**	degradative enzyme synthesis regulation	In *B.sub*, regulates hyperproduction of levansucrase and other extracellular degradative enzymes (AprE, NprE, AmyE), require DegU/S for effect	Not in *L.mono*
degS	no	**y**	stationary phase,	2component Histidine Kinase sensor. In *B.sub*, regulates competence, degradative enzyme synthesis, phosphorylates DegU, stimulated by high salt	Ortholog not in *L.mono*/ paralog in *L.mono*: 2839.1(YvqE)
degU	3158.1	**y**	stationary phase, competence, degradative enzyme synthesis	2-component response regulator. In *B.sub*, regulates competence, protease production, osmotic response; induced by high salt. Phosphorylated by DegS	Paralog in *L.mono*: 2838.1
dps (E.coli) ytkB (B.sub)	3846.1	**y**	general stress, stationary phase	Metallo-DNA binding-protecting protein, non-heme Fe Ferritin. In *B.sub*, induced by heat, EtOH, entry to stationary phase. Controlled by SigB. In *E. coli*, Dps alters DNA conformation to stimulate stationary phase sigma (SigS) transcription.	Unique in *L. mono*, paralog in *B. sub* MrgA
flgM	no	**y**	motility, flagellum synthesis	In *B.sub.*, coupling of late flagellar gene expression (SigD-dep.) to the assembly of the hook-basal body complex (SigA-dep.)	
fnr	2292_1	**y**	anaerobic stress	Transcriptional regulator. In *B.sub.*, induced by oxygen limitation via ResD	
kat(X)	727_1	**y**	stationary phase, oxidative stress	The only catalase in *L.mono*, more similar to the fore-spore specific KatX than KatA or KatB of *B.sub*.	
katA	no	**y**	stationary phase, oxidative stress	In *B.sub*, vegetative catalase1, induced by H_2O_2 and entry into stationary phase under Fe, Mn limitation; secreted in stationary phase in rich medium	
katB	no	**y**	general stress	In *B.sub*, catalase2; induced by heat, salt, EtOH, glucose starvation; not by oxidative stress, not in forespores	

GENE	in *L. mono*? IPF no.	in *B.sub*?	FUNCTION	DESCRIPTION	Paralogs in *L. monocytogenes*
lonA	no	**y**	stress response, transition phase.	Class III stress response-related ATP dependent protease, induced by stress. In B. sub, post-translationally regulate SigH concentration.	Paralogs in *L.mono*: Clp proteases, ClpB 2248, ClpC 0242.
lonB	no	**y**	stress response, transition phase.	Class III stress response-related ATP dependent protease, induced by stress. In *B.sub*, post-translationally regulate SigH concentration.	Paralogs in *L.mono*: 3136.2, 1116.4, ClpE protease 1011.
malL = yvdL	949.1	**y**	sationary phase, maltodextrin utilization	In *B.sub*, maltodextrin utilization; oligo-1,4-1,6-alpha-glucosidase; induced in stationary and by maltose, starch, amylose, glycogen; catabolite repression by glucose and fructose	Many paralogs in *L.mono*. closest is 3616.1 (TreA), and 1329.1, 1898.2
mcpC	no	**y**	stationary phase, chemotaxis	In *B.sub*, chemotaxis towards cysteine, pro, thr, gly,ser,lys,val,arg; SigD dependent, methyl accepting membrane receptor; induced by entry to stationary phase	Ortholog in *L.mono*, paralog in *L.mono* : 697.1 (TlpA)
mecA	799_2	**y**	stationary phase, competence regulation	In B. sub., negatively regulates competence via comK; ClpC/MecA/ComK or ComS complex degraded by ClpC/P when dissociated; activate autolysin (LytC, LytD) synthesis; stimulate sigD motility genes. In *L. mono*: also regulatory role, targets not well defined.	Unique in *L.mono*
nfrA = ywcG	1583_1	**y**	early stationary,	Nitroflavin reductase/ SigD dependent transcription in excess glucose and glutamate, essential	Possible paralogs in *L.mono*: 2486.1 247aa
oppA	3312.2	**y**	stationary phase, competence, sporulation,	In *B.sub.*, Quorum sensing. ABC transporter, di-peptide binding, internalizes CSF(from PhrC) pheromone, which stimulates competence, sporulation, surfactin and degradative enzyme production.	
oppB	3310_1	**y**	stationary phase, competence, sporulation,	In *B.sub.*, Quorum sensing. ABC transporter permease; internalizes CSF pheromone, which stimulates competence, sporulation, surfactin and degradative enzyme production.	
oppC	3309_1	**y**	stationary phase, competence, sporulation,	In *B.sub.*, Quorum sensing. ABC transporter permease; internalizes CSF pheromone, which stimulates competence, sporulation, surfactin and degradative enzyme production.	

GENE	in *L. mono*? IPF no.	in *B.sub*?	FUNCTION	DESCRIPTION	Paralogs in *L. monocytogenes*
oppD	3307_1	**y**	stationary phase, competence, sporulation,	In *B.sub*., Quorum sensing. ABC transporter ATPase; internalizes CSF pheromone, which stimulates competence, sporulation, surfactin and degradative enzyme production.	
oppF	3306_1	**y**	stationary phase, competence, sporulation,	In *B.sub*., Quorum sensing. ABC transporter ATPase; internalizes CSF pheromone, which stimulates competence, sporulation, surfactin and degradative enzyme production.	
phoP	795_1	**y**	Phosphate limitation	2component response regulator. In *B.sub*., regulates alkaline phosphatase synthesis and phosphate levels via *phoA, phoB, phoD, resABCDE, tagAB, tagDEF, tuaA-H*.	
phoR	794_1	**y**	Phosphate limitation	2component histidine kinase sensor. In *B.sub*., regulates alkaline phosphatase synthesis and phosphate levels via *phoA, phoB, phoD, resABCDE, tagAB, tagDEF, tuaA-H.*	
phrA	no	**y**	sporulation regulation, quorum sensing	Pheromone in *B.sub*, cleaved by signal peptidase I, exported, processed to the active penta-peptide inhibitor, and re-imported by the oligopeptide transport (Opp) system, inhibits RapA phosphatase activity	Not in *L.mono*
phrC	no	**y**	competence, sporulation, degradative enzyme synthesis, quorum sensing	Pheromone CSF (5aa) in *B.sub*, regulates competence gene expression, stimulates the ability of cells at low cell density to sporulate at high concentrations	Not in *L.mono*
ppK	no	**n**	stationary phase- *E. coli*	Polyphosphate kinase. Polyphosphate accumulation is essential for *E. coli* stationary phase survival	Not in *L. mono*
rapA	no	**y**	sporulation regulation, quorum sensing	Response regulator, aspartate phosphatase. In *B.sub*, repress sporulation, repressed by Spo0A (not dependent of AbrB), inhibited by PhrA pheromone, induced by the ComP/A signal transduction system; glucose starvation inducible	Not in *L.mono*
rapC	no	**y**	competence, sporulation, degradative enzyme quorum sensing	Response regulator, aspartate phosphatase. In *B.sub*, repress srfA surfactin production, inhibited by *phrC* product CSF pheromone	Not in *L.mono*

GENE	in *L. mono*? IPF no.	in *B.sub*?	FUNCTION	DESCRIPTION	Paralogs in *L. monocytogenes*
resD	3169_1	**y**	aerobic/anaerobic response	2component response regulator (of ResE) in *B.sub*. for aerobic/anaerobic respiration; induced in phosphate limitation; repressed by Spo0A; autoactivate, induces PhoP, suppress PhoP, ResD during sporulation	
resE	3168_1	**y**	aerobic/anaerobic response	2component histidine kinase sensor in *B.sub*. for aerobic/anaerobic respiration; induced in phosphate limitation; repressed by Spo0A; autoactivate, induces PhoP, suppress PhoP, ResD during sporulation	
rmf	no	**n**	stationary phase-*E. coli*	Important in *E. coli* stationary phase: RMFcomplexed with 70S ribosomes render them inactive during stationary phase.	Not in *L. mono*
sacB	no	**y**	degradative enzyme synthesis	Levansucrase; modified by sucrose, DegU, DegQ	Not in *L.mono*
secA	2834_2	**y**	exponential phase secretion, transition phase secretion	Translocase ATPase. In *B.sub*, highest expression (by SigA) at transition phase and greatly increase secretory activity	Paralog in *L.mono*: 2741.1
sigB	1287_1	**y**	alternative sigma factor: general stress, stationary phase	Sigma37. In *B.sub*, activated by either a drop in intracellular ATP or exposure to environmental stress	Unique in *L.mono*
sigD	no	**y**	alternative sigma factor: motility, chemotaxis, autolysis	Sigma28. In *B.sub*, its regulon includes genes involved in flagellar synthesis, motility, chemotaxis, autolysis	Not in *L.mono*
sigH	3391_3	**y**	alternative sigma factor: transition phase	Sigma30. In *B.sub*, appears in early sporulation, repressed by a mechanism responding to amino acids levels; post-translationally negatively regulated in response to external low pH	Unique in *L.mono*
sigS = ropS	no	**n**	*E. coli* alternative sigma factor: stationary phase	Transcription regulator. In *E. coli*, globally activates stationary phase specific genes.	Not in *L. mono*
sinR	no	**y**	stationary phase	Transcription regulator. In *B.sub*, stimulates competence, subtilisin and flagellar synthesis, autolysin; represses sporulation.	Not in *L.mono*

GENE	in *L. mono*? IPF no.	in *B.sub*?	FUNCTION	DESCRIPTION	Paralogs in *L. monocytogenes*
spo0A	no	**y**	sporulation	2-component global response regulator of sporulation in *B.sub*. Represses *abrB, kinA, kinC, spo0A*, stimulates *spoIIA, spoIIE, spoIIG*	Not in *L.mono*
srfAA	no	**y**	competence, surfactin synthesis regulation	In *B.sub*, contains *comS*, stimulated by DegU	Not in *L.mono*/ domains found in fatty acid phospholipid metabolic genes
srfAB	no	**y**	competence, surfactin synthesis regulation	In *B.sub*, stimulated by DegU; srfAB gene contains comS (48aa)	Not in *L.mono*/ domains found in fatty acid phospholipid metabolic genes
srfAC	no	**y**	competence, surfactin synthesis regulation	In *B.sub*, stimulated by DegU	Not in *L.mono*/ domains found in fatty acid phospholipid metabolic genes
srfAD	no	**y**	competence, surfactin synthesis regulation	In *B.sub*, stimulated by DegU	Not in *L.mono*
tlpA	697_1	**y**	chemotaxis	Chemotaxis protein, methyl accepting transmembrane receptor	Unique in *L.mono*/ domain overlap in 3649.1
yaaD	1933_2	**y**	stationary phase, diauxic shock	Highly conserved stress response gene in bacteria, yeast, plants. In Yeast, stationary phase induced (SNZ), coregulated with SNO (YaaE homolog). Oxidative stress induced in *B.sub*. Ethylene stress induced in Para rubber trees	Unique in *L.mono*
yaaE	1932_1	**y**	stationary phase, diauxic shock	Highly conserved stress response gene in bacteria, yeast, plants. In Yeast, stationary phase induced (SNO), coregulated with SNZ (YaaD homolog).	Unique in *L.mono*

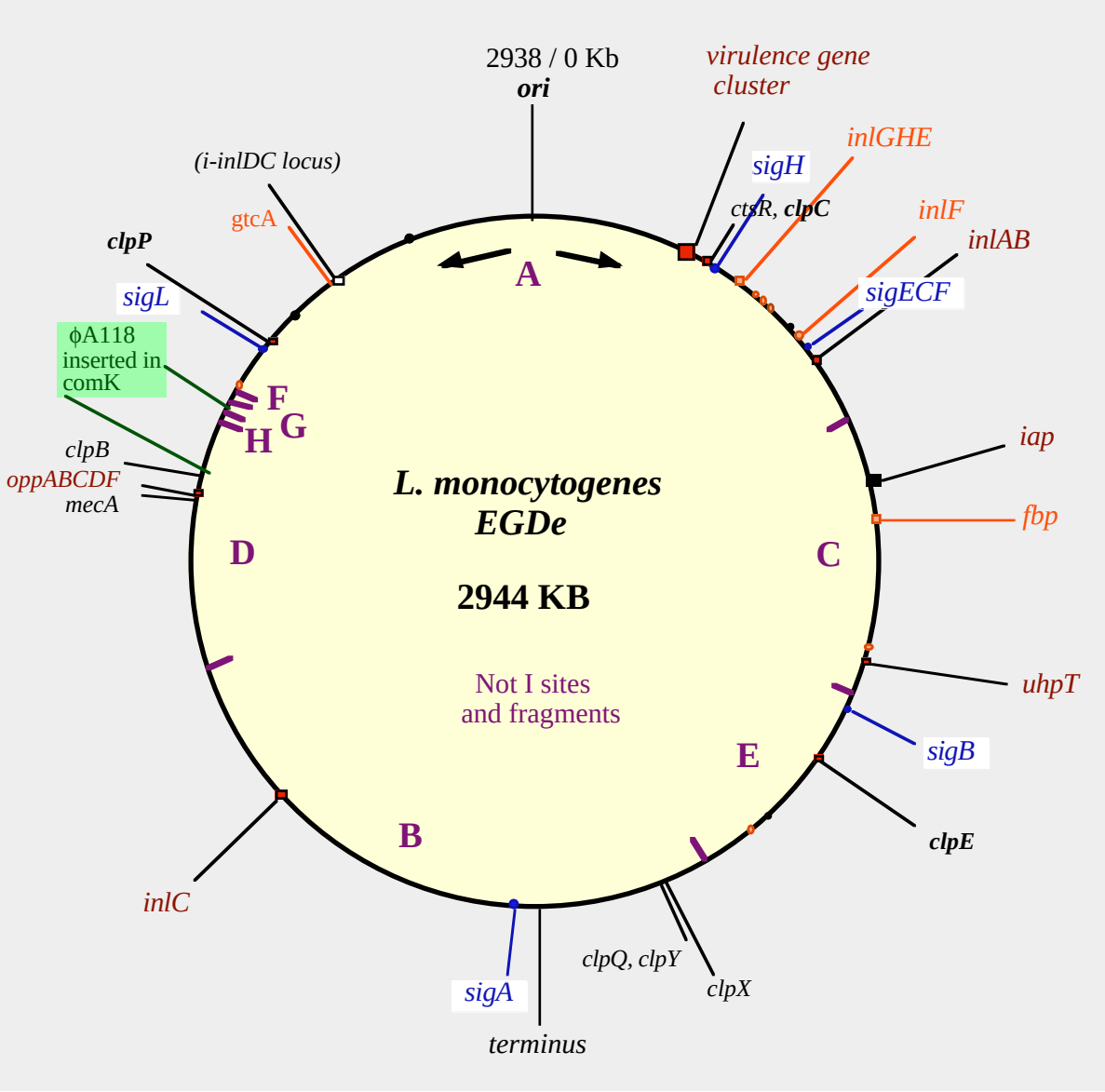


Figure 4.4. The circular map of the *L. monocytogenes* genome.
Red denotes proven virulence associated *loci*: the virulence gene cluster, *iap*, *inlAB*, *inlC*, *oppA* (oligopeptide transporter), and *uhpT* (glucose-6-phosphate transporter, Marcus Beck-personal communications). Black-bold denotes chaperone proteins and proteases: *clpC*, *clpE*, *clpP* (classIII heat shock genes) with demonstrated pleiotropic effects on stress response and virulence functions. Black-not-bold denotes other class III heat shock genes and regulators whose functions are not fully understood in *L. monocytogenes*. Orange denotes loosely defined, "potential" virulence *loci*, these include: *inlGHE*, *inlF* (LPXTG genes), *gtcA* (cell-wall teichoic acid glycolyating protein), and *fbp* (fibronectin binding protein). Black dots denote *iap* and other *iap-like* genes. Blue denotes sigma factors. Green denotes phage A118.

DISCUSSION

To highlight the findings of *L. monocytogenes* EGDe, this genome is compared to its closest, non-pathogenic relative, *L. innocua*, and reference organism *B. subltilis*. See Table 4.3. Because *L. innocua*'s sequence is not entirely completed, the numbers presented are representative but not precise. These comparative figures are obtained from the Consortium and the *B. subtilis* genome (Kunst et al 1997).

Table 4.3. Comparison of *L. monocytogenes* with *L. innocua* and *B. subtilis* genomes.

	Bacillus subtilis 168	***L. monocytogenes EGDe***	***L. innocua Sv6a***
Genome Size	4214.8Kb	2944Kb	~2988Kb
%GC	43.5%	37-38%	--
rRNA operons	10	6	6
Coding sequences (CDS)	4107	2932	+/- 3000
Average CDS size	891bp	912bp	--
Coding Density	87%	90%	--
Insertion Sequences (IS)	0	3 x 1 IS, 1 x Tn916	> *L. mono.*
Phages	10	3	6
Plasmids	0	0	1 (81Kb, contains transposons)
Specific Genes	--	400	230
LPXTG genes	--	41, 19 internalin like	12, 9 internalin like
Sugar transport systems (PTS)	16	42	35
Sigma Factors	18	5	Same 5 as *L. mono.*
2-Component systems	~34	13	--
Synteny	Not well preserved, some small blocks of conservation	Reference	Strong synteny with *L. mono.*
Orthologs	~60% sequence conservation among orthologs with *L. mono.*	Reference	>85% sequence conservation with *L. mono.*

The genomes of *L. monocytogenes* and *L. innocua* are highly similar, most genes exhibit greater than 85% nucleic acid similarity. The pronounced differences are due to the presence of species specific genes interspersed throughout their respective genomes (400 *L. monocytogenes* specific genes, 230 *L. innocua* specific genes), the abundance of phages (1 in *L. monocytogenes*, 6 in *L. innocua*) and the presence of an 81 Kb plasmid in *L. innocua*. The extent that these differences are specific only for these strains as opposed to being consistent differences between the species is not known.

Figure 4.4 illustrates the distribution of various loci of interest on the genome map of *L. monocytogenes*. The known genes associated with virulence capability are scattered on the chromosome. The *L. innocua* Sv6a sequence lack the virulence gene cluster, the *inlAB, inlC, inlGHE, inlF,* and *uhpT* loci. Like *L. monocytogenes*, *L. innocua* contains the orthologous sigma factors (*sigA, sigB, sigH, sigL, sigECF*), *gtcA* (teichoic acid glycosylase), and the many loci simultaneously important for virulence and other cellular functions: the class III stress response *clp* genes, the *oppABDCDF* polypeptide uptake system, and *iap* for cell division. The *L. innocua* Sv6a sequence confirmed all our findings previously reported for *L. innocua* strain Sv6b. Our limited sequence fragments of Sv6b (Chapter 2) are nearly identical to that of Sv6a (data not shown).

A brief comparison between the *Listeria* and *B. subtilis* genomes show poor conservation of synteny (two adjacent genes of one genome remaining adjacent in the other), low level of nucleic acid conservation among orthologous genes, exaggerated multiplicity of PTS systems in *Listeria* (42 in *L. monocytogenes*, 35 in *L. innocua*, 16 in *B. subtilis*), and very different decision making-regulatory networks.

Listeria sigma factors

While *B. subtilis* contains 18 sigma factors and 34 two-component systems, *L. monocytogenes* has 5 sigma factors and 13 two-component systems. The dearth of sigma factors cannot be fully attributed to non-sporulation in *Listeria*. The break down of *B. subtilis*' 18 sigma factors is as follows: SigA is the primary sigma factor, SigB governs general stress, SigI is induced by heat shock, SigD controls flagellar synthesis, motility and chemotaxis, SigL induces the levanase producation, 5 sigma factors (SigE,F,G,K,H) govern post-exponential--sporulation functions, 7 are attributed to extracytoplasmic functions (SigM, SigV, SigW, SigX, SigY, SigZ, ylaC), and an unknown, phage associated sigma Xpf ((Haldenwang 1995; Kroos & Yu 2000; Wosten 1998; Zuber et al 2001) and Subtilist). Of these, only SigA, SigB, SigL, SigH and one ECF-type sigma are present in *Listeria*. *Listeria* appears to have very different regulatory strategies from *B. subtilis,* whose complex differentiation programs are characteristically controlled by specific sigma factors.

Environmental signal relays also differ since *Listeria* has relatively few two-component systems (histadine-kinase sensor + DNA-binding response reguator) and no quorum sensing phenomenon has been reported to date.

These observations raise the following questions. Does *Listeria* favor controlling its physiology in large networks of simultaneously cross-talking PTS intermediary signals over well-defined differentiation programs? Does encounter with eukaryotic host cells trigger discrete differentiation programs via specific alternative sigma factors such as SigH or Sig-ECF? Does *Listeria* effectively occupy different ecological niches from *B. subtilis* though superficially, both are found in soil and decaying matter?

Listeria Stationary phase genes

Regulatory differences are again illustrated by the comparison of the post-exponential functions between the two organisms. Post-exponential—transition—stationary phase is perhaps the dominant growth condition encountered by environmental bacteria. Virtually nothing is known about stationary phase adaptations in *L. monocytogenes*. The best-studied stationary phase network is in *B. subtilis*. In conditions not conducive to exponential growth,

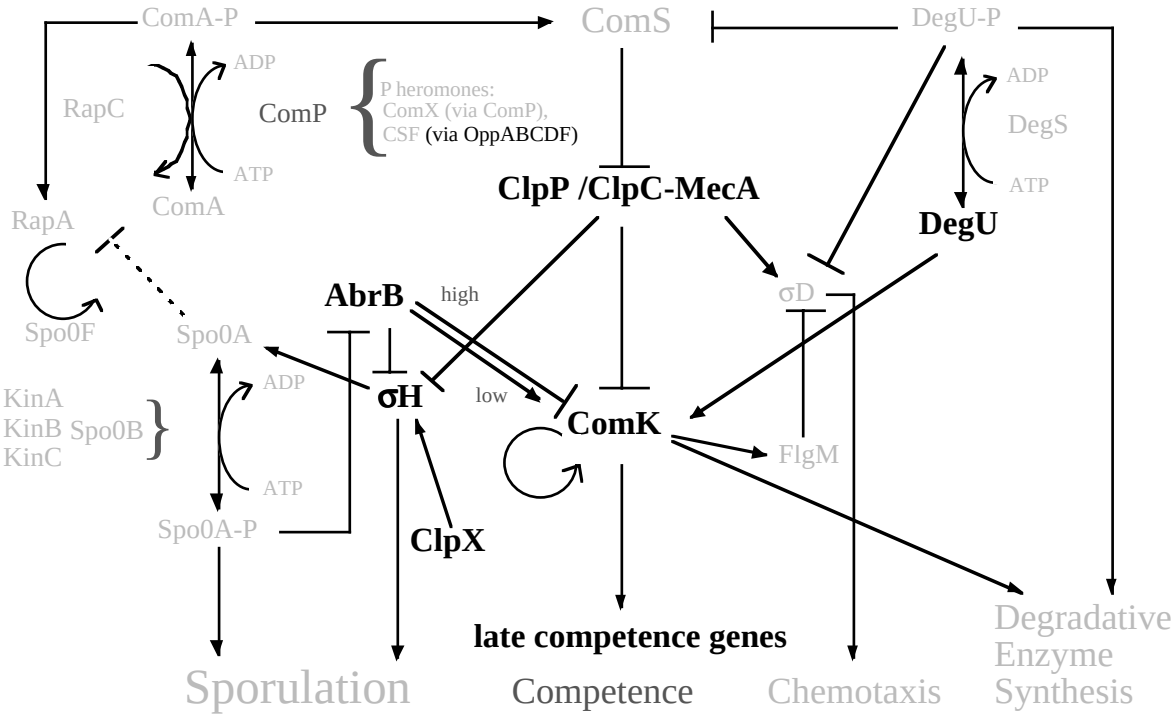


Figure 4.5. Transition—stationary phase regulatory scheme based on *B. subtilis.* Distilled from (Lazazzera 2000; Liu et al 1999; Msadek 1999)). Orthologs that are present in *L. monocytogenes* are shown in black, the absent ones in gray.

the *B. subtilis* complex network regulates the induction of transition state functions such as the secretion of degradative enzymes to scavenge alternative food sources, chemotaxis and motility to search for food or escape noxious agents, and the production of antibiotics to eliminate competitors in the immediate ecosystem. Most importantly, it finely regulates the commitment to two separate developmental pathways: 1) competence for DNA uptake, repair, recombination; and 2) the decision of last resort, sporulation (reviewed in (Lazazzera 2000; Msadek 1999)).

In non-sporeforming *L. monocytogenes*, the entire stationary phase network based on the *B. subtilis* disintegrates; illustrating that sporulation is the heart of all *B. subtilis* transition functions. See figure 4.5. The quorum sensing-pheromone system of *comX* and *phrC* (CSF), and most of the major 2 component regulators of competence and sporulation are absent in *L. monocytogenes*. These components exist to finely regulate the *B. subtilis*' developmental fates of sporulation versus genetic competence and other functions. Although reports vary (Dons et al 1992; Galsworthy et al 1990; Peel et al 1988), chemotaxis / motility is thought to be a temperature dependent phenomenon in *Listeria;* there is presently no evidence that flagellar based motility is a stationary phase function.

However, *Listeria* possesses nearly the entire set of genes encoding the genetic competence machinery (Consortium results: late competence genes, prepared by Berche et al.) This implies that *Listeria* had once been, or still is capable of genetic competence under unknown circumstances. In *L. monocytogenes* strain EGDe as in *L. innocua* strain Sv6a, this function is putatively disrupted by the insertion of phage A118 within the structural gene of *comK*, the key positive regulator of this inferred competence network. A recent report of *L. monocytogenes* containing intact *comK* and deleted *mecA* failed to exhibit competence unlike the positive control *B. subtilis* 168 (Borezee et al 2000a). *B. subtilis* genetics is based on strain 168's competence capability, a trait unnaturally derived from UV and X-ray mutagenesis. Although competence is a highly desirable trait for genetic manipulation in the laboratory, it is apparently a rarely used option in *Listeria* and wild *B. subtilis* strains under natural circumstances.

A comparison of listerial genes against known *E. coli* stationary phase genes yield almost no overlap between the two organisms. In Gram negative *E. coli*, transcription of stationary phase genes are regulated by either stationary phase specific sigma factor RpoS, or by other transcriptional factors independent of RpoS . RpoS is not found in Gram positive bacteria. RpoS driven transcription is enhanced by decreased superhelicity of the chromosomal DNA. In stationary phase, the *E. coli* chromosome drops in superhelicity and becomes more compacted by DNA protecting, histone-like proteins H-NS, IHF, and stationary phase specific Dps/PexB. *dps*, encoding a metallo-DNA binding protein is present in *Listeria* and *B. subtilis (ytkB)*. *B. subtilis* Dps (YtkB) is stress and stationary phase induced, and is under

SigB transcriptional control. Translation is limited in stationary phase *E. coli* by the 55 amino acid ribosome modulation factor (RMF) which converts *E. coli* 70S ribosomes to inactive 100S dimers, rendering them more nuclease and protease insensitive. *rmf* is not found in *Listeria* or *B. subtilis*. The accumulation of inorganic polyphosphate (polyP) is mediated by polyphosphate kinase (PPK) and appears crucial for stationary phase survival of *E. coli*. *ppk* is not found in *Listeria* or *B. subtilis*.

The genome of *L. monocytogenes* offers interesting insight in the regulation of its various functions. The pronounced abundance of phospho-transfer systems and the scarcity of sigma factors both show that Listeria use very different regulatory strategies even though the apparent ecology of saprophytic *Listeria* overlaps that of soil bacteria *B. subtilis*. The discovery of SigH and the ECF-type sigma factor provide intriguing possibilities in their roles in regulating virulence gene expression, since present evidence point to an unidentified sigma factor as essential for this function (see Chapters 1 and 3). Moreover, sporulation alone appear to be the key function in the design of the complex *B. subtilis* post-exponential phase network. With the exception of AbrB, SigH, and the highly conserved Clp chaperons and proteases, non-sporulating *Listeria* do not possess many genes that largely regulate *B. subtilis* stationary phase functions, including the components of two important quorum sensing systems. Neither *L. monocytogenes* nor *B. subtilis* stationary phase genes bear resemblance to Gram negative *E. coli*'s stationary phase functions.

SUMMARY

In conclusion, evidence presented in Chapter 2 indicate that pathogenic capability resided in most if not all *Listeriae* once, and that the presently non-pathogenic species of *L. innocua* and *L. welshimeri* have most likely lost their virulence gene cassettes in two separate events. The genome sequences of *L. monocytogenes* EGDe and *L. innocua* Sv6a support this hypothesis. Known *L. monocytogenes* virulence genes scattered throughout the genome appear deleted or never acquired in *L. innocua*. If our proposed phylogeny remains accurate with accumulating genome-wide data across the genus, then this is an unusual finding. This is unlike other bacterial pathogens, especially those associated with metazoan guts. Of these, the pathogenic enteric "species" and strains such as *Salmonella, Shigella* and pathogenic *E. coli* strains, were commensal, ancestral *E. coli* that have progressively, repeatedly, and often independently acquired pathogenic genes via horizontal gene transfer. Why did *Listeriae* favor losing pathogenic functions while the enteric group showed repeatedly, selective advantage in gaining pathogenic functions? Is this because *Listeriae* are better adapted for a

free-living lifestyle, or that overall selective pressure favors one lifestyle more than another, or that *Listeriae* have not overcome the problem of efficient transmissibility, limiting their capacity to develop into efficient parasites?

This evolutionary scenario is important for functional assessments using *L. innocua* as a testing ground for *L. monocytogenes* virulence genes. While these species are reassuringly closely related, it cannot be assumed that *L. innocua* is truly 'innocent' of virulence genes and adaptations. However, this pair presents a wonderful opportunity to study adaptations and selective trade-offs of the sporadically pathogenic versus entirely free-living lifestyles. The construction of the COR mutants (*L. innocua* with re-instated virulence cluster genes) presented in Chapter 3 will hopefully facilitate the verification of virulence functions, and dissect the regulatory adaptations for pathogenicity. The genome data has yielded a host of internalin-like genes and two important regulators for testing: sigma H and sigma ECF, and other regulatory chaperones and proteases that might regulate them.

Chapter 5

Methods

Chapter 5-A. Materials and Methods for Chapter 2: The Evolution of the genus *Listeria*

Bacterial Strains. All *Listeria* strains used here were from the strain collection maintained at the University of Würzburg, some of these are also obtainable from the Special *Listeria* Culture Collection (SLCC) at the Institute Pasteur or the American Type Culture Collection (ATCC). The strains are: *L. monocytogenes* strains EGD and LO28, of serotype 1/2a and 1/2c respectively; *L. innocua* strain serotype Sv6b, *L. welshimeri* SLCC 5334, *L. ivanovii* ATCC 19119 (SLCC 2379), *L. seeligeri* SLCC 3954, L. *grayi*. Species identity of all strains has been confirmed using species specific primers for the *iap* gene (Bubert et al 1992a; Bubert et al 1992b). *E. coli* host strain TOP10 used for cloning was provided in the TOPO TA Cloning Kit (Invitrogen).

Media. Brain Heart Infusion (BHI from Difco) was used for the growth of all listerial strains. *E. coli* strains used for cloning were grown on Luria-Bertani (LB) medium. Antibiotic selection used in the cloning procedure was Ampicillin 100 µg/ml in LB (stock solution 100mg/ml in dH_20).

DNA extraction methods. Chromosomal DNA was obtained from each *Listeria* in the following manner. A single colony isolate was inoculated into 10ml BHI and grown overnight with rolling at 37°C. The cells were harvested 16 hours after inoculation and pelleted. The pellet was washed with 5ml of 0.1x SSC (1x SSC: 0.15M NaCl, 0.15M trisodium citrate, pH7.0) and re-pelleted. The cells were treated at 37°C for 1 to 2 hours with 0.5ml of 25mg/ml lysozyme (Sigma) dissolved in TES (30mM Tris/HCl pH8.0, 50mM NaCl, 5mM EDTA pH 8.0, 20% sucrose). Lysis was effected with the addition of 4.5ml Lysis Buffer (10mM Tris/HCl pH 8.0, 1mM EDTA pH 8.0, 1% SDS, 0.5mg/ml Proteinase K(Merck)) and incubated at 37°C for 1 hour. DNA was extracted by gentle inversion using one or two extractions with phenol, followed twice with phenol/$CHCl_3$ (1:1) and once with $CHCl_3$. DNA was precipitated by the addition of 500µl of 3M Na(OAc) and 10ml ethanol, and collected by spooling with a glass rod. It was briefly washed in 70% ethanol and finally resuspended in 400µl of sterile distilled water. This spooling method of harvest as opposed to centrifugation gave larger molecular weight molecules and greatly enhanced the success of long range PCR. Plasmid DNA was isolated from *E. coli* cells using Nucleobond AX100 midiprep columns (Macherey-Nagel) according to the instructions of the manufacturers.

PCR Amplifications. The primers used successfully for the amplification of DNA fragments described in this study are listed in Table 5-A.1. Many more primer pairs in different permutations were tried. The range of trail annealing temperatures for various reactions was 45 °C, 50 °C, 52 °C and 54 °C. The PCR machines used were Techne and Perkin Elmer GeneAmp 2400 for all the above protocols.

Table 5-A.1. PCR primers used for amplification of DNA fragments from listerial chromosomal DNA in this study.

Region of Interest	Primer Name	Primer Sequence 5' to 3'
Virulence Gene Cluster	prs1>	GCGCCGATTGCTATTATTGA
	ldh1<	GAATTCCCAGCATGGAGCCA
	ivan-plcb1>	AAATGCGAAACAGACCTGCG
	see-plcb1>	ACAAGGGCTTTCAGATTCTC
	see-vclY3<	TCATATGTAAAGCTGGATGATC
	see-vclY1>	GGTCTATTTAGTTAGAGGAGA
Extending 5' of *prs*	con-prs1>	GTGGTTGTCATGTATATGTTATTCAA
	see-prs1<	GTGGTGCTACAGACAGCTGT
	see-prs2<	GAGCAATGGAGTTAGTAACAACT
	ivan-prs1<	ACAGATGCATTTTCACGTACA
	ivan-prs2<	ACGATTGCTTCACCTAGCAGT
Loci corresponding to *inlC*	rpls2>	TCGAAGGCGCTGCAGTCAAACG
	infC1<	GTCTTCGCACGCTTTTGCA
Loci corresponding to *inlGHE*	pGluco2>	GTAAGTGCCTGCAGAAGCGAAATGTCC
	PGluco1>	AGTAAGTGCCTCCACAAGCG
	desuc1<	TGTAAACATCTACCATCTCCAA
Loci corresponding to *i-inlDC*	li-inlD4>	GAGAGAGCAATCTTTCAAC
	li-emr5<	TTTCACCAACTAAAGCATTCAT
	li-emr6>	GAGGTGTTTTTTTGAAGGAGAA
	li-emr1<	GTGTATCCATCGTTAAGAACAT

Only successful primers are listed.

For products less than 4 to 5 Kb. Generally, the reactions were done using 2mM $MgCl_2$ for Taq DNA Polymerase (Promega or Silver Star) and 2-4 mM $MgSO_4$ for Deep Vent DNA polymerase (New England Biolabs), the respective buffer provided by the manufacturer, 200µM of each dNTP, 30 to 50µM of each primer, 1ul of chromosomal DNA obtained as described above, and 1U of Deep Vent or Taq DNA polymerase in 100µl reaction volumes. For trial reactions under different annealing temperatures, this 100µl was split into 3 x 33ul reactions for economy and convenience. Standard cycling parameters were 30 cycles with denaturation for 30 seconds at 94°C, annealing temperatures are specified in Table 2.1 for each primer pair used, extension times were calculated as 1 minute per Kb, at 72°C.

For products larger than 4 to 5 Kb. The GeneAmp XL PCR kit (Perkin Elmer) was employed. The conditions used for the PCR of the virulence gene cluster fragments up to 12 Kb: 1x XL Buffer II, 200µM of each dNTP provided in the kit (do not substitute), 15µM of each primer, 2mM $Mg(OAc)_2$, 1-2µl chromosomal DNA template obtained by spooling method described above, and 0.5-1U rTth DNA Polymerase XL in a 100µl final volume. Enzyme was added after a 2-3 minute 94°C treatment to denature the template DNA, i.e. the 'Hot Start' method. The cycling parameters for products up to 12Kb were 25 cycles in total: the first 12 cycles consisted of 30 seconds denaturation at 94°C, 30 seconds annealing at 53°C, 11 minutes extension at 68°C; followed by 13 cycles of the above except with an increase of 15 seconds of extension time for each proceeding cycle. This was followed by a 60-minute final extension at 68°C before chilling to 4°C.

Cleaning the PCR reactions. PCR reactions were cleaned of protein and unincorporated primers using the PCR purification kit (Qiagen). Up to 100µl of reaction volume could be used at one time (manufacturer's instructions). For volumes greater than 100µl, repeated loadings of 100µl + 500ul PB buffer can be added onto the same column, spun down, before the WASH step. This was especially useful for concentrating low-yield reactions. DNA was eluted with 30-50µl of dH_20.

Cloning of the PCR products. The TOPO TA cloning kit (Invitrogen) was predominantly employed here as a rapid way to clone PCR products. Because this technique depends on the presence of a 3' dATP characteristically generated by Taq DNA polymerase, blunt-end templates had to be treated for adaptation to this kit. PCR products generated by Deep Vent DNA polymerase and rTth DNA polymerase for the sake of greater amplification accuracy and/or longer products are blunt ended. These reactions are cleaned using PCR purification kit (Qiagen), and the 3' A extension was effected by a 15 minute incubation at 72° in a 25µl reaction containing cleaned PCR product, 1x Taq DNA polymerase buffer, 200µM dATP, and 1U Taq DNA polymerase. The TA cloning protocol (Topoisomerase based 'ligation' and immediate transformation) as directed by the manufacturer followed immediately. To economize, reactions were routinely scaled to 1/2 or even 1/3 of the recommended volumes. Cells were selected and screened on LB-Ampicillin 100ug/ml-Xgal plates, 16-24 hours incubation at 37°C.

DNA Sequencing. Greater than 40Kb of double stranded sequence was obtained for this study. Sequences reported in this study are high quality and obtained for both strands. Most sequence discrepancies were observed to be caused by cloning, and thus multiple clones and independent PCR reactions were often employed.

Approximately half of this sequencing was done in-house on a Perkin-Elmer-Biosystems ABI 310 capillary machine using BigDye chemistry. This included the

sequencing of either PCR products or cloned PCR products. Reactions contained 200ng of plasmid template DNA, or 50ng PCR product template DNA, 3.2pmoles of primer, 3µl BigDye in a 20µl reaction volume. The reactions were cycle-sequenced as recommended by the manufacturer in a Techne PCR machine. The reactions were cleaned by Sodium Acetate and Ethanol precipitation as recommended by the manufacturer. The results were hand checked and assembled. The average reliable read obtained was 0.35-0.4Kb, this was a function of the capillary length, an inherent limitation of the ABI 310. Therefore, for chromosomal walking type sequencing tasks, this was rather inefficient and costly in time and money. Cost is driven by the turnover time, the number of reactions, and the number of primers needed to cover a given stretch of DNA. Because of capacity saturation in-house, the remaining half of the sequences were done by MWG on the Licor system with fluorescent dye-primers (IRD-800) producing 0.7-1Kb long reads.

Computer analyses. Sequence information was managed using GCG version 10.0-UNIX (the Wisconsin Package), DNA strider version 1.2 software (Mac), and Clone Manager (PC). Internet sites used for the various queries are listed in Chapter 5-D.

Annotation was done by hand. The sequences were translated into all 6 frames for examination by eye, while electronic searches using varying sized fragments were subjected to progressive blast searches. Orfs were confirmed, or discovered by eye using probable start sites (ATG, TTG,GTG), ribosome binding sites, transcriptional termination sites, and similarities to known sequences. Consensus motif searches and functional assignment confirmation were done using the ProDom, Japan, and Blocks databases, and by eye. Alignments of motifs were done against ProDom database using 'MultAlin' option. Predicted Orfs were then checked for probable cellular localization using the signal peptide prediction software. When multiple listerial homologs were discovered, these were aligned to examine the extent of divergence using 'pileup' and displayed with 'pretty' or 'prettybox' in GCG. Searches for repeats were done using 'gap', 'bestfit', 'hairpin' in GCG; progressive segments were scanned and queried against itself, and against different sections of the sequence in both orientations.

Phylogenetic analyses. Michael Schmid and Michael Wagner performed these analyses at the *Lehrstuhl für Mikrobiologie* of *Technische Universität München* . For phylogenetic analyses, the ARB software package (Ludwig et al 1998) was used. Phylogenetic trees based on nucleic acids were calculated using "Maximum Parsimony" (PAR), "Maximum Likelihood" (ML) (Felsenstein 1981) and "Neighbor-Joining"(NJ) (Saitou & Nei 1987) methods. Amino acid sequence based trees were calculated using NJ, ML, and the Protein Parsimony methods (Eck 1966). In addition, amino acid trees were inferred from

distances using FITCH with global rearrangements in the "Phylogeny Inference Package" (PHYLIP) version 3.57c, from the Departments of Genetics, U. of Washington, Seattle.

Data for the phylogenetic study: in addition to DNA sequences reported in this study for *vclA, vclB, ldh*(partial) and *prs* (partial), all available sets of molecules for *Listeriae* were employed. Sequence data for 16S rRNA molecules were from X56153 (*L. monocytogenes*), X56149 (*L. welshimeri*), X56151 (*L. ivanovii*), X56148 (*L. seeligeri*), X56150 (L. grayi) (Collins et al 1991), X55473 (*L. innocua*) (Czajka et al 1993). 23S rRNA sequences were derived from X92951 (*L. monocytogenes*), X92949 (*L. innocua*), X92954 (*L. welshimeri*), X92950 (*L. ivanovii*), X92953 (*L. seeligeri*), and X92948 (L. grayi) (Sallen et al 1996). Iap or invasion associated protein (P60) sequences were derived from X52268 (*L. monocytogenes* EGD), M80351 (*L. monocytogenes* Mackeness), M80347 (*L. innocua* Sv6a), M80349 (*L. innocua* Sv6b), M80348 (*L. welshimeri*), M80350 (*L. ivanovii*), M80353 (*L. seeligeri*) and M95579 (L. grayi) (Bubert et al 1992b). For phylogenetic analyses of the *ldh* genes: *Bacillus caldolyticus* (acc.no.: M19394), *Bacillus caldotenax* (M19386), *Bacillus stearothermophilus* (AB033627); *Bifidobacterium longum* (M33585), *Deinococcus radiodurans* (AB005539), *Lactobacillus casei* (M76708), *Lactobacillus sakei* (U26688), *Lactococcus lactis* (M88490), *Mycoplasma genitalium* (U39733), *Mycoplasma hyopneumonia* (X67286), *Streptococcus mutans* (M72545), *Streptococcus pneumonia* (AJ005815), *Thermotoga maritima* (X74302), *Thermus aquaticus* (D00858) were used as outgroups. For phylogenetic analyses of *vclB* genes: *E. coli* (AE000188) and *B. anthracis* (AF188935) were used as outgroups.

Chapter 5-B. Materials and Methods of Chapter 3: Towards reconstituting a pathogenic *Listeria*.

Table 5-B.1. Bacterial Strains and plasmids used in this study

Strains	*Genotype*	*Source*
L. monocytogenes EGD	'wildtype'	S.H.E. Kaufmann
L. monocytogenes PKP1 (EGD *vcl* Δ mutant)	*plcA*$^-$, *hly*$^-$, *mpl*$^-$, *actA*$^-$, *plcB*$^-$	(Engelbrecht et al 1996)
L. monocytogenes PKP1+1	*mpl*$^-$, *actA*$^-$, *plcB*$^-$	constructed here
L. innocua (serotype Sv6b)	*prfA*$^-$, *plcA*$^-$, *hly*$^-$, *mpl*$^-$, *actA*$^-$, *plcB*$^-$	SLCC collection (Andreas Bubert)
L. innocua (serotype Sv6a)	*prfA*$^-$, *plcA*$^-$, *hly*$^-$, *mpl*$^-$, *actA*$^-$, *plcB*$^-$	NCTC 11288
L. innocua COR1	Sv6b; co*mpl*emented with *prfA plcA hly*	constructed here
E. coli DH10B	F' *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ* ΔM15 Δ*lacX*74 *doeR* recA1 *endA*1 *araD*139 Δ(*ara-leu*)7697 *galU galK* λ$^-$ *rpsL* (Str^R) *nupG*	Gibco-BRL
Plasmids	*Genotype*	*Source*
E. coli TOPO10F'	F'{lacI^q Tn10 (Tet^R)} *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lacZ* ΔM15 Δ*lacX*74 *recA*1 *araD*139 Δ(*ara-leu*)7697 *galU galK rpsL* (Str^R) *endA*1 *nupG*	Invitrogen
pWH1509E	ori pE194 ts, Erm^R (Gram +); ori pBR322 Tet^R Amp^R (Gram -)	(Rygus & Hillen 1992)
pLSV1	ori pE194 ts, Erm^R (Gram +); ori ColE1	(Goebel et al 1991)
pWKS30	Amp^R *lacZ*, low copy number	Strategene
pCR2.1-TOPO	Amp^R *lacZ* Kn^R; ori ColE1	
pEN01	pWH1509E; Tet^R Erm^R ts(Gram +) *L. innocua orfZ* in ScaI site, orientation 1	constructed here
pEN02	pWH1509E; Tet^R Erm^R ts(Gram +) *L. innocua orfZ* in ScaI site(orientation 2)	constructed here
pEN03	pWH1509E; insert *L.mono prs*(3') *prfA plcA hly mpl*(5'), *L. inn. orfZ*	constructed here
pEN04	pWH1509E; insert *L. mono. hly*(3') *mpl actA*(5'), *L. inn. orfZ*	constructed here
pEN05	pWH1509E; insert *L. mono. actA*(5') *plcB, orfXYZ, orfB* (5'), *L. inn. orfZ*	constructed here
pEN06	pWH1509E; insert *L. mono. hly*(3') *mpl, actA*(5'), *orfX, L. inn. orfZ*	constructed here

Table 5-B.1. Continue: Plasmids used in this study.

Plasmids	*Genotype*	*Source*
pEN07	pWH1509E; insert *L. mono prs*(3') *prfA plcA hly mpl*(5'), *orfX, L. innocua orfZ*	constructed here
pEN08	pWH1509E, insert *L. inn. prs*(3'); *L.mono prs*(3'), *prfA, plcA, hly, mpl*(5'); *orfX*; *L. inn. orfZ*	constructed here
pTOPO-I	pCR2.1-TOPO; insert *L.mono prs*(3') *prfA plcA hly mpl*(5')	constructed here
pTOPO-II	pCR2.1-TOPO; insert *L. mono. hly*(3') *mpl actA*(5')	constructed here
pTOPO-III	pCR2.1-TOPO; insert *L. mono. actA*(5') *plcB, orfXYZ, orfB* (5')	constructed here

Media. Brain Heart Infusion (BHI from Difco) was used for the growth of all listerial strains. Antibiotic used for transformation procedures were Erythromycin 5 µg/ml in BHI (stock solution 10mg/ml in 95% ethanol), and Kanamycin 50 µg/ml in BHI (stock 50 mg/ml dH_20). *E. coli* strains used for cloning were grown on Luria-Bertani (LB) medium. Antibiotic selection used in various cloning procedures included Ampicillin 100 µg/ml in LB (stock solution 100mg/ml in dH_20), Tetracycline 7.5 µg/ml in LB (stock solution 10 mg/ml in 50% ethanol), and Kanamycin 25 µg/ml in LB (stock 50 mg/ml in dH_20). β-galactosidase driven, blue/white screens for insert presence during certain clonings were carried out using Xgal in the selective media.(2ml of 2% X-gal in DMF added to 1L of LB agar when cooled to 55°C).

DNA isolation methods standard for both *Listeria* and *E. coli* are as described in the Chapter 5-A above, except for small scale, DNA-Zol (BRL), chromosomal DNA isolations used for the screenings of chromosomal integrands. Typically, a small swap of cells was suspended in 250ul TE pH8.2 containing freshly added 5µl of 120mg/ml lysozyme. This suspension was incubated at 37°C for 20 to 30 minutes, spun briefly in a microfuge and the supernatant was removed. 250µl of well suspended DNA Zol solution was added to lyse the cells, which yielded a cleared solution. 125µl (1/2 volume) of 100% ethanol was gently mixed in by inversion. The supernatant was removed after a quick spin, the pellet was washed twice with 200µl 95% ethanol, and dried briefly in air before being resuspended in 30µl 8mM NaOH solution. 1µl of this lysate was used in a 50µl volume PCR reaction.

Standard molecular biology techniques not mentioned here. Please refer to Maniatis et al. (Maniatis 1982).

Table 5-B.2. Primers used in this study.

5' Primer Name	*Sequence*
5'AmpScaI>	AAGAGTATGAGTATTCAACAT
*ActA*2>	CCCAACAGAAGAAGAGTTGAAC
con-*prs*2>	CACCGCAAATCCAAGGTTTCT
DC1-inn>	CTTGTATTAGAATGTTTCGCTAG
DC2-egd>	GATAATTTGTATGTTTCGCTAG
egd-*prs*4>	TAATCGAAACTGCTGGTGCA
*hly*2>	ACCACGCTTTATCCGAAATAT
*hly*3(Aat2)>	GGAAAAATTAACATCGACGTCTCTGGAGGATACGT
orfB2(Eco)>	AGCCGAATTCAAATCTTGCGCTTCGATGACAA
*orfZ*8-inn(XhoI)>	TCATCACTCGAGTTAGGTACTCTTAAATAGTACA
*plcA*1>	TTCGCAAAAGAAATGCATACA
*plcB*2(XhoI)>	AACCTACTCGAGCTAG*ACTA*AGAGATTC
*plcB*2>	CTGGAGCTAG*ACTA*AGAGATT
*prs*1(XhoI)>	TTATCACTCGAGGCCGATTGCTATTATTGAT
*prs*2(Sal)>	TCTAGGGTCGACGCGCCGATTGCTATTATTGA
*prs*4-inn>	TTCGCCTTCTAAGTG*ACTA*TT
*prs*5(Aat2)>	GTAAAATGGCTGGACGTCTGAAAGCGCCGATT
*prs*8-inn(Aat2)>	GATTTAGTTGACGTCTCACCTGACCAC
3' Primer Name	*Sequence*
<3'AmpPst	GATCTTCACCTAGATCCTT
<*ActA*1(XhoI)	TGGTTTCATTTCCTCGAGATCGCTCTCTGTAGC
<emr1-egd	TTCAGCCGAACATTCTGTAA
<emr2-inn	AGCCGAACATTCTGCAAGAA
<*ldh*5(Eco)	TCTAGGGAATTCCCATTACGGTATGAACGATAT
<*mpl*2	CTAATCTGACAGAGAGAGTTA
<*mpl*2(XhoI)	CTAATCTGACAGCTCGAGTTAAGGACACGT
<orfB10	CAAGGCTATGAAGATGAACTTAT
<orfB3	GTGTTGATATTGATTATGAAGTGC
<*orfX*1(XhoI)	TCATCACTCGAGAACCGC*ACTA*TTGCAATAGAT
<*orfX*2(XhoI)	AGATATCTCGAGAACATTTCCCAACAAATAC
<*orfZ*8-inn(XhoI)	TCATCACTCGAGTTAGGTACTCTTAAATAGTACA
<*orfZ*9-inn	CTTCTTTAAAGCCTTGAGATT
<*plcA*1	TTTCAGGTGTATTAGAAACGA
<*plcB*1(XhoI)	TCATCGCTCGAGTATACATTTGGCTTACTTCCT
<*prfA*2	ACAGAAACATCGGTTGGCTAT
<*prfA*3	TTTAGCATGTCCTGCTACTTG
<*prs*9-inn(Aat2)	TTATCCCGACGTCAAGTTATCACACCCTCAT

Restriction sites are described in the primer name, and underlined in the sequences.

Table 5-B.3. A summary of the primers and conditions used in the cloning and integration of the virulence gene cluster (*vcl*) of EGD into *L. innocua* Sv6b and *L. monocytogenes* Δ PKP1.

Experiment	*Purpose*	*Fragment Size*	*Restric-tion used*	*Primer pair*	*PCR anneal temp °C*	*Poly-merase*
vcl as 1 piece	Obtain vcl fragment from EGD	~12Kb	none	*prs*2(Sal)> <*ldh*5(Eco)	50-54	rTth
	Test: insert presence	330bp	–	*hly*2> <*mpl*2	50	Taq
vcl as 2 pieces	Obtain *L. innocua orfZ* fragment	500bp	none	*orfZ*8-inn(XhoI)> <*orfZ*9-inn	52	DVent
(pEN01)	Test: insert, orientation (1) in pWH1509E	800bp	–	5'AmpScaI> <*orfZ*8-inn(XhoI)	52	Taq
(pEN02)	Test: insert, orientation (2) in pWH1509E	800bp	–	5'AmpScaI> <*orfZ*9-inn	52	Taq
	Obtain PKP1 vcl fragment	5Kb	XhoI	*prs*1(XhoI)> <*orfX*1(XhoI)	50	rTth
	Test: insert, orientation in pEN01	1070bp	–	*plcB*2> <*orfZ*9-inn	52	Taq
	Obtain EGD *plcA-plcB* fragment	7.9Kb	none	*plcA*1> <*plcB*1(XhoI)	50	rTth
	Test: insert in pWH1509E	330bp	–	*hly*2> <*mpl*2	50	Taq
vcl as 3 pieces	Obtain fragment "I" from EGD (3'*prs*,*prfA*,*plcA*, *hly*, 5'*mpl*)	4.5Kb	AatII/ XhoI	*prs*5(Aat2)> <*mpl*2(XhoI)	54	rTth
(pEN03)	Test: insert, orientation in pEN01	830bp	–	*hly*2> <*orfZ*9-inn	50	Taq
	Test: 5'integration into *L. innocua* Sv6b	2103bp	–	*prs*4inn> <*plcA*1	48	Taq
	Test: 3'integration into *L. innocua* Sv6b	965bp/ 1085bp	–	*hly*2> <orfB10	48	Taq
	Test: 5'integration into PKP1	2204bp	–	egd-*prs*4> <*plcA*1	48	Taq
	Test: 3'integration into PKP1	1280bp	–	*hly*2> <orfB10	48	Taq
	Obtain fragment "II" from EGD (*mpl*, *actA*-inc*ompl*ete)	4.1Kb	AatII/ XhoI	*hly*3(Aat2)> <*ActA*1(XhoI)	54	rTth
(pEN04)	Test: insert, orientation in pEN01	1043bp	–	*ActA*2> <*orfZ*9-inn	50	Taq
	Obtain fragment "III" from EGD (3'*actA*, *plcB*, *orfX*,Y,Z,5'B)	2.8Kb	none	*ActA*2> <orfB3	54	rTth

Table 5-B.3. Continue...

Experiment	*Purpose*	*Fragment Size*	*Restric -tion used*	*Primer pair*	*PCR anneal temp °C*	*Poly-merase*
(pEN05)	Test: insert, orientation (1) in pWH1509E	476bp	–	orfB2(Eco)> <5'AmpSca	50	Taq
	Test: insert, orientation (2) in pWH1509E	725bp	–	orfB2(Eco)> <3'AmpPst	50	Taq
Modifying "pEN03,4"	Obtain *orfX* fragment from EGD	548bp	XhoI	*plcB*2(XhoI)> <*orfX*2(XhoI)	48	Taq
(pEN07)	Test: insert, orientation in pEN03	863bp	–	*hly*2> <*orfX*2(XhoI)	50	Taq
	Test: insert, number in pEN03	1370bp	–	*hly*2> <*orfZ*9-inn	50	Taq
(pEN06)	Test: insert, orientation in pEN04	1120bp	–	*ActA*2> <*orfX*2	50	Taq
	Test: insert, number in pEN04	1573bp	–	*ActA*2> <*orfZ*9-inn	50	Taq
	Obtain *prs* fragment from *L. innocua*	643bp	AatII	*prs*8-inn(Aat2)> <*prs*9-inn(Aat2)	48	DVent
(pEN08)	Test: insert, orientation in pEN07	1037bp +407bp	–	*prs*5(Aat2)> <*prfA*3	52	Taq
pEN08 in Sv6b	Test: 5'integration into *L. innocua* Sv6b	1.6Kb	–	*prs*4-inn> <*prfA*2	50	Taq
(COR1)	Test: 3'integration into *L. innocua* Sv6b	1.6Kb	–	*hly*2> <orfB10	50	Taq
pEN08 in PKP1	Test: 5'integration into PKP1	4.7Kb	–	con-*prs*2> <*mpl*2(XhoI)	50	rTth
(PKP1+1)	Test: 3'integration into PKP1	1.6Kb	–	*hly*2> <orfB10	50	Taq
Confirm Species	*L. innocua*	2.1Kb	–	DC1-inn> <emr2-inn	50	Taq
	L. monocytogenes	500bp	–	DC2-egd> <emr1-egd	50	Taq

[ ] Strains and plasmids constructed from these experiments are indicated in the left-hand column in parentheses.

PCR primers used in the amplification of DNA fragments in this study are listed in Table 5-B.1. Table 5-B.2 lists their respective uses and the conditions in which they were used in each of the experiments described. The amplification procedures are as described in Chapter 5-A.

DNA sequencing employed BigDye chemistry on an ABI 310 sequencing machine (in house) was described in Chapter 5-A.

Cloning. Table 5-B.3 above describes the clonings of the *L. monocytogenes* virulence gene cluster (vcl) in this study. Direct cloning from PCR products using the TOPO-TA Kit (Invitrogen) was previously described in Chapter 5-A. Additional procedures employed in this study are described below.

Kinasing with T4 polynucleotide kinase (T4 PNK). PCR products made by Deep Vent or rTth/Deep Vent mixed polymerases have dephosphorylated ends. In order to use these products with dephosphorylated vectors, the insert must be kinased for ligation to occur. To enhance activity for blunt end substrates, DNA (1µg) was heat treated at 70°C for 5 minutes and placed on ice. The reaction mix containing 1X T4 PNK buffer, 1mM dNTP, 10-50U T4 PNK (NEB) enzyme in a reaction volume of 50µl was incubated at 37°C for 30 to 40 minutes, and cleaned up with Qiagen Clean Up column, and eluted in dH_2O.

Ligation with dephosphorylated vectors prevents the vectors from self-ligating, and reduces the number of false positives in the transformation. Dephosphorylation with shrimp alkaline phosphatase (SAP) was carried out in the smallest volume possible containing the amount of vector for one ligation reaction (50-200ng DNA). DNA was suspended in 1X SAP buffer with 1U SAP enzyme (Boehringer Mannheim) and incubated at 37°C for 1 hour. The enzyme activity was inactivated by heat at 65 °C for 15 minutes. Ligation followed directly by augmenting buffer concentration to 2X, adding insert DNA and 1U ligase (Gibco-BRL). Alternatively, ligase buffer was added at 1X concentration to the SAP reaction, insert DNA, ≤1U ligase and dH_2O to a volume of 10-20ul. Reactions proceeded as described below.

Ligations with normal, phosphorylated vectors were set up with insert DNA and linearized vector DNA, 1X ligase buffer, ≤1U ligase (Gibco-BRL) and dH_2O to a volume of 10-20µl per reaction (10µl for 100ng vector as guideline). Less than 1U of ligase was used for sticky ends reactions. Ligation reactions were carried out at 16 °C overnight, or at room temperature for 1 to 6 hours.

Desalting of DNAs is crucial for electroporation to prevent arcing. PEG precipitation was used to desalt ligation reactions. dH_2O was added to a ligation mix to achieve a 30µl volume. 15µl of 30% PEG8000 in 1.5M NaCl was mixed in and the DNA was precipitated on ice for 30 minutes. The sample was spun at 4 °C, 30minutes at top speed in a microfuge. The supernatant was removed and the pellet was washed twice with 0.5ml of 70% ethanol that

was stored at -20°C. The DNA was air dried and resuspended in ≤10µl dH_2O. One to 2µl of this was used in each electroporation (equivalent to about 10ng vector DNA).

Electrocompetent *E. coli* cells were prepared from 1L of cells growing in LB, at OD600 = 0.5, and not exceeding 0.7. Cells were chilled on ice 15-30 minutes, spun down at 4000g at 4°C (5000rpm in a Beckman JA-10 rotor). Maximum amount of the supernatant was carefully discarded, and the cells were gently resuspended in ice cold, sterile dH_2O or 10% filter sterilized, glycerol. This procedure was repeated 4 times, and cells resuspended in progressively smaller volumes (1 volume, followed by 0.5 volume of sterile dH_2O, followed by 0.02 volume, and finally in 1ml of 10% filter sterilized glycerol). These cells were divided into 40µl aliquots on ice, flash frozen in an ethanol-dry ice bath, and stored frozen at -70°C. Cells routinely exhibited an efficiency of transformation at 10^6 to 10^7 cfu/µg of pUC or other vector DNAs.

Electroporation was carried out with fresh or frozen competent cells prepared as above. Desalted DNA was gently added to the 40µl of cells on ice. The mix was gently transferred to pre-chilled electroporation cuvettes (0.2cm gap length), and electroporated at 25µF, 200Ω, 2.5KV (BioRad Apparatus). Ice cold SOC (LB supplemented with 10mM $MgSO_4$, 10mM $MgCl_2$, and 20mM glucose) was immediately added to the electroporated cells and gently transferred into sterile, 10ml tubes. The cells were incubated with gentle rolling at 37°C for 1hour for Amp selection, or up to 2 hours for Tet selection. The cells were plated at the dilution desired on selective plates and incubated 16 to 24 hours at 37°C.

"Cracking" was used as a fast method to assess the size of plasmids in *E. coli* transformants. A swap of cells was suspended in 40-50µl of sterile dH_2O. An equal volume of 2X Cracking buffer (0.1M NaOH, 0.05M EDTA pH8.0, 1% SDS, 0.05% Bromocresol Green, 10% glycerol) was added and the mix was vigorously vortexed for 30-60 seconds, left at room temperature for 10 minutes, and 15-20ul of this was directly loaded onto ≤0.8% agarose gel to run against supercoiled DNA standard (Gibco-BRL).

Transformation of *Listeria*. Electroporation was used to transform *Listeria* in this study. *Listeria* cells (1-2ml overnight culture) were inoculated into 50ml BHI containing 0.02% Glycine in side-arm flasks and grown at 37°C with vigorous shaking till early to midlog (80-90 Klett units). Penicillin G at 5ug/ml was added (2.5ul of 100mg/ml Penicillin G stock), and the cells were grown till 120-130 Klett units. The cells were placed on ice and harvested by spinning at 6000 rpm in a Hereaus centrifuge for 10 minutes at 4°C. The pellet was twice washed with 5ml 3.5X SMHEM (952mM sucrose, 3.5mM MgCl2, 7mM HEPES, final pH = 7.2, and sterilized in an autoclave), and spun down as above. The final pellet was resuspended in 0.5ml 3.5X SMHEM, 100-200µl was used for each electroporation. Unused cells were frozen at -70°C.

Two to 5µg of supercoiled DNA suspended in ≤8 µl sterile dH_2O was mixed with 100 - 200µl of cells on ice, and transferred to pre-chilled 0.2cm gap length, electroporation cuvettes. The mix was electroporated at 25µF, 100Ω, 2.25KV (BioRad Apparatus), and immediately suspended in a final volume of 1ml BHI. The cells were transferred to 10ml sterile tubes and incubated with rolling at 30 ºC for 3 hours (all antibiotic selection except Erm) or 6 hours (Erm selection). No more than 200µl of cells were plated on each selective plate. Colonies appeared at or after 3 days of incubation at 30 ºC (for temperature sensitive plasmids).

Chapter 5-C. Materials and Methods of Chapter 4
The Genome of *L. monocytogenes*

MATERIALS: Perkin Elmer / Applied Biosystems

BIGDYE 4303149 (100reactions, at 8ul/reaction), cost: 1200DM

310 Capillaries 61cmx50um, **402820** 2/pkg
(long capillaries, read>600bp), cost: 193DM

Analyser Tubes 401957 (500/pkg) cost: 75DM

Analyser Septa 401956 (500/pkg) cost: 217DM

TSR solution (8vials) + 3ml POP6 resin, 403076, cost: 300DM

CentriSep Columns (Princeton Separations), clean up unincorporated ddNTP
401762 (100/pkg) cost: 338DM **401763** (32/pkg) cost: 120DM

PRIMERS were purchased exclusively from ARC GmbH for quality and 24 hour delivery.

SOFTWARE: Phred / Phrap / Consed mounted on Linux platform was used to assess sequence quality and perform local assembly within gaps. Phred / Phrap / Consed was provided by Phil Green, University of Washington, Seattle. **Sequencher-Demo** (GeneCodes) for Macintosh was used for initial sequence alignments and display to facilitate primer design. Proposed primers were tested against the *L. monocytogenes* bank at Pasteur via Internet linkage.

Verifying Gaps with PCR from EGDe chromosomal DNA:

Conditions determined on case-by-case basis. Generally: Taq DNA polymerase, 94C denaturing for 30sec, 50C annealing for 30sec, 3 min extension at 72C, 20 min final extension after 30 cycles. Criteria: CLEAN products for purification for direct sequencing.

PROTOCOLS DISTRIBUTED FOR GENOME PROJECT

NOTE: Use materials specified for the Genome Project. Avoid Dust--use filtered solutions.

Store samples frozen, in the dark, the dyes are light sensitive. Label everything!

General Sequencing protocol for ABI:

Calculate the volume of template needed to make desired amount of DNA

Label the cycle-PCR tube

Place this DNA into the cycle-PCR tube

Dry DNA in speed vac (heat is OK for this). Assembly to fit in the speed vac:
nest the 0.2ml PCR-tube in a capless 0.5ml blue tube nested in another capless 1.5ml eppendorf tube.

While DNA dries, dilute primers: (Reaction vol. - vol. of BIGDYE = vol.containing total amount of primer in each reaction) *For example*: a 10 µl reaction with 2 µl BIGDYE,
30 to 100ng PCR product as template, 3 pmoles of primer....
3 pmoles = 8 µl of 0.375pmol / µl diluted primer.
(1.8 µl 10pmole/ µl stock into 48.2 µl HPLC water = 0.375 pmol / µl)

Double check before adding this:

Resuspend the correct DNA pellet with the correct diluted primer.

Add BIGDYE, close cap (close cap only once-- too fragile for repeats)

Cycling in Techne or PE2400 PCR machines:

Denature: 96C X 10sec
Anneal: 50C X 5sec (for normal primers) or 55C X 10sec (for long or high GC primers)
Extend: 60C X 4 minutes
25 cycles total, 4C hold.
Check program. Close lid. Start program , when temp = 96C, hit PAUSE button.
Put in samples, wait 30 - 60 seconds.
Hit PAUSE button to resume program.
Denature: 96C X 10sec

CLEAN UP methods for ABI samples:

2. Centri Sep column: (hydrated columns stored at 4C are good for a few days)

One column per sample

Gently tap column to make gel settle to bottom

Remove top cap, add 0.8ml HPLC water

Replace top cap, vortex column briefly to mix

Allow gel to hydrate 2 hours

Remove any bubbles by inverting and sharply tapping column, then let settle.

When bubble-free and settled, remove the top cap.

Then remove the bottom stopper

Allow gravity to drain into the 2ml tube, if this doesn't work, use a latex pipet bulb to apply gentle air pressure to the top of the column. Approximately 200-250 µl will drain out.

Discard this fluid, refit the 2 ml tube to column

Use the orientation of the tube, (line up bump facing out of the rotor)

Spin the column 3000rpm for 2 min in Hereaus microfuge

Discard the 2 ml tube.

Blot the end of tube.

(DO Not at any point, let the gel in the column get dry)

Immediately add sample in the center of the column bed (holding the column up to the light helps), avoid touching the gel, and do not let sample touch the sides of the column. Fit sample collection tube (capless 1.5ml tube) onto column. Place column assembly into microfuge with the orientation mark facing out.

Spin 3000rpm for 2 min

Dry samples in speedvac, about 20min. NO HEAT. Do not over-dry.

2. ETOH precipitation. (Back Up protocol)

0.1 volume 3M Sodium Acetate pH 5.2, 2.5 volume 95% ETOH, undenatured

Mix well, 15min at room temp.

Spin 20min at top speed in microfuge, room temp, (hinge of tube facing out).

Remove all supernatant without disturbing pellet (imaginary, in the bottom of the hinge side)

Wash with 250ul 70% room temp. ETOH, add gently!

Remove Supernatant, as above, careful to remove all.

Air dry 15 min in warm room (take precautions about dust!)

OR Dry in speedvac NO HEAT, 10 minutes. DO NOT OVER-DRY

FINAL PREPARATION:

Resuspend the cleaned-up and dried DNA in 20ul TSR

Transfer to ABI tube, top with rubber stopper.

Label with the designated **Name of the sequence**, (the name Paris wants to see).

No Bubbles in samples.

Rubber tops MUST BE must be placed, **totally flat** against the rim of the tube.

Please double check these 2 things, because failure can stop or break the ABI310 machine.

Store at 4C in the dark. Do Not Freeze.

Double / Triple check tube order and designations.

Deliver samples to the Virology or Hygiene sequencing facility:

as specified by the work sheet for the day, samples on ice in transit.

Chapter 5-D. Annotation tools on the Internet used in this thesis.

Inquiries into general, known protein banks and nucleic acid sequence comparisons were done with BlastX2, BlastP2 and BlastN2 in the NCBI, GenomeNet at University of Kyoto or Prodom databases.

URLs: http://www.ncbi.nlm.nih.gov/BLAST/ (NCBI)
http://www.blast.genome.ad.jp/ (GenomeNet at University of Kyoto)

Inquiries into the genomes of *L. monocytogenes* and *L. innocua* genomes were done at the Pasteur Institute server (these are temporary and not publicly available).

URLs: http://berenice.pasteur.fr/Listeria/genome.cgi (protein searches)
http://berenice.pasteur.fr/listeria/
(contig searches of incomplete *L. monocytogenes* genome)
http://berenice.pasteur.fr/innocua/
(contig searches of incomplete *L. innocua* genome)
http://genomeweb.pasteur.fr/lfrangeu/lm/cgi-bin/grepIPF.html
(keyword search tool for *L. monocytogenes* annotation data)
http://genomeweb.pasteur.fr/lfrangeu/lm/cgi-bin/IPF_reader.CGI?IPF_name=1669.1
(IPF reader, provide IPF information desired after "name=##")

Specific searches against defined databases:

URLs: http://genolist.pasteur.fr/Colibri/ (*E. coli* database: "Colibri")
http://genolist.pasteur.fr/SubtiList/ (*B. subtilis* database: "Subtilist")
http://motif.genome.ad.ip/ (GenomeNet at University of Kyoto, motif search)
http://www.toulouse.inra.fr/prodom/doc/blast_form.html (Protein Domains)
http://www.ncbi.nlm.nih.gov/COG/ (Protein families, evolutionary grouping)
http://ecocyc.PangeaSystems.com:1555/server.html
(*E. coli* metabolic enzymes, functions, definitions, networks)

Signal peptide predictions were performed using Signal P version 1.1 at the Technical University of Denmark database.

URL: http://www.cbs.dtu.dk/services/SignalP/

Blocks searches were done to identify signature blocks of amino acids within a specific class of enzymes. Blocks database is maintained by the Fred Hutchinson Cancer Research Center.

URL: http://blocks.fhcrc.org/blocks/blocks_search.html

Sequences were uploaded to EMBL.

URL: http://www.ebi.ac.uk/embl/Submission/webin.html

Literature searches of publications and published sequences.

URL: http://www.ncbi.nlm.nih.gov/Entrez/

Literature Cited

Antelmann, H., Bernhardt, J., Schmid, R., Mach, H., Volker, U., Hecker, M. 1997. First steps from a two-dimensional protein index towards a response-regulation map for *Bacillus subtilis*. *Electrophoresis* 18:1451-63.

Appelberg, R., Leal, I. S. 2000. Mutants of *Listeria monocytogenes* defective in *In vitro* invasion and cell-to-cell spreading still invade and proliferate in hepatocytes of neutropenic mice. *Infect Immun* 68:912-4.

Becker, L. A., Cetin, M. S., Hutkins, R. W., Benson, A. K. 1998. Identification of the gene encoding the alternative sigma factor sigmaB from *Listeria monocytogenes* and its role in osmotolerance. *J Bacteriol* 180:4547-54.

Bielecki, J., Youngman, P., Connelly, P., Portnoy, D. A. 1990. *Bacillus subtilis* expressing a haemolysin gene from *Listeria monocytogenes* can grow in mammalian cells. *Nature* 345:175-6.

Blaser, M. J., Newman, L. S. 1982. A review of human salmonellosis: I. Infective dose. *Rev Infect Dis* 4:1096-106.

Bockmann, R., Dickneite, C., Goebel, W., Bohne, J. 2000. PrfA mediates specific binding of RNA polymerase of *Listeria monocytogenes* to PrfA-dependent virulence gene promoters resulting in a transcriptionally active complex. *Mol Microbiol* 36:487-97.

Bockmann, R., Dickneite, C., Middendorf, B., Goebel, W., Sokolovic, Z. 1996. Specific binding of the *Listeria monocytogenes* transcriptional regulator PrfA to target sequences requires additional factor(s) and is influenced by iron. *Mol Microbiol* 22:643-53.

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

Bohne, J., Kestler, H., Uebele, C., Sokolovic, Z., Goebel, W. 1996. Differential regulation of the virulence genes of *Listeria monocytogenes* by the transcriptional activator PrfA. *Mol Microbiol* 20:1189-98.

Borezee, E., Msadek, T., Durant, L., Berche, P. 2000a. Identification in *Listeria monocytogenes* of MecA, a homologue of the *Bacillus subtilis* competence regulatory protein. *J Bacteriol* 182:5931-4.

Borezee, E., Pellegrini, E., Berche, P. 2000b. OppA of *Listeria monocytogenes*, an oligopeptide-binding protein required for bacterial growth at low temperature and involved in intracellular survival. *Infect Immun* 68:7069-77.

Braun, E. L., Fuge, E. K., Padilla, P. A., Werner-Washburne, M. 1996. A stationary-phase gene in *Saccharomyces cerevisiae* is a member of a novel, highly conserved gene family. *J Bacteriol* 178:6865-72.

Braun, L., Ghebrehiwet, B., Cossart, P. 2000. gC1q-R/p32, a C1q-binding protein, is a receptor for the InlB invasion protein of Listeria monocytogenes. *Embo J* 19:1458-66.

Braun, L., Nato, F., Payrastre, B., Mazie, J. C., Cossart, P. 1999. The 213-amino-acid leucine-rich repeat region of the *Listeria monocytogenes* InlB protein is sufficient for entry into mammalian cells, stimulation of PI 3-kinase and membrane ruffling. *Mol Microbiol* 34:10-23.

Braun, L., Ohayon, H., Cossart, P. 1998. The InIB protein of *Listeria monocytogenes* is sufficient to promote entry into mammalian cells. *Mol Microbiol* 27:1077-87.

Brehm, K., Kreft, J., Ripio, M. T., Vazquez-Boland, J. A. 1996. Regulation of virulence gene expression in pathogenic *Listeria* [published erratum appears in Microbiologia 1996 Sep;12(3):507]. *Microbiologia* 12:219-36.

Brundage, R. A., Smith, G. A., Camilli, A., Theriot, J. A., Portnoy, D. A. 1993. Expression and phosphorylation of the *Listeria monocytogenes* ActA protein in mammalian cells. *Proc Natl Acad Sci U S A* 90:11890-4.

Bubert, A., Kohler, S., Goebel, W. 1992a. The homologous and heterologous regions within the iap gene allow genus- and species-specific identification of *Listeria spp.* by polymerase chain reaction. *Appl Environ Microbiol* 58:2625-32.

Bubert, A., Kuhn, M., Goebel, W., Kohler, S. 1992b. Structural and functional properties of the p60 proteins from different *Listeria* species. *J Bacteriol* 174:8166-71.

Camilli, A., Goldfine, H., Portnoy, D. A. 1991. *Listeria monocytogenes* mutants lacking phosphatidylinositol-specific phospholipase C are avirulent. *J Exp Med* 173:751-4.

Camilli, A., Tilney, L. G., Portnoy, D. A. 1993. Dual roles of *plcA* in *Listeria monocytogene*s pathogenesis. *Mol Microbiol* 8:143-57.

Clewell, D. B., Flannagan, S. E., Jaworski, D. D. 1995. Unconstrained bacterial promiscuity: the Tn916-Tn1545 family of conjugative transposons. *Trends Microbiol* 3:229-36.

Collins, M. D., Wallbanks, S., Lane, D. J., Shah, J., Nietupski, R., et al. 1991. Phylogenetic analysis of the genus *Listeria* based on reverse transcriptase sequencing of 16S rRNA. *Int J Syst Bacteriol* 41:240-6.

Cossart, P., Vicente, M. F., Mengaud, J., Baquero, F., Perez-Diaz, J. C., Berche, P. 1989. Listeriolysin O is essential for virulence of *Listeria monocytogenes*: direct evidence obtained by gene complementation. *Infect Immun* 57:3629-36.

Cotter, P. D., Emerson, N., Gahan, C. G., Hill, C. 1999. Identification and disruption of *lisRK*, a genetic locus encoding a two-component signal transduction system involved in stress tolerance and virulence in *Listeria monocytogenes*. *J Bacteriol* 181:6840-3.

Cousens, L. P., Wing, E. J. 2000. Innate defenses in the liver during *Listeria* infection. *Immunol Rev* 174:150-9.

Czajka, J., Bsat, N., Piani, M., Russ, W., Sultana, K., et al. 1993. Differentiation of *Listeria monocytogenes* and *Listeria innocua* by 16S rRNA genes and intraspecies discrimination of *Listeria monocytogenes* strains by random amplified polymorphic DNA polymorphisms. *Appl Environ Microbiol* 59:304-8.

Dalton, C. B., Austin, C. C., Sobel, J., Hayes, P. S., Bibb, W. F., et al. 1997. An outbreak of gastroenteritis and fever due to *Listeria monocytogenes* in milk. *N Engl J Med* 336:100-5.

Davies, W. A. 1983. Kinetics of killing *Listeria monocytogenes* by macrophages: rapid killing accompanying phagocytosis. *J Reticuloendothel Soc* 34:131-41.

Decatur, A. L., Portnoy, D. A. 2000. A PEST-like sequence in listeriolysin O essential for *listeria monocytogenes* pathogenicity. *Science* 290:992-5.

Dhar, G., Faull, K. F., Schneewind, O. 2000. Anchor structure of cell wall surface proteins in *Listeria monocytogenes*. *Biochemistry* 39:3725-33.

Dickneite, C., Bockmann, R., Spory, A., Goebel, W., Sokolovic, Z. 1998. Differential interaction of the transcription factor PrfA and the PrfA- activating factor (Paf) of *Listeria monocytogenes* with target sequences. *Mol Microbiol* 27:915-28.

Domann, E., Chakraborty, T. 1989. Nucleotide sequence of the listeriolysin gene from a *Listeria monocytogenes* serotype 1/2a strain. *Nucleic Acids Res* 17:6406.

Domann, E., Leimeister-Wachter, M., Goebel, W., Chakraborty, T. 1991. Molecular cloning, sequencing, and identification of a metalloprotease gene from *Listeria monocytogenes* that is species specific and physically linked to the listeriolysin gene. *Infect Immun* 59:65-72.

Dons, L., Olsen, J. E., Rasmussen, O. F. 1994. Characterization of two putative *Listeria monocytogenes* genes encoding polypeptides homologous to the sensor protein CheA and the response regulator CheY of chemotaxis. *DNA Seq* 4:301-11.

Dons, L., Rasmussen, O. F., Olsen, J. E. 1992. Cloning and characterization of a gene encoding flagellin of *Listeria monocytogenes*. *Mol Microbiol* 6:2919-29.

Dramsi, S., Dehoux, P., Lebrun, M., Goossens, P. L., Cossart, P. 1997. Identification of four new members of the internalin multigene family of *Listeria monocytogenes* EGD. *Infect Immun* 65:1615-25.

Dramsi, S., Kocks, C., Forestier, C., Cossart, P. 1993. Internalin-mediated invasion of epithelial cells by *Listeria monocytogenes* is regulated by the bacterial growth state, temperature and the pleiotropic activator prfA. *Mol Microbiol* 9:931-41.

DuPont, H. L., Levine, M. M., Hornick, R. B., Formal, S. B. 1989. Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis* 159:1126-8.

Ebe, Y., Hasegawa, G., Takatsuka, H., Umezu, H., Mitsuyama, M., et al. 1999. The role of Kupffer cells and regulation of neutrophil migration into the liver by macrophage inflammatory protein-2 in primary listeriosis in mice. *Pathol Int* 49:519-32.

Eck, R. V., and M. O. Dayhoff. 1966. *Atlas of Protein Sequence and Structure.* Silver Springs, Maryland: National Biomedical Research Foundation

Engelbrecht, F. 1999. Die Genfamilie der kleinen Internaline in *Listeria monocytogenes* und *Listeria ivanovii*. In *Doctoral thesis, Lehrstuhl für Mikrobiologie*. Würzburg: Bayerischen Julius-Maximilians-Universität Würzburg

Engelbrecht, F., Chun, S. K., Ochs, C., Hess, J., Lottspeich, F., et al. 1996. A new PrfA-regulated gene of *Listeria monocytogenes* encoding a small, secreted protein which belongs to the family of internalins. *Mol Microbiol* 21:823-37.

Engelbrecht, F., Dickneite, C., Lampidis, R., Gotz, M., DasGupta, U., Goebel, W. 1998a. Sequence comparison of the chromosomal regions encompassing the internalin C genes (*inlC*) of *Listeria monocytogenes* and *L. ivanovii*. *Mol Gen Genet* 257:186-97.

Engelbrecht, F., Dominguez-Bernal, G., Hess, J., Dickneite, C., Greiffenberg, L., et al. 1998b. A novel PrfA-regulated chromosomal locus, which is specific for *Listeria ivanovii*, encodes two small, secreted internalins and contributes to virulence in mice. *Mol Microbiol* 30:405-17.

Farber, J. M., Peterkin, P. I. 1991. *Listeria monocytogenes*, a food-borne pathogen [published erratum appears in Microbiol Rev 1991 Dec;55(4):752]. *Microbiol Rev* 55:476-511.

Farber, J. M., Ross, W. H., Harwig, J. 1996. Health risk assessment of *Listeria monocytogenes* in Canada. *Int J Food Microbiol* 30:145-56.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17:368-76.

Flanary, P. L., Allen, R. D., Dons, L., Kathariou, S. 1999. Insertional inactivation of the *Listeria monocytogenes* cheYA operon abolishes response to oxygen gradients and reduces the number of flagella. *Can J Microbiol* 45:646-52.

Freitag, N. E., Portnoy, D. A. 1994. Dual promoters of the *Listeria monocytogenes prfA* transcriptional activator appear essential in vitro but are redundant in vivo. *Mol Microbiol* 12:845-53.

Freitag, N. E., Rong, L., Portnoy, D. A. 1993. Regulation of the *prfA* transcriptional activator of *Listeria monocytogenes*: multiple promoter elements contribute to intracellular growth and cell-to-cell spread. *Infect Immun* 61:2537-44.

Fuge, E. K., Braun, E. L., Werner-Washburne, M. 1994. Protein synthesis in long-term stationary-phase cultures of *Saccharomyces cerevisiae*. *J Bacteriol* 176:5802-13.

Gaillard, J. L., Berche, P., Frehel, C., Gouin, E., Cossart, P. 1991. Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive *cocci*. *Cell* 65:1127-41.

Gaillard, J. L., Berche, P., Sansonetti, P. 1986. Transposon mutagenesis as a tool to study the role of hemolysin in the virulence of *Listeria monocytogenes*. *Infect Immun* 52:50-5.

Gaillot, O., Pellegrini, E., Bregenholt, S., Nair, S., Berche, P. 2000. The ClpP serine protease is essential for the intracellular parasitism and virulence of *Listeria monocytogenes*. *Mol Microbiol* 35:1286-94.

Galsworthy, S. B., Girdler, S., Koval, S. F. 1990. Chemotaxis in *Listeria monocytogenes*. *Acta Microbiol Hung* 37:81-5.

Gedde, M. M., Higgins, D. E., Tilney, L. G., Portnoy, D. A. 2000. Role of listeriolysin O in cell-to-cell spread of *Listeria monocytogenes*. *Infect Immun* 68:999-1003.

Gilot, P., Andre, P., Content, J. 1999. *Listeria monocytogenes* possesses adhesins for fibronectin. *Infect Immun* 67:6698-701.

Goebel, W., Chakraborty, T., Domann, E., Kohler, S., Kuhn, M., et al. 1991. Studies on the pathogenicity of *Listeria monocytogenes*. *Infection* 19:S195-7.

Goldfine, H., Johnston, N. C., Knob, C. 1993. Nonspecific phospholipase C of *Listeria monocytogenes*: activity on phospholipids in Triton X-100-mixed micelles and in biological membranes. *J Bacteriol* 175:4298-306.

Gouin, E., Dehoux, P., Mengaud, J., Kocks, C., Cossart, P. 1995. *iactA* of *Listeria ivanovii*, although distantly related to *Listeria monocytogenes actA*, restores actin tail formation in an *L. monocytogenes actA* mutant. *Infect Immun* 63:2729-37.

Gouin, E., Mengaud, J., Cossart, P. 1994. The virulence gene cluster of *Listeria monocytogenes* is also present in *Listeria ivanovii*, an animal pathogen, and *Listeria seeligeri*, a nonpathogenic species. *Infect Immun* 62:3550-3.

Gregory, S. H., Sagnimeni, A. J., Wing, E. J. 1997. Internalin B promotes the replication of *Listeria monocytogenes* in mouse hepatocytes. *Infect Immun* 65:5137-41.

Greisen, K., Loeffelholz, M., Purohit, A., Leong, D. 1994. PCR primers and probes for the 16S rRNA gene of most species of pathogenic bacteria, including bacteria found in cerebrospinal fluid. *J Clin Microbiol* 32:335-51.

Guzman, C. A., Rohde, M., Chakraborty, T., Domann, E., Hudel, M., et al. 1995. Interaction of *Listeria monocytogenes* with mouse dendritic cells. *Infect Immun* 63:3665-73.

Haas, A., Dumbsky, M., Kreft, J. 1992. Listeriolysin genes: complete sequence of *ilo* from *Listeria ivanovii* and of *lso* from *Listeria seeligeri*. *Biochim Biophys Acta* 1130:81-4.

Haldenwang, W. G. 1995. The sigma factors of *Bacillus subtilis*. *Microbiol Rev* 59:1-30.

Hanawa, T., Yamamoto, T., Kamiya, S. 1995. *Listeria monocytogenes* can grow in macrophages without the aid of proteins induced by environmental stresses. *Infect Immun* 63:4595-9.

Hartford, T., Sneath, P. H. 1993. Optical DNA-DNA homology in the genus *Listeria*. *Int J Syst Bacteriol* 43:26-31.

He, W., Luchansky, J. B. 1997. Construction of the temperature-sensitive vectors pLUCH80 and pLUCH88 for delivery of Tn917::NotI/SmaI and use of these vectors to derive a circular map of *Listeria monocytogenes* Scott A, a serotype 4b isolate. *Appl Environ Microbiol* 63:3480-7.

Hess, J., Gentschev, I., Szalay, G., Ladel, C., Bubert, A., et al. 1995. *Listeria monocytogenes* P60 supports host cell invasion by and in vivo survival of attenuated *Salmonella typhimurium*. *Infect Immun* 63:2047-53.

Hodgson, D. A. 2000. Generalized transduction of serotype 1/2 and serotype 4b strains of *listeria monocytogenes*. *Mol Microbiol* 35:312-23.

Ishihama, A. 1997. Adaptation of gene expression in stationary phase bacteria. *Curr Opin Genet Dev* 7:582-8.

James B. Kaper, J. H., ed. 1999. *Pathogenicity Islands and Other Mobile Virulence Elements*. first ed. Washington, D.C.: ASM Press. 352 pp.

Jones, D. a. S., H. 1992. The Genus *Listeria*. In *The Prokaryotes : a handbook on the biology of bacteria : ecophysiology, isolation, identification, applications*, ed. e. a. Balows A. pp. 4 v. (xliii, 4126). Vol. v.2. New York: Springer-Verlag

Karunasagar, I., Lampidis, R., Goebel, W., Kreft, J. 1997. Complementation of *Listeria seeligeri* with the *plcA-prfA* genes from *L. monocytogenes* activates transcription of seeligerolysin and leads to bacterial escape from the phagosome of infected mammalian cells. *FEMS Microbiol Lett* 146:303-10.

Klarsfeld, A. D., Goossens, P. L., Cossart, P. 1994. Five *Listeria monocytogenes* genes preferentially expressed in infected mammalian cells: *plcA, purH, purD, pyrE* and an arginine ABC transporter gene, *arpJ*. *Mol Microbiol* 13:585-97.

Kobe, B., Deisenhofer, J. 1995. Proteins with leucine-rich repeats. *Curr Opin Struct Biol* 5:409-16.

Kreft, J., Dumbsky, M., Theiss, S. 1995. The actin-polymerization protein from *Listeria ivanovii* is a large repeat protein which shows only limited amino acid sequence homology to ActA from Listeria monocytogenes [published erratum appears in FEMS Microbiol Lett 1995 Oct 1;132(1-2):181]. *FEMS Microbiol Lett* 126:113-21.

Kroos, L., Yu, Y. T. 2000. Regulation of sigma factor activity during *Bacillus subtilis* development. *Curr Opin Microbiol* 3:553-60.

Kuhn, M., Goebel, W. 1989. Identification of an extracellular protein of *Listeria monocytogenes* possibly involved in intracellular uptake by mammalian cells. *Infect Immun* 57:55-61.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., et al. 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390:249-56.

Lampidis, R., M. Emmerth, I. Karunasagar, and J. Kreft. 2000. The virulence gene cluster from *Listeria seeligeri* contains large insertions and a partial gene duplication. *(manuscript in preparation)*

Lampidis, R., Gross, R., Sokolovic, Z., Goebel, W., Kreft, J. 1994. The virulence regulator protein of *Listeria ivanovii* is highly homologous to PrfA from *Listeria monocytogenes* and both belong to the Crp-Fnr family of transcription regulators. *Mol Microbiol* 13:141-51.

Lazazzera, B. A. 2000. Quorum sensing and starvation: signals for entry into stationary phase. *Curr Opin Microbiol* 3:177-82.

Lecuit, M., Dramsi, S., Gottardi, C., Fedor-Chaiken, M., Gumbiner, B., Cossart, P. 1999. A single amino acid in E-cadherin responsible for host specificity towards the human pathogen *Listeria monocytogenes*. *Embo J* 18:3956-63.

Lecuit, M., Ohayon, H., Braun, L., Mengaud, J., Cossart, P. 1997. Internalin of *Listeria monocytogenes* with an intact leucine-rich repeat region is sufficient to promote internalization. *Infect Immun* 65:5309-19.

Leimeister-Wachter, M., Domann, E., Chakraborty, T. 1991. Detection of a gene encoding a phosphatidylinositol-specific phospholipase C that is co-ordinately expressed with listeriolysin in *Listeria monocytogenes*. *Mol Microbiol* 5:361-6.

Leimeister-Wachter, M., Domann, E., Chakraborty, T. 1992. The expression of virulence genes in *Listeria monocytogenes* is thermoregulated. *J Bacteriol* 174:947-52.

Leimeister-Wachter, M., Haffner, C., Domann, E., Goebel, W., Chakraborty, T. 1990. Identification of a gene that positively regulates expression of listeriolysin, the major virulence factor of *Listeria monocytogenes*. *Proc Natl Acad Sci U S A* 87:8336-40.

Leimeister-Wachter M, H. C., Domann E, Goebel W, Chakraborty T. 1990. Identification of a gene that positively regulates expression of listeriolysin, the major virulence factor of *Listeria monocytogenes*. *Proc Natl Acad Sci U S A* 87:8336-40.

Lingnau, A., Domann, E., Hudel, M., Bock, M., Nichterlein, T., et al. 1995. 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* 63:3896-903.

Liu, J., Cosby, W. M., Zuber, P. 1999. Role of Lon and ClpX in the post-translational regulation of a sigma subunit of RNA polymerase required for cellular differentiation in *Bacillus subtilis*. *Mol Microbiol* 33:415-28.

Liu, J., Zuber, P. 2000. The ClpX protein of *Bacillus subtilis* indirectly influences RNA polymerase holoenzyme composition and directly stimulates sigma-dependent transcription. *Mol Microbiol* 37:885-97.

Loessner, M. J., Inman, R. B., Lauer, P., Calendar, R. 2000. Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. *Mol Microbiol* 35:324-40.

Loessner, M. J., Krause, I. B., Henle, T., Scherer, S. 1994. Structural proteins and DNA characteristics of 14 *Listeria* typing bacteriophages. *J Gen Virol* 75:701-10.

Lonetto, M. A., Brown, K. L., Rudd, K. E., Buttner, M. J. 1994. Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. *Proc Natl Acad Sci U S A* 91:7573-7.

Lorber, B. 1997. Listeriosis. *Clin Infect Dis* 24:1-9; quiz 10-1.

Ludwig, W., schleifer, K.H., and Stackebradt, E. 1984. 16S rRNA analysis of *Listeria monoctogenes* and *Brochothrix thermosphacta*. *FEMS Microbiology Letters* 25:199-204.

Ludwig, W., Seewaldt, E., Kilpper-Balz, R., Schleifer, K.H., Magrum, L., Woese, C.R., Fox, G.E., Stackebrandt, E. 1985. The phylogenetic position of *Streptococcus* and *Enterococcus*. *J Gen Microbiol* 131:543-51.

Ludwig, W., Strunk, O., Klugbauer, S., Klugbauer, N., Weizenegger, M., et al. 1998. Bacterial phylogeny based on comparative sequence analysis. *Electrophoresis* 19:554-68.

Ly, T. M., Muller, H. E. 1990a. Ingested *Listeria monocytogenes* survive and multiply in protozoa. *J Med Microbiol* 33:51-4.

Ly, T. M., Muller, H. E. 1990b. Interactions of *Listeria monocytogenes, Listeria seeligeri,* and *Listeria innocua* with protozoans. *J. Gen. Appl. Microbiol.* 36:143-150.

Majewski, J., Cohan, F. M. 1998. The effect of mismatch repair and heteroduplex formation on sexual isolation in *Bacillus*. *Genetics* 148:13-8.

Majewski, J., Cohan, F. M. 1999. DNA sequence similarity requirements for interspecific recombination in *Bacillus*. *Genetics* 153:1525-33.

Majewski, J., Zawadzki, P., Pickerill, P., Cohan, F. M., Dowson, C. G. 2000. Barriers to genetic exchange between bacterial species: *Streptococcus pneumoniae* transformation. *J Bacteriol* 182:1016-23.

Maniatis, T., Fritsch, F. and J. Sambrook. 1982. *Molecular Cloning: a laboratory manual* New York: Cold Spring Harbour Laboratory

Marino, M., Braun, L., Cossart, P., Ghosh, P. 1999. Structure of the lnlB leucine-rich repeats, a domain that triggers host cell invasion by the bacterial pathogen *L. monocytogenes*. *Mol Cell* 4:1063-72.

Marino, M., Braun, L., Cossart, P., Ghosh, P. 2000. A framework for interpreting the leucine-rich repeats of the *Listeria* internalins. *Proc Natl Acad Sci U S A* 97:8784-8.

Marquis, H., Doshi, V., Portnoy, D. A. 1995. The broad-range phospholipase C and a metalloprotease mediate listeriolysin O-independent escape of *Listeria monocytogenes* from a primary vacuole in human epithelial cells. *Infect Immun* 63:4531-4.

McLauchlin, J. 1990. Distribution of serovars of *Listeria monocytogenes* isolated from different categories of patients with listeriosis. *Eur J Clin Microbiol Infect Dis* 9:210-3.

Mengaud, J., Braun-Breton, C., Cossart, P. 1991a. Identification of phosphatidylinositol-specific phospholipase C activity in *Listeria monocytogenes*: a novel type of virulence factor? *Mol Microbiol* 5:367-72.

Mengaud, J., Geoffroy, C., Cossart, P. 1991b. Identification of a new operon involved in *Listeria monocytogenes* virulence: its first gene encodes a protein homologous to bacterial metalloproteases. *Infect Immun* 59:1043-9.

Mengaud, J., Lecuit, M., Lebrun, M., Nato, F., Mazie, J. C., Cossart, P. 1996. Antibodies to the leucine-rich repeat region of internalin block entry of *Listeria monocytogenes* into cells expressing E-cadherin. *Infect Immun* 64:5430-3.

Metzger, R., Brown, D. P., Grealish, P., Staver, M. J., Versalovic, J., et al. 1994. Characterization of the macromolecular synthesis (MMS) operon from *Listeria monocytogenes*. *Gene* 151:161-6.

Michel, E., Cossart, P. 1992. Physical map of the *Listeria monocytogenes* chromosome. *J Bacteriol* 174:7098-103.

Milenbachs, A. A., Brown, D. P., Moors, M., Youngman, P. 1997. Carbon-source regulation of virulence gene expression in *Listeria monocytogenes*. *Mol Microbiol* 23:1075-85.

Milohanic, E., Pron, B., Berche, P., Gaillard, J. L. 2000. Identification of new loci involved in adhesion of *Listeria monocytogenes* to eukaryotic cells. European *Listeria* Genome Consortium. *Microbiology* 146:731-9.

Missiakas, D., Raina, S. 1998. The extracytoplasmic function sigma factors: role and regulation. *Mol Microbiol* 28:1059-66.

Msadek, T. 1999. When the going gets tough: survival strategies and environmental signaling networks in *Bacillus subtilis*. *Trends Microbiol* 7:201-7.

Nair, S., Derre, I., Msadek, T., Gaillot, O., Berche, P. 2000a. CtsR controls class III heat shock gene expression in the human pathogen *Listeria monocytogenes*. *Mol Microbiol* 35:800-11.

Nair, S., Frehel, C., Nguyen, L., Escuyer, V., Berche, P. 1999. ClpE, a novel member of the HSP100 family, is involved in cell division and virulence of *Listeria monocytogenes*. *Mol Microbiol* 31:185-96.

Nair, S., Milohanic, E., Berche, P. 2000b. ClpC ATPase is required for cell adhesion and invasion of *Listeria monocytogenes*. *Infect Immun* 68:7061-8.

Nanamiya, H., Ohashi, Y., Asai, K., Moriya, S., Ogasawara, N., et al. 1998. ClpC regulates the fate of a sporulation initiation sigma factor, sigmaH protein, in *Bacillus subtili*s at elevated temperatures. *Mol Microbiol* 29:505-13.

Navarre, W. W., Schneewind, O. 1999. Surface proteins of gram-positive bacteria and mechanisms of their targeting to the cell wall envelope. *Microbiol Mol Biol Rev* 63:174-229.

Ochman, H., Lawrence, J. G., Groisman, E. A. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* 405:299-304.

Padilla, P. A., Fuge, E. K., Crawford, M. E., Errett, A., Werner-Washburne, M. 1998. The highly conserved, coregulated SNO and SNZ gene families in *Saccharomyces cerevisiae* respond to nutrient limitation [published erratum appears in J Bacteriol 1998 Dec;180(24):6794]. *J Bacteriol* 180:5718-26.

Parida, S. K., Domann, E., Rohde, M., Muller, S., Darji, A., et al. 1998. Internalin B is essential for adhesion and mediates the invasion of *Listeria monocytogenes* into human endothelial cells. *Mol Microbiol* 28:81-93.

Park, S. F., Kroll, R. G. 1993. Expression of listeriolysin and phosphatidylinositol-specific phospholipase C is repressed by the plant-derived molecule cellobiose in *Listeria monocytogenes*. *Mol Microbiol* 8:653-61.

Peel, M., Donachie, W., Shaw, A. 1988. Temperature-dependent expression of flagella of *Listeria monocytogenes* studied by electron microscopy, SDS-PAGE and western blotting. *J Gen Microbiol* 134:2171-8.

Piffaretti, J. C., Kressebuch, H., Aeschbacher, M., Bille, J., Bannerman, E., et al. 1989. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. *Proc Natl Acad Sci USA* 86:3818-22.

Poyart-Salmeron, C., Carlier, C., Trieu-Cuot, P., Courtieu, A. L., Courvalin, P. 1990. Transferable plasmid-mediated antibiotic resistance in *Listeria monocytogenes*. *Lancet* 335:1422-6.

Poyart-Salmeron, C., Trieu-Cuot, P., Carlier, C., MacGowan, A., McLauchlin, J., Courvalin, P. 1992. Genetic basis of tetracycline resistance in clinical isolates of *Listeria monocytogenes*. *Antimicrob Agents Chemother* 36:463-6.

Pron, B., Boumaila, C., Jaubert, F., Sarnacki, S., Monnet, J. P., et al. 1998. Comprehensive study of the intestinal stage of listeriosis in a rat ligated ileal loop system. *Infect Immun* 66:747-55.

Raffelsbauer, D., Bubert, A., Engelbrecht, F., Scheinpflug, J., Simm, A., et al. 1998. The gene cluster *inlC2DE* of *Listeria monocytogenes* contains additional new internalin genes and is important for virulence in mice. *Mol Gen Genet* 260:144-58.

Rasmussen, O. F., Skouboe, P., Dons, L., Rossen, L., Olsen, J. E. 1995. *Listeria monocytogenes* exists in at least three evolutionary lines: evidence from flagellin, invasive associated protein and listeriolysin O genes. *Microbiology* 141:2053-61.

Raveneau, J., Geoffroy, C., Beretti, J. L., Gaillard, J. L., Alouf, J. E., Berche, P. 1992. Reduced virulence of a *Listeria monocytogenes* phospholipase-deficient mutant obtained by transposon insertion into the zinc metalloprotease gene. *Infect Immun* 60:916-21.

Renzoni, A., Cossart, P., Dramsi, S. 1999. PrfA, the transcriptional activator of virulence genes, is upregulated during interaction of *Listeria monocytogenes* with mammalian cells and in eukaryotic cell extracts. *Mol Microbiol* 34:552-61.

Renzoni, A., Klarsfeld, A., Dramsi, S., Cossart, P. 1997. Evidence that PrfA, the pleiotropic activator of virulence genes in *Listeria monocytogenes*, can be present but inactive. *Infect Immun* 65:1515-8.

Ripio, M. T., Brehm, K., Lara, M., Suarez, M., Vazquez-Boland, J. A. 1997. Glucose-1-phosphate utilization by *Listeria monocytogenes* is PrfA dependent and coordinately expressed with virulence factors. *J Bacteriol* 179:7174-80.

Ripio, M. T., Dominguez-Bernal, G., Suarez, M., Brehm, K., Berche, P., Vazquez-Boland, J. A. 1996. Transcriptional activation of virulence genes in wild-type strains of *Listeria monocytogenes* in response to a change in the extracellular medium composition. *Res Microbiol* 147:371-84.

Ripio, M. T., Vazquez-Boland, J. A., Vega, Y., Nair, S., Berche, P. 1998. Evidence for expressional crosstalk between the central virulence regulator PrfA and the stress response mediator ClpC in *Listeria monocytogenes*. *FEMS Microbiol Lett* 158:45-50.

Robichon, D., Gouin, E., Debarbouille, M., Cossart, P., Cenatiempo, Y., Hechard, Y. 1997. The *rpoN* (sigma54) gene from *Listeria monocytogenes* is involved in resistance to mesentericin Y105, an antibacterial peptide from *Leuconostoc mesenteroides*. *J Bacteriol* 179:7591-4.

Rocourt, J., Francine G., Grimont, P., Seeliger, H. 1982. DNA relatedness Among Serovars of *Listeria monocytogenes sensu lato*. *Current Microbiology* 7:383-388.

Rocourt, J. 1988. Taxonomy of the genus *Listeria*. *Infection* 16:S89-91.

Rocourt, J., Hof, H., Schrettenbrunner, A., Malinverni, R., Bille, J. 1986. [Acute purulent *Listeria seelingeri* meningitis in an immunocompetent adult]. *Schweiz Med Wochenschr* 116:248-51.

Rocourt, J., Seeliger, H. P. 1985. [Distribution of species of the genus *Listeria*]. *Zentralbl Bakteriol Mikrobiol Hyg [A]* 259:317-30.

Rouquette, C., de Chastellier, C., Nair, S., Berche, P. 1998. The ClpC ATPase of *Listeria monocytogenes* is a general stress protein required for virulence and promoting early bacterial escape from the phagosome of macrophages. *Mol Microbiol* 27:1235-45.

Rouquette, C., Ripio, M. T., Pellegrini, E., Bolla, J. M., Tascon, R. I., et al. 1996. Identification of a ClpC ATPase required for stress tolerance and in vivo survival of *Listeria monocytogenes*. *Mol Microbiol* 21:977-87.

Rygus, T., Hillen, W. 1992. Catabolite repression of the *xyl* operon in *Bacillus megaterium*. *J Bacteriol* 174:3049-55.

Saitou, N., Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4:406-25.

Sallen, B., Rajoharison, A., Desvarenne, S., Quinn, F., Mabilat, C. 1996. Comparative analysis of 16S and 23S rRNA sequences of *Listeria species*. *Int J Syst Bacteriol* 46:669-74.

Schirmer, E. C., Glover, J. R., Singer, M. A., Lindquist, S. 1996. HSP100/Clp proteins: a common mechanism explains diverse functions. *Trends Biochem Sci* 21:289-96.

Schlech, W. F., 3rd. 1997. *Listeria* gastroenteritis--old syndrome, new pathogen. *N Engl J Med* 336:130-2.

Schneewind, O., Mihaylova-Petkov, D., Model, P. 1993. Cell wall sorting signals in surface proteins of gram-positive bacteria. *Embo J* 12:4803-11.

Siebers, A., Finlay, B. B. 1996. M cells and the pathogenesis of mucosal and systemic infections. *Trends Microbiol* 4:22-9.

Sivasubramaniam, S., Vanniasingham, V. M., Tan, C. T., Chua, N. H. 1995. Characterisation of HEVER, a novel stress-induced gene from *Hevea brasiliensis*. *Plant Mol Biol* 29:173-8.

Smith, G. A., Marquis, H., Jones, S., Johnston, N. C., Portnoy, D. A., Goldfine, H. 1995. The two distinct phospholipases C of *Listeria monocytogenes* have overlapping roles in escape from a vacuole and cell-to-cell spread. *Infect Immun* 63:4231-7.

Sokolovic, Z., Riedel, J., Wuenscher, M., Goebel, W. 1993. Surface-associated, PrfA-regulated proteins of *Listeria monocytogenes* synthesized under stress conditions. *Mol Microbiol* 8:219-27.

Sorokin, A., Bolotin, A., Purnelle, B., Hilbert, H., Lauber, J., et al. 1997. Sequence of the *Bacillus subtilis* genome region in the vicinity of the lev operon reveals two new extracytoplasmic function RNA polymerase sigma factors SigV and SigZ. *Microbiology* 143:2939-43.

Stackebrandt, E., and Woese, C.R. 1981. The evolution of prokaryotes. In *Molecular and cellular aspects of evolution. Society for General Microbiology, Symposium 32.* pp. 1-31. Molecular and cellular aspects of evolution. Society for General Microbiology, Symposium 32. Cambridge: Cambridge University Press

Strauch, M. A. 1993. AbrB, a Transition State Regulator. In *Bacillus subtilis and other Gram-Positive Bacteria*, ed. A. L. Sonenshein. pp. 757-764. Washington, DC: American Society for Microbiology

Tilney, L. G., Portnoy, D. A. 1989. Actin filaments and the growth, movement, and spread of the intracellular bacterial parasite, *Listeria monocytogenes*. *J Cell Biol* 109:1597-608.

Vaneechoutte, M., Boerlin, P., Tichy, H. V., Bannerman, E., Jager, B., Bille, J. 1998. Comparison of PCR-based DNA fingerprinting techniques for the identification of *Listeria species* and their use for atypical *Listeria* isolates. *Int J Syst Bacteriol* 48 Pt 1:127-39.

Vazquez-Boland, J. A., Kocks, C., Dramsi, S., Ohayon, H., Geoffroy, C., et al. 1992. Nucleotide sequence of the lecithinase operon of *Listeria monocytogenes* and possible role of lecithinase in cell-to-cell spread. *Infect Immun* 60:219-30.

von Both, U., Otten, S., Darbouche, A., Domann, E., Chakraborty, T. 1999. Physical and genetic map of the *Listeria monocytogenes* EGD serotype 1/2a chromosome. *FEMS Microbiol Lett* 175:281-9.

Vulic, M., Dionisio, F., Taddei, F., Radman, M. 1997. Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. *Proc Natl Acad Sci U S A* 94:9763-7.

Wagner, M., Schmid, M., Juretschko, S., Trebesius, K. H., Bubert, A., et al. 1998. In situ detection of a virulence factor mRNA and 16S rRNA in *Listeria monocytogenes*. *FEMS Microbiol Lett* 160:159-68.

Wang, R. F., Cao, W. W., Johnson, M. G. 1992. 16S rRNA-based probes and polymerase chain reaction method to detect *Listeria monocytogenes* cells added to foods. *Appl Environ Microbiol* 58:2827-31.

Wang, R. F., Cao, W. W., Wang, H., Johnson, M. G. 1993. A 16S rRNA-based DNA probe and PCR method specific for *Listeria ivanovii*. *FEMS Microbiol Lett* 80:85-92.

Wiedmann, M., Bruce, J. L., Keating, C., Johnson, A. E., McDonough, P. L., Batt, C. A. 1997. Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. *Infect Immun* 65:2707-16.

Wiedmann M, A. T., Hurley RJ, Boor KJ. 1998. General stress transcription factor sigmaB and its role in acid tolerance and virulence of *Listeria monocytogenes*. *J Bacteriol* :3650-6.

Williams, J. R., Thayyullathil, C., Freitag, N. E. 2000. Sequence variations within PrfA DNA binding sites and effects on *Listeria monocytogenes* virulence gene expression. *J Bacteriol* 182:837-41.

Wosten, M. M. 1998. Eubacterial sigma-factors. *FEMS Microbiol Rev* 22:127-50.

Wuenscher, M. D., Kohler, S., Bubert, A., Gerike, U., Goebel, W. 1993. The *iap* gene of *Listeria monocytogenes* is essential for cell viability, and its gene product, P60, has bacteriolytic activity. *J Bacteriol* 175:3491-501.

Zuber, U., Drzewiecki, K., Hecker, M. 2001. Putative Sigma Factor SigI (YkoZ) of Bacillus subtilis Is Induced by Heat Shock. *J Bacteriol* 183:1472-1475.

List of abbreviations.

A	Adenosine
Amp, $Amp^{R/S}$	Ampicillin, Ampicillin-Resistance/Sensitivity
ATCC	American Type Culture Collection, Rockville, MD, USA
ATP	Adenosine-triphosphate
B.	*Bacillus*
BHI	Brain Heart Infusion
bp	Basepairs
C	Cytosine
°C, C	Degree Celsius / Centigrade
CDS	Coding Sequence (sequence unit predicted to code for a protein)
cfu	colony forming units
C-Terminus	Carboxy-terminus
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleoside-triphosphate
ddNTPs	Di-deoxyribonucleoside-triphosphate
DTT	Dithiothreitol
E.	*Escherichia (coli)*
EDTA	Ethylene-diamine-tetraacetate
e.g.	for example (*L. exempli gratia*)
Erm, $Em^{R/S}$	Erythromycin, Erythromycin-Resistance/Sensitivity
et al.	and others
EtOH	Ethanol
G	Guanosine
gfp, GFP	green fluorescent protein
hr	hour
i.v.	intravenous
i.p.	intraperitoneal
IP	Isoelectric Point
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kb	Kilobase
kDa	Kilodalton
Kn, $Km^{R/S}$	Kanamycin, Kanamycin-Resistance/Sensitivity
L.	*Listeria*
L.	Liter
LB	Luria-Bertani
LD_{50}	50% letal dose
LRR	Leucine-rich repeat
M	Molar
MEM	Minimum Essential Medium
µl	micro liter
µg	micro gram
µM	micro Molar
ml	milli liter
mg	milli gram
mM	milli Molar
min	Minutes
MOI, moi	Multiplicity of infection (number of infecting bacteria per host cell)

mRNA	Messenger RNA
NCTC	National Collection of Type Culture, London, UK
ng	nano grams
nmoles	nano moles
SLCC	Special *Listeria* Culture Collection, Institut für Mikrobiologie und Hygiene, Würzburg
N-Terminus	Amino-terminus
OD_{260}	Optical density, (260) nm
ORF	Open reading frame
P	Phosphate
PBS	Phosphate-buffered Saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PenG	Penicillin G
per se	In or by itself, intrinsically (*L. per sé*)
PC-PLC	Phosphatidylcholine-specific Phospholipase C
PI-PLC	Phosphatidylinositol-specific Phospholipase C
pmol	pica moles
RBS, rbs	ribosomal binding site
rDNA	ribosomal RNA genes--chromosomal DNA derived sequence
RNA	Ribonucleic acid
rRNA	ribosomal RNA
rpm	rotations per minute
RT	Room temperature (20-25 degrees)
S	sedimentation units (e.g. 16S, 23S rRNAs)
sec	Seconds
spp.	species
Strep.	Streptomycine
subsp.	subspecies
Sv	Serovar
T	Thymidine
Temp	Temperature
Tet, $Tet^{R/S}$	Tetracycline, Tet Resistance/Sensitivity
Tris	Tris(Hydroxymethyl)-aminomethan
tRNA	transfer RNA
UV	Ultraviolet
vs.	versus
vol.	volume
X-Gal	5-Bromo-4-Chlor-Indolyl-β-D-Galactopyranoside
WT	wildtype

The *L. monocytogenes* Consortium Participants

(List may be Incomplete)

Institut Pasteur, Paris, France

Pascale Cossart
Lionel Frangeul
Carmen Buchrieser
Dehoux Pierre
Christophe Rusniok
Hafida Fsihi
Elisabeth Couvé
Farid Chétouani
Louis-Marin Jones
Aboubakar Maitournam
Frank Kunst
Philippe Glaser

Universität Würzburg, Germany

Werner Goebel
Jurgen Kreft
Eva Ng

Universidad Complutense, Madrid, Spain

Gustavo Domínguez-Bernal
Patricia Garrido
José Antonio Vázquez-Boland

Institute für Medical Microbiology, Giessen, Germany

Alexandra Amend
Trinad Chakraborty
Eugen Domann
Torsten Hain

Faculté de Médecine Necker-Enfants Malades, Paris, France

Lionel Durand
Alain Charbit
Patrick Berche

Hospital Ramon y Cajal, Madrid, Spain

Jose Claudio Pérez-Diaz
Fernando Baquero Mochales
Jorge Mata Vicente

Universidad Autonoma de Madrid, Spain

Nuria Gomez-Lopez
Encarna Maduenio
Beatriz de Pablos
Francisco Garcia Del-Portillo

GBF- Braunschweig, Germany

Jurgen Wehland
Helmut Bloecker
Petra Brandt
Gabriele Nordsiek

SRD, Frankfurt/Maine, Germany

Karl- Dieter Entian
Jörg Hauf
Matthias Rose

LION bioscience AG, Heidelberg, Germany

Bettina Remmel
Thomas Schluerer
David Jackson
Laurent Gautier
Günther Kurapkat
Antoine de Daruvar
Hartmut Voss

PUBLICATIONS

1. The *Listeria* Genome Consortium. 2001. The comparative genomes of *Listeria monocytogenes* and *L. innocua*. (manuscript in preparation).

2. Ng, E.Y.W., Schmid, M., Wagner, M., Goebel, W. 2001. Pathogenic capability has been lost at least twice during the evolution of the genus *Listeria*. (submitted to Infection, Genetics and Evolution).

3. Kreft, J., Vazquez-Boland, J., Ng, E., Goebel, W. 1999. Virulence gene clusters and putative pathogenicity islands in *Listeriae*. In Pathogenicity Islands and Other Mobile Virulence Elements. Eds. James B. Kaper and Jörg Hacker. Washington D.C.: ASM Press.

4. Coulter, S.N., Schwan, W.R., Ng, E.Y., Langhorne, M.H., Ritchie, H.D., Westbrock-Wadman, S., Hufnagle, W.O., Folger, K.R., Bayer, A.S., Stover, C.K. 1998. *Staphylococcus aureus* genetic loci impacting growth and survival in multiple infection environments. Mol. Microbiol. 30:393-404.

5. Schwan, W.R., Coulter, S.N., Ng, E.Y., Langhorne, M.H., Ritchie, H.D., Brody, L.L., Westbrock-Wadman, S., Bayer, A.S., Folger, K.R., Stover, C.K. 1998. Identification and characterization of the PutP proline permease that contributes to in vivo survival of *Staphylococcus aureus* in animal models. Infect. Immun. 66:567-72.

6. Breines, D.M., Quabdesselam, S., Ng, E.Y., Tankovic, J., Shah, S., Soussy, C.J., Hooper, D.C. 1997. Quinolone resistance locus *nfxD* of *Escherichia coli* is a mutant allele of the *parE* gene encoding a subunit of Topoisomerase IV. Antimicrob. Agents Chemother. 41:175-179.

7. Ng, E.Y.W., Trucksis, M., Hooper, D.C. 1996. Quinolone resistance mutations in Topoisomerase IV: Relationship to the *flq*A locus and genetic evidence that Topoisomerase IV is the primary target and DNA Gyrase is the secondary target of Fluoroquinolones in *Staphylococcus aureus*. Antimicrob. Agents Chemother. 40:1881-1888.

8. Ng, E.Y.W., Trucksis, M., Hooper, D.C. 1994. Quinolone resistance mediated by *norA*: Physiologic characterization and relationship to *flqB*, a Quinolone resistance locus on the *Staphylococcus aureus* chromosome. Antimicrob. Agents Chemother. 38:1345-1355.

9. Soussy, C.J., Wolfson, J.S., Ng, E.Y., Hooper, D.C. 1993. Limitations of plasmid complementation test for determination of Quinolone resistance due to changes in the Gyrase A protein and identification of conditional Quinolone resistance locus. Antimicrob. Agents Chemother. 37:2588-2592.

10. Hooper, D.C., Wolfson, J. S., Bozza, M., Ng, E.Y.W. 1992. Genetics and regulation of outer membrane protein expression by Quinolone resistance loci *nfxB*, *nfxC*, and *cfxB*. Antimicrob. Agents Chemother. 36:1151-1154.

11. McIsaac, R., Ng, E.Y.W., Nocente-McGrath, C., Ernst, S.G. 1992. Histone H2A.F/Z mRNA is stored in the egg cytoplasm and basally regulated in the sea urchin embryo. Dev. Biol. 153:402-406.

12. Hooper, D.C., Wolfson, J.S., Souza, K.S., Ng, E.Y., McHugh, G.L., Swartz, M.N. 1989. Mechanisms of Quinolone resistance in *Escherichia coli*: Characterization of *nfx* B and *cfx* B, two mutant resistance loci decreasing Norfloxacin accumulation. Antimicrob. Agents Chemother. 33:283-290.

13. Hooper, D.C., Wolfson, J.S., Ng, E.Y., Swartz, M.N. 1987. Mechanisms of action of and resistance to Ciprofloxacin. Am. J. Med. 82 (suppl 4A):12-20. (Review)

14. Wolfson, J.S., Hooper, D.C., Ng, E.Y., Souza, K.S., McHugh, G.L., Swartz, M.N. 1987. Antagonism of wild-type and resistant *Escherichia coli* and its DNA gyrase by the Tricyclic 4-Quinolone analogs Ofloxacin and S-25930 stereoisomers. Antimicrob. Agents Chemother. 31:1861-1863.

15. Wolfson, J.S., McHugh, G.L., Swartz, M.N., Ng, E.Y.W., Hooper, D.C. 1987. Visualization of the mitochondrion-kinetoplast in viable *Trypanosoma cruzi* with Rhodamine 123. J. Parasit. 73:866-869.