ABSTRACT

YILDIRIM, SÜLEYMAN. Molecular Characterization of *Listeria monocytogenes* Epidemic Clone I (ECI) like isolates from Food and Food Environments. (Under the direction of Dr. Sophia Kathariou.)

Listeria monocytogenes is an opportunistic foodborne pathogen, associated with serious morbidity and mortality in both humans and animals. A small number of genetically related strains of serotype 4b, designated epidemic clone I (ECI), have been involved in numerous epidemics of foodborne listeriosis. The ecology, transmission, and population structure of ECI strains from different sources, especially foods, have not been adequately addressed. In this study, fifty eight isolates of *L. monocytogenes* serotype 4b recovered from ready-to-eat foods and food processing environments were characterized. All isolates were found to harbor serotype 4b specific antigen determinants (*gltB, gtcA*) and reacted with monoclonal antibodies recognizing serotype 4b, 4d, and 4e antigens. The isolates were then screened for cytosine methylation at GATC sites in their genomes, and the accompanying resistance to digestion by Sau3AI, a well-established hallmark of ECI strains. Eighteen isolates were found to display resistance to Sau3AI digestion. Analysis of genome sequencing data for a prototype ECI strain enabled us to identify the genomic region harboring the putative restriction-modification system responsible for methylation at GATC sites. DNA probes were designed based on this region as well as other ECI-specific genomic regions, identified in the genome sequence. DNA-DNA

hybridizations demonstrated that all serotype 4b food strains whose genome was methylated also harbored these additional ECI- specific DNA fragments.

The eighteen ECI-like food isolates were further characterized by employing Pulsed-Field Gel Electrophoresis (PFGE) with AscI and ApaI, and DNA-DNA hybridizations in a macro-array format. The ECI outbreak strain (F2365) specific genes were spotted on nylon membranes and hybridized to genomic DNA of the ECI-like food strains. Analysis of PFGE and macroarray data indicated that the ECI-like food strains were unambiguously clustered together with confirmed outbreak strains, and were in a different cluster from other serotype 4b strains from foods or clinical cases that do not harbor ECI specific marker genes. Presence of an ECI outbreak strain (F2365)-specific genes in the genome of all eighteen ECI like strains, in addition to the PFGE fingerprints, suggest that these isolates were the descendants of one epidemic clone, the founder of the ECI clonal complex.

In the course of routine screening of food isolates we also identified eight strains of other serotypes (six serotype 1/2a and two serotype 1/2b) showing methylation at GATC sites. The restriction-modification cassette harbored by ECI strains was also harbored by several of these serotype 1/2a and 1/2b isolates, and sequence analysis of the genes from one of these strains (serotype 1/2a, isolated in 1999 from ready to eat food) revealed that the sequences were highly conserved. However, other ECI-specific genes and cassettes were absent from the genome of these strains. These findings suggested horizontal transfer of the restriction-modification cassette within *L. monocytogenes*.

The findings of this work suggest that ECI strains from food are closely related to those implicated in numerous human outbreaks. Regions specific to ECI strains of either clinical or food origin have been identified, which would facilitate application of several DNA-based tools for accurate detection and monitoring of strains of this epidemic clone in foods or the food processing environment.

MOLECULAR CHARACTERIZATION OF *LISTERIA MONOCYTOGENES* EPIDEMIC CLONE I (ECI) LIKE ISOLATES FROM FOOD AND FOOD ENVIRONMENTS

by

Süleyman Yıldırım

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

FOOD SCIENCE

Raleigh, NC
August, 2006

APPROVED BY:

Dr. Craig Altier

Dr. Todd R. Klaenhammer

Dr. Jeffrey Thorne

Dr. LeeAnn Jaykus
Co-chair of Advisory Committee

Dr. Sophia Kathariou
Chair of Advisory Committee

To my parents,

Feride and İsmail Yıldırım,

who taught me how and why I should learn

BIOGRAPHY

Süleyman Yıldırım was born in a lovely industrial city in Black Sea Region, Karabük- TURKEY to İsmail and Feride Yıldırım. He lived most of his teenage years and attended public schools in this city until 1989, the time he decided to pursue a college degree in engineering. He successfully passed the university entrance exam, which placed him in Food Engineering, a serendipity which determined the career path he has followed up until he has written this dissertation. He attended English prep class for one year before he began his studies at Hacettepe University located in the capital city, Ankara. He obtained his B.S. degree in 1994 and was admitted into the Graduate School in Food Biotechnology program in the Department of Food Technology at Ankara University. He was awarded Ministry of the Turkish National Education fellowship to pursue graduate degree in the United States in 1995. In a hot and muggy day, he landed in Iowa ("You make me smile") on May 18, 1996 where he participated in a research project on transgenic corn at Iowa State University and obtained his Masters of Science degree under the mentorship of Dr. Zivko Nikolov. He then spent some time working for food and pharmaceutical companies, respectively, before he joined Dr. Sophia Kathariou's research group to pursue doctoral degree at North Carolina State University in Food Science with a minor Functional Genomics.

ACKNOWLEDGEMENTS

I would like to extend my sincerest appreciation to those individuals, who, through their support, friendship or love, have been the real heroes behind the scene.

At first, I would like to thank Dr. Sophia Kathariou, my mentor, for her support, patience with, and trust in me; I am sure she had been aware that she took a serious risk by accepting me into her laboratory since I came in with limited prior exposure to Molecular Microbiology. I am grateful to her for allowing me to pursue my dreams during my graduate studies and for her generosity. She has been not only a mentor but also a friend and like my older sister at times. Over the years, I have also come to value her style; humble, compassionate, and a true friend of nature as well as her commitment to science and dedication to science education. I also observed that she tries hard to tailor the project according to her student's scientific interests and curiosity.

I was honored and felt privileged by having well-rounded scientists in my dissertation committee; Drs. Todd R. Klaenhammer, Lee-Ann Jaykus, Craig Altier, and Jeffrey Thorne. I just regret that I could not sufficiently take advantage of their expertise. I sincerely thank them all, for their contribution to my academic training.

I also thank Dr. Driss Elhanefi for his friendship and helping me learn genetic manipulations of genes in prokaryotic systems.

Also, Robin Slatzky in our laboratory and Evelyn Durmaz in Dr. Klaenhammer laboratory deserve special recognition; these two hard-working women are assets to the laboratory they manage. They play crucial role in organizing events and managing. More importantly, they perform these functions well and happily.

Finally, I wish to thank my parents for their continued support, prayers, and encouragement all my life. I can never thank them enough for the sacrifices they have made for my education. My father taught me how to read even before I started primary school while he was working on the floor of a steel mill and set a role model for me by reading volumes of books almost on a daily basis. They also taught me the importance of pursuing “ilim” (wisdom) throughout ones life.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF FIGURES

(Continued)

Page

IDENTIFICATION AND CHARACTERIZATION OF A TYPE-II RESTRICTION-MODIFICATION SYSTEM SPECIFIC TO *LISTERIA MONOCYTOGENES* EPIDEMIC CLONE SEROTYPE 4B IN THE GENOMES OF NON-4B STRAINS ISOLATED FROM FOODS AND FOOD ENVIRONMENTS.

*Both the natural and the experimental disease have interesting and characteristic features and their consideration has forced us to the conclusion that the causative organism either has not been described previously, or has been inadequately described and so cannot be traced in the literature. In either case, we feel justified in naming it. Its salient character is the production of a large mononuclear leucocytosis. This is far the most important and most striking character we have discovered and we name the microorganism we shall describe in this paper "**Bacterium monocytogenes**".* […]

Murray, **E.G.D.**., R.A. Webb, and M.B.R. Swann. **1926**. A disease of rabbit characterized by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus *Bacterium monocytogenes*. J. Pathol. Bacteriol. 29:407

CHAPTER I

LITERATURE REVIEW

INTRODUCTION

Listeria monocytogenes is one of the six recognized species of the genus *Listeria*. The phylogenetic position of the genus *Listeria* is close to *Bacillus, Clostridium, Enterococcus, Streptococcus,* and *Staphylococcus* (Sallen et al., 1996). Phylogenetic analyses based on the sequences of selected genes (Fig. 1) suggested that these six species were divided into two lines of descent: (i) *L. grayi* and (ii) the remaining five *Listeria* species: *L. monocytogenes*, *L. innocua*, *L. ivanovii*, *L. seeligeri*, and *L. welshimeri*. Furthermore, the latter group was consistently subdivided into two distinct subgroups: *L. monocytogenes* and *L. innocua* formed a separate group from the remaining three species. *L grayi* was the species that was most distantly related to any of the others, regardless to the tree algorithm used (Schmid et al., 2005).

L. monocytogenes are gram-positive, non-spore forming, facultative anaerobic rods of 0.4 by 1 to 1.5 µm that have no capsule, and are optimally motile at 10 to 25 °C. *L. monocytogenes* can grow between -0.4 and 50 °C and utilize a wide range of carbon sources, including C5 and C6 sugars (Pine et al., 1990; Farber and Peterkin, 1991). It can also survive and even proliferate under a variety of unfavorable conditions such as low temperatures and at high salt concentrations (up to 10% sodium chloride) by accumulating compatible solutes within the cell (Sleator et al., 2003). The bacterium expresses β-hemolysin, which produces zones of clearing (hemolysis) on blood agar. The natural habitat of these bacteria is thought to be decomposing plant matter. Hence, strains of *L. monocytogenes* are isolated from a wide range of environmental sources, including decaying vegetation, soil, water, effluents, a large variety of foods, and the feces of humans and animals. Domesticated ruminants are thought to play a key role in the maintenance of *Listeria spp.* in the rural environment via a continuous fecal-oral enrichment cycle (Vazquez-Boland et al., 2001).

Population Structure. *L. monocytogenes* has a strongly clonal population based on linkage equilibrium analysis. In addition, genetically indistinguishable isolates have been recovered in diverse parts of the world. The observed clonal structure in *L. monocytogenes* suggests that horizontal gene flow between the different serotype-associated clonal lineages is limited (Pifaretti et al., 1989; Call et al., 2003).

Thirteen serotypes of *Listeria* have been recognized on the basis of somatic (O) and flagellar (H) antigens: 1/2a, 1/2b, 1/2c, 3a, 3b, 3c, 4a, 4ab, 4b, 4c, 4d, 4e, and 7. Molecular and other typing methods including Multi Locus Enzyme Electrophoresis (MLEE), ribotyping, DNA arrays, and Pulsed-Field Gel Electrophoresis (PFGE) have consistently classified the strains of *L. monocytogenes* into distinct phylogenetic lineages (also termed genomic divisions) separating serotypes 4b, 3b and 1/2b (lineage II, or genomic division II) from 1/2a, 3a, 3c and 1/2c (lineage I, or genomic division I) and serotype 4a and 4c (lineage III) (Pifaretti et al., 1989; Brosch et al., 1996; Wiedmann et al., 1997; Nadon et al., 2001; Doumith et al., 2004). Thus, the antigenic structure of *L. monocytogenes* correlates well with the lineages of the organism. Although clonal lineages are common in bacterial pathogens, the clonal partitioning of the species along serotypic groups is not routinely seen in other pathogens that possess multiple serotypes.

Even though several serotypes have been identified in foods and in environmental samples, only three serotypes (4b, 1/2b, and 1/2a) are responsible for more than 95% of the outbreaks. Furthermore, the majority of *L. monocytogenes* strains found in conjunction with large outbreaks of human food-borne listeriosis are serotype 4b. Unlike serotype 4b strains, serotype 1/2c strains are often recovered from foods and food processing environments, but rarely cause human infection. Serotype 1/2a strains are also overrepresented among food isolates. However, serotype 4b strains are overrepresented in clinical isolates and this discrepancy is still not well understood (Farber and Peterkin, 1991; Rocourt et al., 1997; Kathariou, 2003).

Subtyping systems for *L. monocytogenes*. Several typing schemes have been applied on *L. monocytogenes* strains. The following factors have been used to determine the value of a typing method: typeability of a strain, reproducibility, discrimination index, resolution, lab-to-lab variation, standardization, portability, cost, and practicality (Hunter, 1990). Traditionally, serotyping (Seeliger and Hohne, 1979), phage typing, and Multilocus Enzyme Electrophoresis (MLEE) (Boerlin and Piffaretti, 1991) have been extensively used in order to subtype *L. monocytogenes* strains. Recently, molecular subtyping methods, such as Ribotyping (Gendel et al., 2000), and Pulsed-Field Gel Electrophoresis (PFGE) (Graves et al., 1994) as well as sequence-based methods such as Multilocus Sequence Typing (MLST) (Salcedo et al., 2003; Revazishvili et al., 2004) have been adapted and are being increasingly used for subtyping purposes.

Serotyping. *Listeria* species are divided into serotypes based on somatic (O) and flegallar (H) antigens. Unlike most other well-known pathogenic bacteria. in *L. monocytogenes* the antigenic structure correlates well with its epidemiological and population structure. However, as mentioned above. more than 95% of clinical isolates belong to only three serotypes (1/2a, 1/2b, and 4b) (Mead et al., 1999). Also, some strains are not typeable. Therefore, serotyping per se is of limited practical value in epidemiological investigations. However, because only three serotypes are known to be involved in most of the clinical cases serotyping is of value in that the strains of three serotypes can be rapidly screened for further analysis.

Phage Typing. Serotyping and phage typing of *L. monocytogenes* strains have been extensively used in epidemiological investigations before molecular subtyping methods became available (Graves et al., 1999). A standard set of phages are utilized to assign phage types to isolates according to their susceptibility. Nevertheless, standardization of phage typing is highly problematic due to significant biologic and experimental variability (Wiedmann, 2002). In addition, the WHO Multicenter *L. monocytogenes* subtyping study reported that 20-51% of the isolates were untypeable using an international phage set (McLauchlin et al., 1996). On the other hand, phage typing has the advantages of high discriminating power, low cost, and easy applicability to large numbers of strains in outbreak investigations.

Multi-Locus Enzyme Electrophoresis (MLEE). MLEE is based on separation of metabolic enzymes in a gel matrix (non-denaturing starch gels) and comparison of mobility of the same set of enzymes from different strains. The observed differences stem from amino acid substitutions in polypeptide sequences which in turn result from nucleotide substitutions in the gene sequence encoding the proteins, and which change the charge of the protein. Mobility variants of each enzyme are considered to be different electromorphs and combinations of a set of electromorphs represent an electrophoretic type (ET) or multilocus genotype (Graves et al., 1999; Wiedmann, 2002). Most strains are typeable using this method, and this method has been frequently used for population genetics studies of pathogenic bacteria (Moller et al., 1992; Johnson et al., 1997) as well as epidemiological investigations of listeriosis outbreaks (Pifaretti et al., 1989; Boerlin et al., 1991;Norrung and Skovgaard, 1993). In spite of being a powerful tool for population genetic, taxonomic, and evolutionary studies, it often appears to be less discriminatory than other DNA-based subtyping methods used in epidemiological investigations. Also, it is labor-intensive and standardization between different laboratories is difficult (Wiedmann, 2002).

Ribotyping. Polymorphisms associated with ribosomal RNA operons are detected in DNA-DNA hybridizations employing Southern hybridization. An rRNA gene probe is appropriately labeled and is allowed to hybridize on a nylon membrane. For the most part *Eco*RI has been the preferred restriction enzyme to digest the genomic DNA of *L. monocytogenes* (Graves et al., 1999). Ribotyping has been extensively employed for subtyping *L. monocytogenes* (Bruce et al., 1995; Graves et al., 1994; Wiedmann et al., 1997). The Riboprinter (DuPont-Qualicon, Wilmington, DE) is an automated ribotyping system that produces and analyzes ribotyping patterns of bacteria. Therefore, this method has the advantage of automation and standardization. However, resolution is relatively limited, and ribotyping using the Riboprinter is typically limited to reference laboratories due to the high cost of the equipment.

Pulsed-Field Gel Electrophoresis (PFGE). In this method, large fragments of intact genomic DNA which are obtained by embedding live cells in agarose gel plugs are separated in a periodically changing electrical field across the gel. Large DNA fragments are the result of digestion of unsheared genomic DNA released from the cells within the agarose blocks by one or more restriction endonucleases whose recognition sites are infrequently distributed around the genome. Subsequent Restriction Fragment Length Polymorphism (RFLP) analysis using computers helps to discriminate related isolates from unrelated ones. PFGE was successfully applied to *L. monocytogenes* for the first time in the early nineties and high discrimination of closely related strains of serotype 4b was obtained (Brosch et al., 1991). In addition, genotypic data generated by PFGE correlate well with the phenotypic data generated by serotyping for establishing the overall relatedness of isolates (Brosch et al., 1994). This method offers high discriminatory power, reproducibility between different laboratories and is standardized in the case of *L. monocytogenes* (Graves et al., 1999; 2001). The restriction enzymes AscI and ApaI are routinely used to digest genomic DNA of *L. monocytogenes*. Compared to other subtyping methods, PFGE has yielded the highest discrimination index (Louie et al., 1996; Borucki et al., 2003; 2004) and thus is considered to be the "Gold Standard" for subtyping *L. monocytogenes.* In addition, the availability of a U. S. national network (PulseNet consisting of state health departments, local health departments, and federal agencies (CDC, USDA/FSIS, FDA) for comparison of subtypes makes PFGE the preferred subtyping method. However, novel DNA sequence-based methods are currently under development, which may eventually replace PFGE (Gerner-Smidt et al., 2006).

DNA Sequence-Based Subtyping. The major advantage of this approach over band-based methods is that DNA sequence data can be analyzed with greater accuracy and can be readily shared with other laboratories around the world via the internet (i.e. they are "portable") (Spratt et al., 1999). Recent advances in genome sequence technology have facilitated sequencing of entire genomes, often from multiple strains of the same species of a bacterium. In parallel to this several sequence-based subtyping methods, such as Multilocus Sequence Typing (MLST) and Variable Number Tandem Repeat (VNTR) have been developed. In general, MLST refers to sequencing of multiple housekeeping genes (Maiden et al., 1998); however, sequencing of multiple virulence genes can also be similarly used as a subtyping method (Zhang et al., 2004). *L. monocytogenes* is an attractive candidate for DNA-sequenced based subtyping since there is an excellent set of well-characterized, epidemiologically related and unrelated strains of this bacterium that have been already included in other subtyping methods (Graves et al., 1999). Not surprisingly, several candidate loci in the genome of *L. monocytogenes* have been identified and shown to be highly discriminatory in MLST schemes compared to other subtyping methods (Cai et al., 2002; Revazishvili et al., 2004; Meinersmann et al., 2004). However, DNA-sequence based subtyping approaches for *L. monocytogenes* or foodborne pathogenic bacteria at large are still in the early developmental stages and optimal target genes are still being defined for these pathogens.

Key Aspects of *Listeria* Epidemiology. *L. monocytogenes* is responsible for severe food-borne infections with a high case fatality rate (ca. 30%) in susceptible animal and human hosts. Human listeriosis occurs as sporadic disease or in the form of outbreaks (epidemics) which can affect certain segments of the population, especially the elderly, neonates, pregnant women or immunocompromised individuals. While rarely diagnosed prior to 1960, more than 10,000 cases of listeriosis were recorded in the medical literature between 1960 and 1982 (Rocourt, 1991). Epidemiologic and laboratory investigations in 1983 following an outbreak in Nova Scotia showed for the first time that foods served as vehicle for the transmission of this pathogen (Schlech et al., 1983). Most commonly implicated in listeriosis are deli and other processed meats, soft cheeses, and processed seafood that are eaten without cooking or reheating (Kathariou, 2002).

The majority of outbreaks in Europe and North America have been caused by a small number of closely related serotype 4b strains. Also, serotype 4b strains have been implicated in a substantial fraction of sporadic infections. The presence of bacteria of this serotype in an RTE food appears to increase the food's risk of implication in listeriosis (Pinner et al., 1992). The other key serotypes that contribute to the burden of listeriosis are serotypes 1/2a and 1/2b.

EVOLUTIONARY AND GENOMICS STUDIES OF *L. MONOCYTOGENES*

Lineages and genetic diversity within each lineage. *L. monocytogenes* strains display both genetic and serotypic diversity. As mentioned above, phylogenetic analyses have shown that antigenic structure of this organism correlates well with the lineage or genomic division; serotype 4b, 1/2b, and 3b strains are found mostly in lineage I and serotype 1/2a, 1/2c, 3a, and 3c strains are found predominantly in lineage II (Bibb et al., 1989; Brosch et al., 1994; Rasmussen et al., 1991). A third lineage, lineage III (serotypes 4a and 4c) was proposed due to the genetic distance of serotypes 4a and 4c based on ribotype and virulence gene polymorphism data (Rasmussen et al., 1995; Wiedmann et al., 1997). Phylogenetic data suggest that some serotype 4b strains were included in this lineage and subgroups of these strains were reported to show unusual phenotypes compared to other serotype 4b strains, such as lack of the ability to ferment rhamnose and absence of putative virulence gene (*lmaA*) (Ward et al., 2004; Roberts et al.,2006).

Strains of *L. monocytogenes* in different lineages display significant differences in terms of prevalence in human and animal listeriosis. Phylogenetic analysis of polymorphisms in different housekeeping genes and the *prfA* virulence gene cluster indicated that the strains of lineage I seem to be strongly clonal, in contrast to strains of other lineages (Ward et al., 2004; Nightingale et al., 2005). Lineage II showed greater diversity and evidence of horizontal gene transfer. Genetic exchange between the two lineages appears to be rare. Genes from lineage III organisms were mostly similar to lineage I genes, with instances of genes appearing to be mosaics with lineage II genes. In fact, both neighbor-joining and maximum-parsimony analyses of the combined *prfA* virulence gene cluster data strongly support a close genetic relationship between lineages I and III (Meinersmann et al., 2004; Ward et al., 2004; Nightingale et al., 2005). Nevertheless, Lineage I isolates tend to be more frequently found among human cases than among animal cases, while lineage III isolates are more prevalent in animal than in human cases due possibly to host specificity for non-primate mammals (Jeffers et al., 2001). Furthermore, lineage I isolates tend to show greater cytopathogenicity in vitro than lineage II isolates (Pohl et al., 2006). The frequency that the bacteria are associated with cases of encephalitis, septicemia, and fetal infection may differ between lineages I and II, possibly because allelic forms of virulence genes such as *inlA* and *ActA* may be associated with distinct outcomes of the bacterial infection (Pohl et al., 2006). These observations may be explained to a certain extent by differences in evolutionary histories of lineage I and II organisms, i.e., horizontal gene transfer, positive selection of virulence genes, and recombination (Meinersmann et al., 2004; Ward et al., 2004; Nightingale et al., 2005).

Serotype-specific differences. Recent studies of different serotypes by comparative genomic sequence analysis and subtractive hybridization suggest that lineage and serotype-specific patterns of genome diversity exist among *L. monocytogenes* strains. Comparative genomic analyses using both subtractive libraries and DNA arrays revealed that as much as 5 to 8 % of the serotype 4b genome is not present in serotype 1/2a strains (Herd and Kocks, 2001). Several independent comparative genome sequencing studies of *L. monocytogenes* consistently show a collective bias toward cell surface-related differences in genome content among the *L. monocytogenes* serotypes and *Listeria* species. In particular, internalin family proteins are often variable in different serotypes. Serotype 4a strains, for example, do not harbor genes encoding the *inlGHE* cluster (Doumith et al., 2004). In addition to cell surface molecules many putative transcription regulators specific to lineage II were identified. Based on this finding it was hypothesized that lineage-specific gene expression might contribute to the differences in terms of host-pathogen interaction (Zhang et al., 2003; Call et al., 2004) and one such genome segment belonging to lineage II (*lstR* –lineage specific thermal regulator) has been recently characterized (Zhang et al., 2005).

GENOMES AND PRIMARY FEATURES

Comparative genomics of *L. monocytogenes* and *L. innocua*. *L. monocytogenes* and *L. innocua* are closely related species and are often isolated from the same ecological habitat. If isolated together, *L. innocua* can overgrow *L. monocytogenes* in enrichment broths, which may preclude *L. monocytogenes* from being detected (Gnanou et al., 2005). A line of evidence suggests that the virulence gene cluster (LIPI-1) was present in the common ancestor of *L. monocytogenes, L. innocua, L. ivanovii, L. seeligeri,* and *L. welshimeri* and that this pathogenicity island has been lost in two separate genetic events in *L. innocua* and *L. welshimeri* (Johnson et al, 2004;Schimid et al., 2005) while stabilized in *L. monocytogenes and L. ivanovii.* Intriguingly, LIPI-1 is kept in *L. seeligeri* genome in intact but inactivated form, suggesting a relatively recent event (Vaquez-Boland et al., 2001).

The first *L. monocytogenes* genome was sequenced in 2001 and its sequence was compared to the non-pathogenic counterpart (*L. innocua*) to explain the evolution of *Listeria* species and emergence of the pathogen (Glaser et al., 2001). The strains chosen for this purpose were *L. monocytogenes* EGD-e (serotype 1/2a), a derivative of the first *L. monocytogenes* isolate which was isolated from a rabbit (Murray et al., 1926) and *L. innocua* strain CLIP 11262 (serovar 6a). Both genomes encoded many putative surface and secreted proteins, transporters, and transcriptional regulators, consistent with the ability of both species to survive under diverse environmental stresses and colonize various ecological niches (Glaser et al., 2001). Nevertheless, 30 of the 133 surface proteins identified in *L. monocytogenes* EGD-e were absent from *L. innocua* and 20 of these 30 *L. monocytogenes* EGD-e specific surface proteins harbored the LPXTG motif typical of surface proteins of Gram-positive bacteria. This observation suggested that *L. monocytogenes* EGD-e may be able to better adapt to different ecological niches than *L. innocua.* In addition to surface proteins, a rich collection of putative regulatory proteins, the largest family being BglG regulatory proteins, and transport proteins, especially those of the phosphoenolpyruvate-dependent phosphotransferase (PTS) system may account for *L. monocytogenes* virulence and its ability to adapt, colonize, and respond to a wide variety of different environments and its virulence (Glaser et al., 2001; Buchrieser et al., 2003). Comparative genomic analysis between the two genomes revealed 270 (9.5%) *L. monocytogenes* EGD-e specific genes and 149 (5%) *L. innocua* specific genes. The distribution of these genes within the different functional categories showed that the pathogenic *L. monocytogenes* strain EGD-e possesses about twice as many cell wall, transport, and putative transcriptional regulatory proteins as its non-pathogenic counterpart (Glaser et al., 2001).

Comparison of the two genomes revealed a synteny of the ORFs and little inversion and/or recombination. In addition, only few insertion sequences (IS) in both genomes were found, which may partially account for the conservation of the orthologous genes and limited chromosomal rearrangement. Even though both genomes contained phage genes they did not seem to have played a critical role in the emergence of the pathogenic species. However, strain-specific regions seem to have been contributed by mobile genetic elements. Interestingly, only 21% of the EGD-e entire genome (616 Kb) harbored almost half of the total strain-specific regions, which suggested that this region may be a recombinational hot spot (Buchrieser et al., 2003).

The G+C content of *L. monocytogenes* strain-specific genes varied from 24 to 46% and the average G+C content for the entire genome is 38%. Similarly, many of the *L. innocua* specific surface proteins were located in regions with distinct G+C content (Glaser et al., 2001). Thus, strain-specific regions in both genomes could be due to horizontal transfers from other species. On the other hand, *L. innocua* CLIP 11262 carried a 100-kb plasmid (pLM100) not present in *L. monocytogenes* EGD-e. The plasmid harbored genes putatively encoding heavy metal resistance and shared strong similarity in several loci with plasmid pLM80 from the recently sequenced serotype 4b strain (H7858), which caused a multistate outbreak in 1998-1999 in the USA (*MMWR Weekly*; Nelson et al., 2004).

Whole genome comparisons of serotype 4b and 1/2a strains of *L. monocytogenes.* A comparative genomic approach with multiple strains was taken by Nelson et al, 2004, three years after the first *L. monocytogenes* genome was published. The strains chosen for genome sequencing represented two serotypes (1/2a and 4b) that were associated with human foodborne illness. More specifically, *L .monocytogenes* strain F2365 (serotype 4b) was a cheese isolate from the Jalisco cheese outbreak of 1985 in California (Linnan et al., 1988); *L.monocytogenes* strain F6854 (serotype 1/2a) was isolated from turkey frankfurters implicated in a sporadic case in 1988 in Oklahoma, representing a PFGE type indistinguishable from that associated with a strain causing a multistate outbreak in 2001 (Olsen et al., 2005); and *L. monocytogenes* strain H7858 was a serotype 4b frankfurter isolate from the multistate outbreak of 1998-1999 in the USA (MMWR, 1999) .The genomes of these strains were compared to each other and to the previously sequenced genomes (EGD-e and *L. innocua*), and the key findings are summarized below

Mobile genetic elements. Serotype 4b strains lacked intact insertion sequences (IS elements) although both strains were found to harbor in homologous locations four copies of transposases of the IS3 family Nelson et al., 2004). Likewise serotype 1/2a strains (F6854 and EGD-e) contained three copies of the same transposase ORFA in the same location. In addition to these transposases, serotype 1/2 a strains harbored an intact IS element, namely ISLmo1, flanking internalin-like genes and likely to interfere with their expression as an outward facing promoter flanking the internalin-like proteins could produce antisense RNA. Polymorphisms in the five strains were also compared in a class of short, extragenic DNA repeats called the clustered regularly interspaced palindromic repeats (CRISPRs). CRISPR repeats, if targeted in a genotypic method, may facilitate identification of strains (Schouls et al., 2003). CRISPRs are thought to be involved in horizontal transfer of foreign DNA (Godde and Bickerton, 2006).

Bacteriophages and conjugative plasmids are involved in prokaryotic population control, evolution, and pathogenicity and their impact on the genome is substantial. In species that are not naturally transformable it is probable that phage-mediated transduction is the most important mechanism of localized genetic exchange. In the case of *Listeria,* bacteriophages seem to be important players in the evolution of this organism even though relatively few genomic loci containing phage-related sequences were identified. Comparison of the genomes revealed nine putative prophages and five defective or satellite prophage regions. Interestingly, six of the nine putative prophages had at least 11 ORFs that were homologous to ORFs of ϕA118, a temperate phage whose genome has been sequenced (Loessner et al., 2000). Surprisingly, the isolate (F2365) from the Jalisco cheese outbreak of 1985 in California did not contain an intact prophage in its genome (Nelson et al., 2004). A summary of the key genomic findings of these five strains is presented in Table 1.

Serotype- and strain-specific genes were determined for all strains using a comparative genomic analysis. Strain-specific genes could partially account for the survival, fitness, and pathogenicity attributes of these strains and serotypes. A total of 51, 97, 69 and 61 genes were identified to be unique to strains F2365, F6854, H7858 and EGD-e, respectively. A putative type II restriction-modification system responsible for cytosine methylation at GATC sites was present in the genome of the F2365 strain only. The chromosome of epidemic serotype 4b strains were previously shown to be methylated at GATC sites and resistance to Sau3AI digestion which fails to cut in the presence of cytosine methylation, has been used to identify epidemic-associated serotype 4b strains (Zhang and Kathariou, 1997; Herd and Kocks, 2001; Yildirim et al., 2004).

Earlier, the genome sequence analysis of *L. monocytogenes* EGD-e and subtractive hybridization data had revealed that surface associated proteins, particularly internalin family proteins, and sugar transport proteins, together with hypothetical proteins, comprised the majority of the strain specific genes (Herd and Kocks, 2001; Buchrieser et al., 2003). Likewise, Nelson et al. (2004) reported that putative phage-related genes occupied a large category among serotype 4b strain specific genes, besides unknown and hypothetical proteins.

Role of mobile elements in resistance to antimicrobial agents. In addition to the mobile elements that were found to be present in *Listeria* genomes, an approximately 80 Kb plasmid (pLM80) was identified in one of the serotype 4b strains (H7858) and was completely sequenced (Nelson et al., 2004). The plasmid harbored several transposable elements that were not found on the chromosome and had a high level of gene organization and sequence similarity to the *L. innocua* CLIP 11262 plasmid pLM100 (Glaser et al, 2001) and to the *Bacillus anthracis* plasmid pXO2 (Read et al., 2002). Comparison of the three plasmids revealed that pLM80 had unique regions decorated by genes for heavy metal resistance such as cadmium and arsenic. Interestingly, five of the six mobile genetic elements found in this region had >80% similarity to the genes of pLM100. The plasmid pLM80 had a region which displayed high similarity to pXO2, including putative plasmid transfer proteins (e.g. TraD/TraG). Therefore, it is likely that the plasmid is conjugative and can increase the frequency of heavy metal resistant strains in *L. monocytogenes* populations. Hence the survival and colonization of the bacterium in food environments or in hosts may be partly facilitated by these mobile elements. In fact, the acquisition of multiple antibiotic resistance in *L. monocytogenes* was shown to be mediated by self-transmissible plasmids, such as pIP811 (Poyart-Salmeron et al., 1990), pUBX1, and pWDB100 (Hadorn et al., 1993). Furthermore, resistance to heavy metals such as cadmium, which is toxic to cells, may provide crucial advantages for survival in food production environments. The resistance seems to be primarily plasmid-borne, since only a few strains were shown to be resistant and plasmid-free (Lebrun et al., 1992). Margolles et al. (1998) characterized plasmids from *L. monocytogenes* and *L. innocua* isolated from short-ripened cheeses. Different strains were found to harbor plasmids with variable sizes and numbers. The plasmids shared homology in regions of at least 20 kb. In addition, presence of plasmids in some strains were reported to correlate well with the sanitizer resistance phenotype when the strains were grown in the presence of quaternary ammonium compounds (QAC), widely used in the food industry (Romanova et al., 2002). However, the gene (*mdrl*) encoding an efflux pump that confers resistance to QACs was present both on the plasmids and chromosome of these strains (Romanova et al., 2002). On the other hand, cadmium resistance was attributed to presence of plasmid DNA, especially in serotype 1/2 strains and strains isolated from food environments (Lebrun et al., 1992; McLauchlin et al., 1997).

The first described natural transposon of *L. monocytogenes* (Tn*5422*), belonging to Tn3 family was found to be associated with *cadAC* genes conferring cadmium resistance (Lebrun et al., 1994). Interestingly, the transposon was identified on a 20.5 kb plasmid (pLm74) harboring Tn*5422* associated *cadAC*. The cadmium resistance phenotype was shown to be transferable at high frequency among *Listeria* species as well as between *L. monocytogenes* and *Staphylococcus aureus,* with similar transfer frequencies (Lemaitre et al., 1998). Not surprisingly, the Tn*5422*-associated cadmium resistance genes (*cadAC*) found on a *L. monocytogenes* plasmid (pLm74) hybridized with the *cadAC* genes of plasmid pI258 of *Staphylococcus aureus* (Lebrun et al., 1994). Thus, horizontal transfers of the genes conferring resistance are likely to be mediated by transposons and/or plasmids.

Restriction-Modification (RM) Systems. Restriction–modification systems are considered to represent bacterial mechanisms of defense against invasion by foreign (e.g. virus) DNA. They consist of a restriction endonuclease that recognizes a specific DNA sequence and hydrolyzes phosphodiester bonds making up the backbone of DNA, and a cognate modification enzyme (methyltransferase) that can methylate the same sequence to prevent the restriction of the organism's own DNA. Such systems are frequently found in bacterial genomes. There are three main types (I, II, and III) of restriction enzymes, each of which have distinct characteristics in terms of cofactor requirement, subunit composition, and mode of action (Wilson and Murray, 1991).

As mentioned, the primary biological function of restriction-modification genes is considered to be to protect the host genome against invasion of phage DNA (Bickle et al., 1993). On the other hand, the genetic barrier provided by an RM system against phages is limited, since the barriers cannot protect the genome completely (Matic et al., 1996). Thus, the traditional view, which considers RM systems as a barrier against foreign DNA invasion, has recently evolved to include new biological functions. RM systems can modulate frequency of gene exchange by limiting acquisition of foreign DNA. Curiously, they can also facilitate integration of a foreign DNA into the chromosome by creating smaller pieces of the incoming DNA, which in turn can be inserted in different loci in the genome (McKane & Milkman, 1995; Milkman et al., 1999; Kobayashi, 1998; Arber, 2005). Additionally, they may be instrumental to the spread of some genes by separating them from linked alleles (Arber, 2000; Milkman, R., 1999). Thus, RM systems may play a role in emergence of novel genetic variants of bacteria and contribute to the evolution of the resident genome.

RM genes are often found linked to mobile or mobility-related genetic elements such as plasmids, prophage in the chromosome, transposons, conjugative transposons, and integrons (Brassard et al, 1995; Kita et al., 1999; Rochepeau et al., 1997; Twomey et al., 1998; Burrus et al., 2001; Rowe-Magnus et al. 2001). RM systems have undergone extensive horizontal transfer (Xu et al., 2000; Sekizaki et al., 2001; Mruk and Kaczorowski, 2003, Lin et al., 2004) to the extent that close homologues occur in distantly related organisms as manifested by sequence and phylogenetic analysis. Indeed, the codon usage and the GC content of RM genes are often different from that of the host genome, and they are known to appear in distinct genomic context in different strains of the same organism. (Kobayashi, 2004; Jeltsch and Pingoud, 1996; Nobusato et al., 2000a, 2000b; Chinen et al., 2000, Bujnicki, 2001).

Methylation of GATC sites and bacterial pathogenesis. Methylation in bacterial genomes usually occurs at 5-methylcytosine (m5C), N4-methylcytosine, and N6-methyladenine (m6A). Base modification in bacterial genomes is performed by two classes of DNA methyltransferases: methylases functioning as part of RM systems and the so-called solitary methylases that lack cognate endonucleases, such as Dam, Dcm, and CcrM (Wion and Casadesus, 2006). Dam transfers a methyl group from S-adenosyl-methionine to the amino group of the adenine in 5′-GATC-3 (Jeltsch, 2002). DNA methylation at GATC sites has been shown to have enormous impact on nucleoid stability, replication, the cell cycle, mismatch repair, and gene expression (Jeltsch and Gumport, 2004). The first solitary DNA adenine methyltransferase that was found to play a biological role in the above functions was the Dam enzyme of *Escherichia coli* (Marinus, 1996). Recently, Dam methylase and its homologs were characterized among bacteria in the gamma and alpha subdivision of *Proteobacteria*, such as *Salmonella typhimurium* (Heithoff et al., 1999), *Yersinia pseudotuberculosis* and *Vibrio cholerae* (Julio et al., 2001), *Escherichia coli* (Oshima et al., 2002), and *Caulobacter crescentus* (Stephens et al., 1996; Wright et al., 1997). Deletion of the *dam* gene produces a variety of phenotypes in several bacteria, ranging from cell death (Julio et al., 2001) to attenuation of virulence (Watson et al., 2004), indicating multiple functions for GATC methylation in modulating gene expression, DNA mismatch repair, initiation of chromosome replication, and nucleoid stabilization.

Computational analysis of the genome of *Escherichia coli* (Henaut et al., 1996) suggested the presence of a 'GATC network', which may play a role in expression of the genes. Microarray and protein profiling analysis of wild type and *dam* mutant of *E. coli* indeed revealed that methylation of GATC sites by the Dam enzyme contributes to expression of a set of genes controlling metabolic and respiratory pathways. However, whether the clusters of GATC sites within the coding regions rather than the clusters in the binding sites for transcriptional proteins are most likely to play a role in expression is not clear (Oshima et al., 2002; Riva et al., 2003).

The effect of methylation on gene regulation in bacterial pathogenesis is well established (Julio et al., 2001). However, not all methyltransferases play a role in transcriptional regulation. For instance, putative S-adenosyl-methionine dependant methyltransferase (lmo0581) in *L. monocytogenes* EGD-e (serotype 1/2a) genome did not result in significant changes in this strain under suboptimal environmental conditions and during murine infection (Rea et al., 2004). Similarly, methyltransferases functioning as part of RM systems were not reported to directly involve in gene regulation. Convsersely, one recent report showed that methylases associated with type III RM system display sequence futures that are consistent with phase-variable expression (a reversible, high-frequency of on/off switching of expression of usually surface proteins with antigenic character) (Waude and Baumler, 2004). In fact, the comparison of global gene expression in the genome of *Haemophilus influenzae* in wild type and a mutant lacking methyltransferases associated with type III RM system identified a number of genes that were either up- or down-regulated including surface, heat-schock, and transport proteins (Srikhanta et al., 2005). Also, a recent example showed for the first time that a methylase gene that is part of a type II RM system, LlaKR2I, recognizing GATC sites (Twomey et al., 1998), has evolved a novel role within the cell besides its assumed role of protecting the host DNA from its cognate restriction enzyme; the experimental data showed that the methylase function independently from the RM system invoking the function of Dam and CcrM methylases. It was demonstrated that the abortive infection (Abi) mechanism is toxic in *Lactococus lactis* without the presence of the LlaKR2I methylase (Yang et al., 2006).

REFERENCES

1. Arber W., 2005. Gene products with evolutionary functions. Proteomics. 5:2280-4

2. Arber, W. 2000. Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol. Rev. 24:1–7

3. Bibb, W. F., B. Schwartz, B. G. Gellin, B. D. Plikaytis, and R. E. Weaver. Analysis of *Listeria monocytogenes* by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. Int. J. Food Microbiol. 8:233–239

4. Bickle, T. A., and D. H. Kruger. 1993. Biology of DNA restriction. Microbiol. Rev. 57:434–450

5. Boerlin P and Piffaretti JC., 1991. Typing of human, animal, food, and environmental isolates of *Listeria monocytogenes* by multilocus enzyme electrophoresis. Appl Environ Microbiol. 57:1624-9

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

7. Borucki MK, Kim SH, Call DR, Smole SC, Pagotto F., 2004. Selective discrimination of *Listeria monocytogenes* epidemic strains by a mixed-genome DNA microarray compared to discrimination by pulsed-field gel electrophoresis, ribotyping, and multilocus sequence typing. J Clin Microbiol. 42:5270-6

8. Borucki MK, Krug MJ, Muraoka WT, Call DR. 2003. Discrimination among *Listeria monocytogenes* isolates using a mixed genome DNA microarray. Vet Microbiol. 92:351-62

9. Brassard S, Paquet H, Roy PH. 1995. A transposon-like sequence adjacent to the AccI restriction-modification operon. Gene. 157:69-72

10. Brosch R, Buchrieser C, Rocourt J., 1991. Subtyping of *Listeria monocytogenes* serovar 4b by use of low-frequency-cleavage restriction endonucleases and pulsed-field gel electrophoresis. Res Microbiol. 142:667-75.

11. Brosch R, Chen J, Luchansky JB., 1994. Pulsed-field fingerprinting of *Listeria*e: identification of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. Appl Environ Microbiol. 60:2584-92

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

13. Bruce JL, Hubner RJ, Cole EM, McDowell CI, Webster JA., 1995. Sets of EcoRI fragments containing ribosomal RNA sequences are conserved among different strains of *Listeria monocytogenes*. Proc Natl Acad Sci U S A. 92:5229-33

14. Buchrieser C, Rusniok C, Kunst F, Cossart P, Glaser P; *Listeria* Consortium., 2003. Comparison of the genome sequences of *Listeria monocytogenes* and

Listeria innocua: clues for evolution and pathogenicity. FEMS Immunol Med Microbiol. 35:207-13

15. Bujnicki JM. 2001. Understanding the evolution of restriction-modification systems: clues from sequence and structure comparisons. Acta Biochim Pol. 48:935

16. Burrus V, Bontemps C, Decaris B, Guedon G. 2001. Characterization of a novel type II restriction-modification system, Sth368I, encoded by the integrative element ICESt1 of *Streptococcus thermophilus* CNRZ368. Appl Environ Microbiol. 67:1522

17. Cai, S., D. Y. Kabuki, Y. Kuaye, T. G. Cargioli, M. S. Chung, R. Nielsen, and M. Wiedmann. 2002. Rational design of DNA sequence-based strategies for subtyping *Listeria monocytogenes*. J. Clin. Microbiol. 40:3319

18. Call DR, Borucki MK, Besser TE., 2003. Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of *Listeria monocytogenes.* J Clin Microbiol. 41:632-9

19. Chinen A, Uchiyama I, Kobayashi I. 2000. Comparison between *Pyrococcus horikoshii* and *Pyrococcus abyssi* genome sequences reveals linkage of restriction-modification genes with large genome polymorphisms. Gene. 259:109-21

20. Davies RL, Arkinsaw S, Selander RK., 1997. Genetic relationships among Pasteurella trehalosi isolates based on multilocus enzyme electrophoresis. Microbiology. 143:2841

21. Doumith M, Cazalet C, Simoes N, Frangeul L, Jacquet C, Kunst F, Martin P, Cossart P,Glaser P, Buchrieser C., check authors 2004 New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. Infect Immun. 72:1072

22. Gendel, S. M., and J. Ulaszek. 2000. Ribotype analysis of strain distribution in *Listeria monocytogenes*. J. Food Prot. 63:179

23. Gerner-Smidt P, Hise K, Kincaid J, Hunter S, Rolando S, Hyytia-Trees E, Ribot EM, Swaminathan B; Pulsenet Taskforce. 2006. PulseNet USA: a five-year update. Foodborne Pathog Dis. 3:9

24. Glaser, P., L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero 2001. Comparative genomics of *Listeria* species. Science 294:849

25. Gnanou Besse, N., N. Audinet, A. Kerouanton, P. Colin, and M. Kalmokoff. 2005. Evolution of *Listeria* populations in food samples undergoing enrichment culturing. Int J Food Microbiol. 104:123

26. Godde JS, Bickerton A., 2006. The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. J Mol Evol. 62:718-29

27. Graves LM, Swaminathan B, Reeves MW, Hunter SB, Weaver RE, Plikaytis BD, Schuchat A. 1994. Comparison of ribotyping and multilocus enzyme electrophoresis for subtyping of *Listeria monocytogenes* isolates. J Clin Microbiol. 32:2936-43

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

29. Graves, LM, Swaminathan, B., and Susan B. Hunter, 1999. Subtyping *Listeria monocytogenes.* in *Listeria*, listeriosis, and food safety, pg. 279-99. Edited by Elliot T. Ryser, Elmer H. Marth, New York, Marcel Dekker.

30. Hadorn, K., H. Hachler, A. Schaffner, and F. H. Kayser. 1993. Genetic characterization of plasmid-encoded multiple antibiotic resistance in a strain of *Listeria monocytogenes* causing endocarditis. Eur J Clin Microbiol Infect Dis 12:928

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

32. Heithoff DM, Sinsheimer RL, Low DA, Mahan MJ., 1999. An essential role for DNA adenine methylation in bacterial virulence. Science.284:967

33. Henaut A, Rouxel T, Gleizes A, Moszer I, Danchin A., 1996. Uneven distribution of GATC motifs in the *Escherichia coli* chromosome, its plasmids and its phages. J Mol Biol. 257:574-85

34. Herd M, Kocks C., 2001. Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. Infect Immun. 69:3972-9

35. Ho, J. L., K. N. Shands, G. Friedland, P. Eckind, and D. W. Fraser. 1986. An outbreak of type 4b *Listeria monocytogenes* infection involving patients from eight Boston hospitals. Arch Intern Med 146:520-4.

36. Hunter, P.R., 1990. Reproducibility and indices of discriminatory power of microbial typing methods. J Clin Microbiol. 28:1903-5

37. Jacquet, C., B. Catimel, R. Brosch, C. Buchrieser, P. Dehaumont, V. Goulet, A. Lepoutre, P. Veit, and J. Rocourt. 1995. Investigations related to the epidemic strain involved in the French listeriosis outbreak in 1992. Appl Environ Microbiol 61:2242-6.

38. Jeffers GT, Bruce JL, McDonough PL, Scarlett J, Boor KJ, Wiedmann M., 2001. Comparative genetic characterization of *Listeria monocytogenes* isolates from human and animal listeriosis cases. Microbiology. 147:1095-104

39. Jeltsch A., 2002. Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem. 3:274-93

40. Jeltsch A., and Gumport, RI., 2004. DNA Methyltransferases, Bacterial. Encyclopedia of Biological Chemistry, Edited by William J. Lennarz, M. Daniel Lane. ISBN 0124437109. Academic Press Academic Press

41. Jeltsch, A., and A. Pingoud. 1996. Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. J. Mol. Evol. 42:91–96

42. Johnson J, Jinneman K, Stelma G, Smith BG, Lye D, Messer J, Ulaszek J, Evsen L, Gendel S, Bennett RW, Swaminathan B, Pruckler J, Steigerwalt A, Kathariou S, Yildirim S, Volokhov D, Rasooly A, Chizhikov V, Wiedmann M, Fortes E, Duvall RE, Hitchins AD., 2004. Natural atypical *Listeria* innocua strains with *Listeria monocytogenes* pathogenicity island 1 genes. Appl Environ Microbiol. 70:4256-66

43. Julio SM, Heithoff DM, Provenzano D, Klose KE, Sinsheimer RL, Low DA, Mahan MJ. 2001. DNA adenine methylase is essential for viability and plays a role in the pathogenesis of *Yersinia pseudotuberculosis* and *Vibrio cholerae*. 1: Infect Immun. 69:7610

44. Kathariou,S. (2002) year format *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J. Food Prot., 65, 1811-1829 consistent format for journal vol., pages

45. Kita, K., J. Tsuda, T. Kato, K. Okamoto, H. Yanase, and M. Tanaka. 1999. Evidence of horizontal transfer of the EcoO1091 restriction-modification gene to *Escherichia coli* chromosomal DNA. J. Bacteriol. 181:6822–6827

46. Kobayashi I., 1998. Selfishness and death: raison d'etre of restriction, recombination and mitochondria. Trends Genet. 14:368-74

47. Kobayashi, I. 2004. Restriction-Modification Systems as Minimal Forms of Life. p. 19-62. In H. J. Gross (Series ed) and A. Pingoud (ed), Restriction Endonucleases. Springer-Verlag, complete the ref

48. Lebrun M, Audurier A, Cossart P., 1994. Plasmid-borne cadmium resistance genes in *Listeria monocytogenes* are similar to cadA and cadC of *Staphylococcus aureus* and are induced by cadmium. J. Bacteriol. 176:3040-8

49. Lebrun M, Loulergue J, Chaslus-Dancla E, Audurier A., 1992. Plasmids in *Listeria monocytogenes* in relation to cadmium resistance. Appl Environ Microbiol. 58:3183-6.

50. Lemaitre JP, Echchannaoui H, Michaut G, Divies C, Rousset A., 1998. Plasmid-mediated resistance to antimicrobial agents among *Listeriae.* J Food Prot. 61:1459-64

51. Lin TL, Shun CT, Chang KC, Wang JT. 2004. Isolation and characterization of a HpyC1I restriction-modification system in *Helicobacter pylori*. J Biol Chem. 279:11156-62

52. Linnan, M. J., L. Mascola, X. D. Lou, V. Goulet, S. May, C. Salminen, D. W. Hird, M. L. Yonekura, P. Hayes, R. Weaver, and et al. 1988. Epidemic listeriosis associated with Mexican-style cheese. N Engl J Med 319:823-8.

53. Loessner MJ, Inman RB, 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

54. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG., 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A. 95:3140-5

55. Marinus, M. G. Methylation of DNA. In *Escherichia coli* and *Salmonella:* Cellular and Molecular Biology (eds Neidhardt, F. C. et al.) 782–791. ASM Press, Washington DC, 1996

56. Matic I, Taddei F, Radman M. 1996. Genetic barriers among bacteria. Trends Microbiol. 4:69-72

57. McKane M, Milkman R. 1995. Transduction, restriction and recombination patterns in *Escherichia coli*. Genetics. 139:35-43

58. McLauchlin J, Audurier A, Frommelt A, Gerner-Smidt P, Jacquet C, Loessner MJ, van der Mee-Marquet N, Rocourt J, Shah S, Wilhelms D. 1996. WHO study on subtyping *Listeria monocytogenes* results of phage-typing. Int J Food Microbiol. 32:289-99

59. Meinersmann RJ, Phillips RW, Wiedmann M, Berrang ME., 2004. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes reveals clonal and recombination histories of three lineages. Appl Environ Microbiol. 70:2193-203

60. Milkman R, Raleigh EA, McKane M, Cryderman D, Bilodeau P, McWeeny K. 1999. Molecular evolution of the *Escherichia coli* chromosome. V. Recombination patterns among strains of diverse origin. Genetics. 153:539-54

61. Milkman, R. 1999. Gene transfer in *E. coli* p. 291-309. In R. L. Charlebois (ed.), Organization of the prokaryotic genome. ASM Press, Washington, D.C.

62. MMWR Weekly, Multistate outbreak of listeriosis in the United States, 1998. MMWR Weekly, 47, 1085-1086

63. Moller K, Nielsen R, Andersen LV, Kilian M., 1992. Clonal analysis of the Actinobacillus pleuropneumoniae population in a geographically restricted area by multilocus enzyme electrophoresis. J Clin Microbiol. 30:623-7

64. Mruk I, Kaczorowski T. 2003. Genetic organization and molecular analysis of the EcoVIII restriction-modification system of *Escherichia coli* E1585-68 and its comparison with isospecific homologs. Appl Environ Microbiol. 69:2638-50

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

66. Nadon, C.A, Woodward, D.I., Young,C., Rodgers, F.G., Wiedmann, M. 2001. Correlations between molecular subtyping and serotyping of *Listeria monocytogenes.* J. Clin. Micro., 39:2704

67. Nelson KE, Fouts DE, Mongodin EF, et al., 2004. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. Nucleic Acids Res. 32:2386-9

68. Nightingale KK, Windham K, Wiedmann M., 2005. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. J Bacteriol. Aug; 187 (16):5537-51.

69. Nobusato A, Uchiyama I, Kobayashi I. 2000a. Diversity of restriction-modification gene homologues in *Helicobacter pylori*. Gene. 259:89-98.

70. Nobusato A, Uchiyama I, Ohashi S, Kobayashi I. 2000b. Insertion with long target duplication: a mechanism for gene mobility suggested from comparison of two related bacterial genomes. Gene. 259:99-108

71. Norrung B, Skovgaard N., 1993. Application of multilocus enzyme electrophoresis in studies of the epidemiology of *Listeria monocytogenes* in Denmark. Appl Environ Microbiol. 59:2817-22

72. Olsen SJ, Patrick M, Hunter SB, Reddy V, Kornstein L, MacKenzie WR, Lane K, Bidol S, Stoltman GA, Frye DM, Lee I, Hurd S, Jones TF, LaPorte TN, Dewitt W, Graves L, Wiedmann M, Schoonmaker-Bopp DJ, Huang AJ, Vincent C, Bugenhagen A, Corby J, Carloni ER, Holcomb ME, Woron RF, Zansky SM, Dowdle G, Smith F, Ahrabi-Fard S, Ong AR, Tucker N, Hynes NA, Mead P. Multistate outbreak of *Listeria monocytogenes* infection linked to delicatessen turkey meat. Clin Infect Dis. 2005 Apr 1;40:962-7

73. Oshima T, Wada C, Kawagoe Y, Ara T, Maeda M, Masuda Y, Hiraga S, Mori H: 2002. Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*. Mol Microbiol 45:673-695

74. Piffaretti JC, Kressebuch H, Aeschbacher M, Bille J, Bannerman E, Musser JM, Selander RK, Rocourt J., 1989. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. Proc Natl Acad Sci; format 86:3818-22

75. Pine L, Malcolm GB, Brooks JB, Daneshvar MI., 1989. Physiological studies on the growth and utilization of sugars by *Listeria* species. Can J Microbiol. 35:245-54

76. Pinner, R. W., A. Schuchat, B. Swaminathan, P. S. Hayes, K. A. Deaver, R. E. Weaver, B. D. Plikaytis, M. Reeves, C. V. Broome, and J. D. Wenger. 1992. Role of foods in sporadic listeriosis. II. Microbiologic and epidemiologic investigation. JAMA 267:2046–2050.

77. Pohl MA, Wiedmann M, Nightingale KK., 2006. Associations among *Listeria monocytogenes* genotypes and distinct clinical manifestations of listeriosis in cattle. Am J Vet Res. 67:616-26.

78. Rasmussen, O. F., P. Skouboe, L. Dons, L. Rossen, and J. E. Olsen. 1995. *Listeria monocytogenes* exists in at least three evolutionary lines: evidence from flagellin, invasive associated protein, and listeriolysin O genes. Microbiology 141:2053–2061

79. Rasmussen, O. F., T. Beck, J. E. Olsen, L. Dons, and L. Rossen. 1991. *Listeria monocytogenes* isolates can be classified into two major types according to the sequence of the listeriolysin gene. Infect. Immun. 59:3945–3951

80. Rea RB, Gahan CG, Hill C., 2004. Disruption of putative regulatory loci in *Listeria monocytogenes* demonstrates a significant role for Fur and PerR in virulence. Infect Immun. 72:717

81. Read,T.D., Salzberg,S.L., Pop,M., Shumway,M., Umayam,L., Jiang,L., Holtzapple,E., Busch,J.D., Smith,K.L., Schupp,J.M. et al. (2002) Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis.* Science, 296: 2028-2033

82. Revazishvili T, Kotetishvili M, Stine OC, Kreger AS, Morris JG Jr, Sulakvelidze A., 2004. Comparative analysis of multilocus sequence typing and pulsed-field gel electrophoresis for characterizing *Listeria monocytogenes* strains isolated from environmental and clinical sources. J Clin Microbiol. 42:276-85

83. Roberts A, Nightingale K, Jeffers G, Fortes E, Kongo JM, Wiedmann M.2006. Genetic and phenotypic characterization of *Listeria monocytogenes* lineage III. Microbiology. 152:685-93

84. Rochepeau, P., L. B. Selinger, and M. F. Hynes. 1997. Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum* VF39SM. Mol. Gen. Genet. 256:387–396

85. Rocourt, J., Jacquet, Ch., Bille, J., 1997. Human listeriosis, 1991/1992. WHO/FNU/FOS/ 97.1. World Health Organization, Geneva, Switzerland.

86. Rowe-Magnus DA, Guerout AM, Ploncard P, Dychinco B, Davies J, Mazel D. 2001. The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. Proc Natl Acad Sci U S A. 98:652-7

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

88. Schlech, W. F. III, P. M. Lavigne, R. A. Bortolussi, A. C. Allen, E.V. Haldane, A. J. Wort, A. W. Hightower, S. E. Johnson, S. H. King, E. S. Nicholls, and C. V. Broome. 1983. Epidemic listeriosis—evidence for transmission by food. N. Engl. J. Med. 308:203–206.

89. Schmid MW, Ng EY, Lampidis R, Emmerth M, Walcher M, Kreft J, Goebel W, Wagner M, Schleifer KH., 2005. Evolutionary history of the genus *Listeria* and its virulence genes. Syst Appl Microbiol. 28:1-18

90. Schouls LM, Reulen S, Duim B, Wagenaar JA, Willems RJ, Dingle KE, Colles FM, Van Embden JD., 2003. Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination. J Clin Microbiol. 41:15

91. Schwartz, B., D. Hexter, C. V. Broome, A. W. Hightower, R. B. Hirschhorn, J. D. Porter, P. S. Hayes, W. F. Bibb, B. Lorber, and D. G. Faris. 1989. Investigation of an outbreak of listeriosis: new hypotheses for the etiology of epidemic infections. J Infect Dis 159:680-5.

92. Seeliger, H. P. R. & Hohne, K. (1979). Serotyping of *L.monocytogenes* and related species. Methods Microbiol 13, 31–49

93. Sekizaki T, Otani Y, Osaki M, Takamatsu D, Shimoji Y. 2001. Evidence for horizontal transfer of SsuDAT1I restriction-modification genes to the *Streptococcus suis* genome. J Bacteriol. 183:500-11

94. Sleator RD, Gahan CG, Hill C., 2003. A postgenomic appraisal of osmotolerance in *Listeria monocytogenes* AEM, 69:1-9

95. Srikhanta YN, Maguire TL, Stacey KJ, Grimmond SM, Jennings MP., 2005. The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. Proc Natl Acad Sci 102:5547

96. Stephens C, Reisenauer A, Wright R, Shapiro L., 1996. A cell cycle-regulated bacterial DNA methyltransferase is essential for viability. Proc Natl Acad Sci U S A. 93:1210-4.

97. Twomey DP, McKay LL, O'Sullivan DJ., 1998. Molecular characterization of the *Lactococcus lactis* LlaKR2I restriction-modification system and effect of an IS982 element positioned between the restriction and modification genes. J Bacteriol. 180:5844

98. Twomey, D. P., L. L. McKay, and D. J. O’Sullivan. 1998. Molecular characterization of the *Lactococcus lactis* LlaKR2I restriction-modification

system and effect of an IS982 element positioned between the restriction and modification genes. J. Bacteriol. 180:5844–5854

99. van der Woude MW, Baumler AJ., 2004. Phase and antigenic variation in bacteria. Clin Microbiol Rev. 17:581

100. Vazquez-Boland JA, Kuhn M, Berche P, Chakraborty T, Dominguez-Bernal G, Goebel W, Gonzalez-Zorn B, Wehland J, Kreft J. 2001. *Listeria* Pathogenesis and Molecular Virulence Determinants. Clin Microbiol Rev., 14:584

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

102. Watson ME Jr, Jarisch J, Smith AL: Inactivation of deoxyadenosine methyltransferase (dam) attenuates *Haemophilus influenzae* virulence. Mol Microbiol 2004, 53:651-664.

103. Wiedmann M, Bruce JL, Keating C, Johnson AE, McDonough PL, Batt CA. 1994. Ribotypes and virulence gene polymorphisms suggest three distinct *Listeria monocytogenes* lineages with differences in pathogenic potential. Infect Immun. 65:2707-16

104. Wiedmann, M., J. L. Bruce, C. Keating, A. E. Johnson, P. L. McDonough, and C. A. Batt. 1997. Ribotypes and virulence gene polymorphisms suggest three distinct

Listeria monocytogenes lineages with differences in pathogenic potential. Infect Immun 65:2707-16.

105. Wilson, G.G. & Murray, N.E. (1991) Restriction and modification systems. Annu. Rev. Genet. 25, 585–627

106. Wion D, Casadesus J., 2006. N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. Nat Rev Microbiol. 4:183-92

107. Wright R, Stephens C, Shapiro L., 1997. The CcrM DNA methyltransferase is widespread in the alpha subdivision of proteobacteria, and its essential functions are conserved in *Rhizobium meliloti* and *Caulobacter crescentus*. J Bacteriol.179:5869-77.

108. Xu Q, Stickel S, Roberts RJ, Blaser MJ, Morgan RD. 2000. Purification of the novel endonuclease, Hpy188I, and cloning of its restriction-modification genes reveal evidence of its horizontal transfer to the *Helicobacter pylori* genome. J Biol Chem. 275:17086-93

109. Yang JM, Deurraza PJ, Matvienko N, O'Sullivan DJ., 2006. Involvement of the LlaKR2I methylase in expression of the AbiR bacteriophage defense system in *Lactococcus lactis* subsp. lactis biovar diacetylactis KR2. 1: J Bacteriol. 188:1920

110. Yildirim, S., W. Lin, A. D. Hitchins, L. A. Jaykus, E. Altermann, T. R. Klaenhammer, and S. Kathariou. 2004. Epidemic clone I-specific genetic markers

in strains of *Listeria monocytogenes* serotype 4b from foods. Appl Environ Microbiol 70:4158-64.

111. Zhang C, Nietfeldt J, Zhang M, Benson AK., 2005. Functional consequences of genome evolution in *Listeria monocytogenes*: the lmo0423 and lmo0422 genes encode sigmaC and LstR, a lineage II-specific heat shock system. J Bacteriol. 187:7243-53

112. Zhang C, Zhang M, Ju J, Nietfeldt J, Wise J, Terry PM, Olson M, Kachman SD, Wiedmann M, Samadpour M, Benson AK. 2003. Genome diversification in phylogenetic lineages I and II of *Listeria monocytogenes*: identification of segments unique to lineage II populations. J Bacteriol. 185:5573-84

113. Zheng W, Kathariou S., 1997. Host-mediated modification of Sau3AI restriction in *Listeria monocytogenes*: prevalence in epidemic-associated strains. Appl Environ Microbiol. 63:3085-9

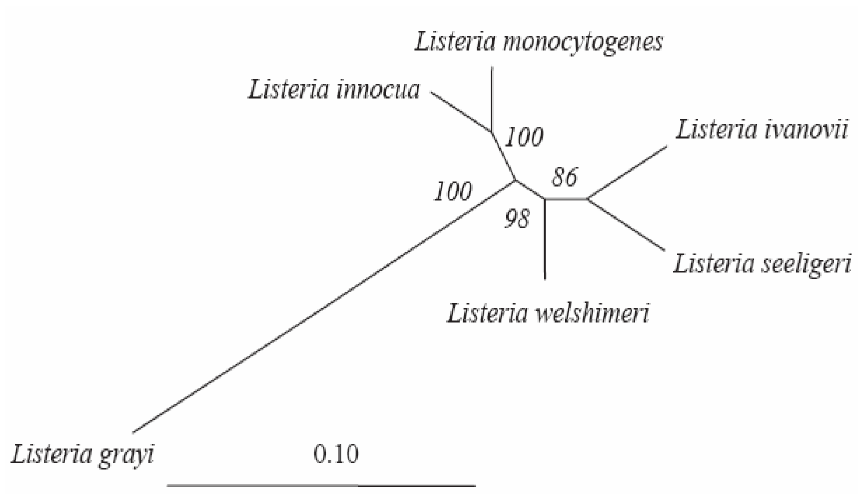


Fig.1. The composite phylogenetic tree of the genus *Listeria* based on concatenated sequences of 16S and 23S rDNA, *iap, prs, vclB,* and, ldh (**Reproduced with the permission from** Schimd et al, 2005)

Table 1. A summary of comparative genomic analyisis of *L. monocytogenes*

Genome Future	F2365* (4b)	H7858* (4b)	EGD-e* (1/2a)	F6854* (1/2a)	CLIP** 11262
Chromosome					
Size (Mb)	2.90	2.89	2.94	2.95	3.01
No. of protein coding genes	2847	3024	2853	2973	2973
(G+C) %	38	38	39	37.8	37
Phage regions	2	2	1	3	5
Monocins	1	1	1	1	1
No. of strain specific genes	51	69	61	97	149
Plasmid	None	1	None	None	1
Size (kb)	-	82.2	-	-	81.9
No. of ORFs	-	94	-	-	79

**L. monocytogenes, **L. innocua*
Adopted from Nelson et al., 2004

CHAPTER II

LISTERIA MONOCYTOGENES EPIDEMIC CLONES AND UNIQUE GENOMIC REGIONS

INTRODUCTION

Listeria monocytogenes is an opportunistic pathogen that causes rare but frequently fatal (ca. 30%) infections termed listeriosis in both human and animals. Human listeriosis occurs as sporadic disease or in the form of outbreaks (epidemics), and can affect especially the elderly, neonates, pregnant women or immunocompromised individuals. While rarely diagnosed prior to 1960, more than 10,000 cases of listeriosis were recorded in the medical literature between 1960 and 1982 (Rocourt and Reilly, 2000). Epidemiologic and laboratory investigations in 1983 following an outbreak in Nova Scotia showed for the first time that food served as a vehicle for the transmission of this pathogen. The ubiquitous distribution of this bacterium in the environment, its ability to replicate in the cold, and its pathogenic potential make it especially problematic for the safety of refrigerated and ready to eat (RTE) foods (Kathariou, 2002). Most commonly implicated in listeriosis are highly processed, ready-to-eat (RTE) foods that are kept refrigerated for extended periods of time, especially deli and other processed meats, soft cheeses, and processed seafood.

Molecular and other typing methods have consistently classified the species of *L. monocytogenes* into distinct phylogenetic lineages separating serotypes 4b, 3b and 1/2b from 1/2a, 3a, 3c and 1/2c (Pifaretti et al., 1989; Bibb et al., 1989, 1990; Brosch et al., 1994; Graves et al., 1994; Rasmussen et al., 1995; Wiedmann et al., 1997; Nadon et al., 2001). However, not all strains of the bacterium are equally likely to be recovered from cases of human listeriosis (Hof and Rocourt, 1992). In fact, only three serotypes (4b, 1/2b, and 1/2a) are responsible for more than 95% of the cases of illness (Rocourt et al., 1997). The overall incidence of serotype 4b in sporadic as well as in epidemic listeriosis ranges from 50 to 70%. Thus, the presence of bacteria of this serotype in RTE foods appear to increase the food's risk of implication in listeriosis (Kathariou, 2002) and these strains need to be accurately characterized for rational design of nucleotide based detection strategies and epidemiological studies.

Epidemics of food borne listeriosis have been characterized by repeated involvement of a small group of strain clusters consisting of genetically closely related strains. One such cluster of strains of serotype 4b have been responsible for numerous temporally and spatially distinct outbreaks in North America and Europe and represent a "clone" from the epidemiological standpoint, named epidemic clone-I (EC-I) (Kathariou, 2003).

Numerous studies have addressed several relevant molecular and microbiological questions related to *L. monocytogenes.* In this review, emphasis is placed on genetic markers with which derivatives of epidemic clones can be identified. We will begin with a brief overview of population genetics postulates with respect to clonality of bacterial populations and the population structure of *L. monocytogenes.* We will then present the identification and characterization of genomic regions specific to ECI, which can serve as genetic markers in detection of these isolates.

Clonality in bacterial populations and epidemic clone concept. The term 'clone' has been frequently used in molecular epidemiology, especially, after a widely accepted definition of the concept was introduced by Ǿrskov and Ǿrskov (1983): "…the word clone will be used to denote bacterial cultures isolated independently, from different sources in different locations, and perhaps at different times, but showing so many identical phenotypic and genotypic traits that the most likely explanation for this identity is a common origin" (Ǿrskov and Ǿrskov, 1983). From the evolutionary point of view, the clone is monophyletic, i.e., all strains are descendants of the same ancestor, although clones may vary in their degree of genetic homogeneity due to recombination (Dijkshoorn et al., 2000; Gupta and Maiden, 2001). Hence the clonality of a bacterial population is the result of a complex interplay between *de novo* mutations (vertical transmission) and recombination occurring during lateral transmission of foreign DNA sequences.

The relative contribution of recombination generating new bacterial genotypes inversely affects the clonality of a bacterial population because recombination allows different regions to have different evolutionary histories by breaking up linkages between loci (Awadalla, 2003). Furthermore, frequent recombination brings about instability in the genome it is impacting, and thereby the strains of interest will not readily lend themselves to molecular epidemiological analysis. In this respect, it is essential that the degree to which recombination affects a pathogen population be accurately measured. In the scarcity of lateral genetic exchange, however, the presence of a given mutation will correlate with other mutations that have accumulated in the chromosome during the history of the lineage. Accordingly, a non-random distribution of genetic polymorphisms (linkage disequilibrium) will be observed within a clonal population (Spratt and Maiden, 1999). In such a population, the genotypes at different loci are inherited by the daughter cells of the same progenitor as blocks of genes (haplotype) without reshuffling. Different parts of the genome, therefore, evolve in parallel, which can be detected using population genetics methods, e.g. a statistical correlation between genetic distance estimates obtained from different markers (Tibayrenc, 2005). In addition to analysis of linkage disequilibrium, which for the most part was applied to MLEE data in the pre-genomics era (Piffaretti et al., 1989), phylogenetic studies of nucleotide sequence data (lack of congruence between the housekeeping gene trees), and studies of particular lineages (genes) as they diverge are used to determine bacterial population structure (Spratt and Maiden, 1999; Awadalla, 2003; Feil, 2004). In bacterial species with clonal population structure such as *L. monocytogenes*, the non random associations of alleles at different genes (linkage disequilibrium) may result in certain clones or families of clones that are more highly pathogenic than others, or are etiologic agents associated with certain diseases and clinical syndromes (Piffaretti et al., 1989).

In general, clonal bacterial populations comprise independent lineages (clonal complexes) whose genomes display pronounced stability over long evolutionary periods (Feil, 2004). The longevity of certain clones of pathogenic bacteria is in the magnitude of decades, if not hundreds of years or longer (Land and Reeves, 2001). Nucleotide substitutions in these populations occur mainly by *de novo* mutations, although diversity may also arise by mobile genetic element-mediated mechanisms, e.g. gain and loss of plasmids, insertion sequences, prophage, etc. Genetic diversity of these populations is affected by adaptive mutations that increase fitness in certain lineages within the population, increasing their frequency at the expense of other lineages (periodic selection) (Levin, 1981, Cohan, 2002). Similarly, clonal bacterial populations undergo clonal expansion, reductions, and/or clonal replacement due to competition and immune selection (Achtman, 1997; Zhu et al., 2001; Gupta and Maiden, 2001).

Clonality of *L. monocytogenes*. It is clearly established that the species of *L. monocytogenes* is divided by the boundaries of distinct lineages. Analysis of allelic frequencies in a given population and the phylogenetic analysis of selected genes have indicated that each division represents a distinct clonal group with limited gene flow between the strains of different lineages (Wiedmann et al., 1997; Doumith et al., 2004; Ward et al., 2004). Wiedmann and co-workers (1997) used ribotyping and PCR-RFLP to analyze polymorphisms in virulence gene alleles of the strains. The alleles sampled from strains of different lineages showed strong linkage disequilibrium, which suggested clonality of the population. In parallel to these observations, Call et al. (2003), performed “mixed-genome” microarrays to analyze the population structure of *L. monocytogenes* and identified twenty nine serotype-specific probes mapping onto the previously reported genomic framework of the species. Although they detected several sequences of mobile genetic elements which are indicative of horizontal gene transfer, recombination was not sufficiently frequent to alter the clonal structure of the population. Rather, *L. monocytogenes* exhibited strong clonal population structure along the lines of genomic division (lineages) as manifested by the linkage disequilibrium index calculated in this study. Similarly, MLST analysis based on hypervariable and housekeeping genes showed that lineage I of *L. monocytogenes* is highly clonal, while strains of the other two lineages displayed relatively greater diversity in their evolutionary history (Meinersmann et al., 2004; Nightingale et al., 2005). Although different lineages of the bacterium can share habitats, the degree of horizontal gene transfer events disrupting clonal structure in each lineage was found to be variable (Nightingale et al., 2005).

Based on the phylogenetic analysis the strong clonal structure of lineage I has been attributed to evolutionary changes that occurred relatively recently, (Ward et al., 2004). The results of the phylogenetic analysis of the *prfA* virulence gene cluster suggested that lineage I underwent a recent sweeping change (bottleneck). Furthermore, random genetic drift and horizontal gene transfer in the background of positive selection are the main evolutionary forces directing the evolution of this particular lineage (Ward et al., 2004; Nightingale et al., 2005).

Comparison of relative frequencies of the three *L. monocytogenes* lineages in isolates from food products and clinical cases by using subtyping methods such as PCR-RFLP, PFGE, ribotyping, and MLST demonstrated that lineage I isolates are overrepresented and lineage II isolates are underrepresented among clinical cases, while lineage III isolates are rare among clinical isolates and appear to be animal-adapted pathogens (Wiedmann, 1997; Jeffers et al., 2001; Gray et al., 2004; Roberts et al., 2006). Serotype 4b strains in lineage I deserve special attention due to their epidemiological importance in that most of the well-known outbreaks and a significant fraction of sporadic illness were caused by serotype 4b strains (Kathariou, 2003). Theoretical models based on DNA arrays point toward the possibility that serotype 4b populations in lineage I may have evolved from a serotype 1/2b ancestor by multiple gene transfer events (Doumith et al., 2004). In parallel observation, some genetically identical or closely related strains based on high discriminatory typing methods (Multilocus Sequence

Typing and Pulsed-Field Gel Electrophoresis) belonged to different serotypes (1/2b and 3b), suggesting the possibility of "antigen switching" during evolution of *L. monocytogenes* (Revazishvilli et al., 2004). It is striking that, in accordance with the clonal structure of its population and epidemic spread of the clones, serotype 4b strains have subpopulations whose genetic repertoire is different from other strains of the same serotype (Herd and Kocks, 2001; Evans et al., 2004; Yildirim et al., 2004).

Epidemic clones of *Listeria monocytogenes*. The majority of outbreaks of food-borne listeriosis reported in the literature have involved a small group of genetically closely related serotype 4b strains. Newly recognized clonal groups in an epidemic may shift the focus and breadth of the epidemiological study. The species of *L. monocytogenes* has four such epidemic clonal groups that have been identified thus far and are referred to as epidemic clones I, Ia, II, and III (ECI, ECIa, ECII, and ECIII) (Table I) (Pifaretti et al., 1989; Bibb et al., 1990; Kathariou, 2003).

Epidemic Clone I (ECI). Epidemic clone-I (ECI) has earned special recognition in that this clonal group is widespread geographically and has caused numerous outbreaks both in Europe and North America since the early 1980s (Rocourt et al., 1991, Kathariou, 2003). Examples of the major outbreaks are listed in Table I. Recent reports suggest that this clone or closely related strains (clonal complex) still continue to cause outbreaks in Europe (Grif et al., 2006; Garcia-Alvarez M, 2006) and in the United States as in the example of a recent outbreak that occurred in Texas due to the consumption of contaminated Mexican style cheese (CDC, 2004). Genetic subtyping studies conducted since the early 1990s have indicated that the strains from different outbreaks are in fact closely related. Pifaretti and co-workers (1989) employed MLEE to show for the first time the clonal relationship of the strains isolated during major epidemics and referred to as Electrophoretic Type 1 (ET-1). Bibb and co-workers (1991) used MLEE as applied to a larger collection of strains, including strains from major epidemics, and corroborated the clonal relationship of the strains that caused epidemics in Europe and in North America. Two years later Buchrieser and co-workers (1993) applied PFGE on 75 serotype 4b strains, and reported that 77% of these strains showed close genomic relatedness. In particular, the outbreak strains isolated during major epidemics in Switzerland (1983-1987), the United States (California, 1985) and Denmark (1985-1987) demonstrated indistinguishable DNA restriction patterns (Buchrieser et al., 1993). Jacquet and co-workers (1995) characterized strains involved in the French listeriosis outbreak of 1992 by employing several subtyping methods and found that the epidemic strain was closely related to epidemic strains responsible for major outbreaks of listeriosis previously observed in Europe and North America. Overall, these studies suggest the presence of a subpopulation with remarkable stability within the serotype 4b population. It must be pointed out, however, that the population of outbreak strains is not always genetically uniform in that up to 29% of each major epidemic outbreak consisted of strains with modification of antigenic determinants, as manifested by non-reactivity with serotype-specific monoclonal antibodies and resistance to serotype 4b phage (Clark et al., 2000).

EC-I unique genomic regions will be presented in detail later in this review. However, it is interesting to note that in spite of well-documented outbreaks and global spread of this clone only recently have genomotyping studies been reported. This is in part because genome sequence data were unavailable until Nelson and co-workers (2004) sequenced the genome of one of the ECI strains (F2365) that caused the California Jalisco cheese outbreak in 1985. In the absence of the genome sequence, however, the hallmarks of this clone, such as polymorphism at certain loci (Zhang and Kathariou, 1995; Tran and Kathariou, 2002) and methylation of cytosines at GATC sites (Zhang and Kathariou, 1997) have helped to identify strains that belong to this lineage. Furthermore, Herd and Kocks (2001) subtracted a virulent prototype epidemic strain (F2365) from virulent *L. monocytogenes* strain EGD-e and demonstrated that certain genomic fragments were indeed unique to F2365 and closely related strains compared to other serotype 4b strains. Yildirim and co-workers (2004) demonstrated that the genetic markers identified in the genome of outbreak strains were also found in EC-I like strains isolated from food with no known epidemiological links to outbreaks. Finally, the genome sequence and comparative genomics work of Nelson et al. (2004) identified 51 genes that were specific to this epidemic clone. In terms of role category the majority of the clone-specific genes encoded surface proteins and sugar transport proteins in addition to proteins whose functions are yet unknown.

Epidemic Clone Ia (ECIa). This clone was first isolated from pasteurized milk in an outbreak that occurred in Massachusetts in 1983. The significance of this clone is that it consists of serotype 4b strains which are genetically related to EC-I. However, it has distinct features such as lack of methylation of cytosines at GATC sites lack of polymorphism at certain loci that are characteristic of EC-I (Zhang and Kathariou, 1995; Tran et al., 2002) and distinct Multilocus Enzyme Electrophoresis (MEE) and ribotypes.

Epidemic Clone II (ECII). Epidemic clone-II strains were implicated in a nationwide outbreak in 1998-1999 that involved contaminated hotdogs and resulted in 108 cases, 14 deaths, and four stillbirths (MMWR, 1999; Stone and Shoenberger, 2001; Mead et al, 2005). This clone is relatively well-characterized among other epidemic clones because its genome was partially sequenced and strain-specific genes (61 in total) were identified (Nelson et al., 2004). Also, several genotypic studies using highly discriminatory typing methods have been conducted (Mead et al., 2005, Graves et al., 2005). These studies revealed that the outbreak strain was a different strain type of serotype 4b with unprecedented genetic fingerprint, thereby representing a novel epidemic-associated lineage (Graves et al., 2005). Evans and co-workers (2004) employed comparative genomic analysis and nucleotide hybridizations to identify genetic markers for this clonal group. Mazotta and Gombas (2001) studied heat resistance of this clone in hot dog batter to compare its physiological potential with other strains of the organism, but they did not observe any differences.

Apart from 61 strain-specific genes which for the most part include hypothetical genes as well as phage remnants, the genome of this strain had extensive single nucleotide polymorphisms scattered around the genome. Also, it harbored an 80-kb plasmid decorated with genes that may confer to the bacterium adaptive ability in certain environments, such as heavy metal resistance (Nelson et al., 2004). It is intriguing to note that the outbreak strain has substantial divergence/deletions compared to the F2365 genome (e.g. LMOf2365_0448 to LMOf2365_0471) in certain loci, which otherwise were conserved in the genome of serotype 4b strains (Evans et al., 2004). Several genes in this region were not detected in the genome of ECII strains and only three ORFs were conserved between this clone and other serotype 4b strains (Evans et al., 2004). The fact that the region was adjacent to a well-known virulence gene of the bacterium *inlAB* and the G+C% content of this region (19-35%) was lower than the genome average (38%) evoked the possibility that the corresponding genes may have been horizontally transferred from an unknown source (Evans et al., 2004).

Epidemic Clone III (ECIII). The strain (F6854) is serotype 1/2a, which represents genomic division I, and its genome was partially sequenced by the Institute of Genomic Research with funding through the USDA (Nelson et al., 2004). The strain contaminated turkey franks in 1988 causing listeriosis in a cancer patient in Oklahoma (MMWR Weekly, 1989). Remarkably, a derivative of this strain was responsible for a multistate outbreak in 2000 involving turkey deli meats (MMWR, 2000). The outbreak strain originated from the same food processing plant and shared indistinguishable pulsotypes with the strain that caused the sporadic case in 1988 (Olsen et al., 2005). Thus, this strain also represents a 'persistent strain' of the organism, which can persist in food processing plants and repeatedly contaminate food products (Autio et al., 2002; Kathariou, 2003). The genomic and bacteriological features of this strain are not well understood yet. The genome sequence of this strain facilitated identification of strain specific genes, which included transport (PTS systems) and cell-wall associated proteins (internalin family), and genes related to mobile genetic elements. Unknown and hypothetical proteins made up most of the strain specific genes.

Epidemiology and ecology of epidemic strains. The historical account of listeriosis in the literature dates as far back as 1924 when Murray and co-workers isolated the organism from rabbits. However, the first recorded listeriosis epidemic was an outbreak affecting newborns in Halle, Germany in 1949 (Rocourt, 1999). The outbreak is thought to be an epidemic “landmark” in that the disease prompted public awareness and instigation of intense epidemiological research on the disease (Hof, 2003). In fact, a plethora of scientific records on listeriosis can be found, including documentation of nine listeriosis outbreaks in the 20-year period from 1960 to 1980. Nevertheless, a clear-cut link between the etiological agent of the disease and foods could not be established until the coleslaw outbreak in Nova Scotia, Canada in 1981 (Rocourt, 1991; Ryser and Marth, 1991; Kathariou, 2003). Examples of important outbreaks caused by serotype 4b strains are listed in the Table 2. Even though strains of serotype 4b have been responsible for most of the outbreaks that have occurred in Europe and North America (McLauchlin, 1990; Kathariou, 2002) epidemic clones are not commonly isolated from routine surveys of foods (Boerlin et al., 1991; 1997; Gendel and Ulaszek, 2000; Gilbreth et al., 2005).

The ecology of epidemic clones of *L. monocytogenes* and their reservoirs in nature or in the processing plant environment remain poorly understood (Kathariou, 2003; Eifert et al., 2005). This paradoxical observation could be ascribed to the paucity of epidemiological surveys and prevalence studies targeting clonal complexes that are closely related to epidemic strains. Clearly, the periodic involvement of the same epidemic clone in temporally and spatially unrelated outbreaks suggests that the implicated strains have reservoirs serving as a conduit to the spread of the epidemic strains. Thus, it is imperative that such reservoirs and ecological niches be identified to understand the evolution of epidemic lineages and to protect the health of consumers.

Even though data with respect to the ecology of epidemic strains are limited, several hypotheses based on the available data have pointed toward sewage water, animal farms, food processing plants, and retail environments as the primary sources of the clonal groups. Boerlin and Pifaretti (1991) employed MLEE to characterize ECI strains that caused the Swiss outbreak between 1983-1987 and hypothesized that food animals could be major reservoir. Recently, evidence is building to support this hypothesis that the farm environment may be a primary reservoir (Borucki et al., 2004; Okwumabua et al., 2005; Kabuki et al., 2004. In fact, in a recent prevalence study conducted at ruminant farms, Nightingale et al. (2005) detected presence of *L. monocytogenes* in cattle farm feedstuffs throughout the year at a constant rate, even though ruminant fecal material showed seasonal variation. Similarly, Van Kessel and co-workers (2004) determined the prevalence of food-borne pathogens in bulk tank milk collected from farms in 21 states in the United States and reported that 93% of the *L. monocytogenes* isolates were of serotypes 1/2a, 1/2b, and 4b. Furthermore, Borucki and co-workers (2004) compared PFGE profiles of dairy farm-associated strains to human sporadic and epidemic disease strains and reported distinct but overlapping profiles between the two populations. Interestingly, three farm environmental strains and one human sporadic strain showed identical pulsotypes with the 1985-California Jalisco outbreak strain, whose genome was recently sequenced (Borucki et al., 2004). In contrast, Eifert and co-workers (2005) reported that the pulsotypes of the serotype 4b complex (serotype 4b, and closely related 4d, and 4e) isolated from turkey processing plants were clearly distinct from those of confirmed epidemic-associated strains, while they observed that isolates with genetic markers typical of two major serotype 4b epidemic clonal groups, ECI and ECII, from environmental and raw product samples persisted. In addition, Kabuki and coworkers (2004) analyzed contamination patterns in three Latin-style fresh-cheese processing plants over a six-month period. Surprisingly, more than half of 36 serotype 4b isolates had indistinguishable subtyping pattern with EC-II, based on ribotyping and virulence gene polymorphisms.

There is still little data suggesting a strong association between subtypes of this pathogen and specific sources. Sauders and co-workers (2004; 2006) employed ribotyping to characterize several food strains from various retail outlets and from foods that were prepared or handled at retail over a five year period. They demonstrated that some subtypes that have been linked to human disease could persist in retail environments. On the other hand, a survey of *L. monocytogenes* subtypes from farm sources (animals, animal feedstuff, and farm environments), food processing environments, foods, and human clinical isolates using the same subtyping method showed a lack of clear-cut association between subtypes of this pathogen and potential reservoir from which the isolates were collected.

Sewage water and sewage sludge used for land applications deserve special attention as potential environmental reservoir for epidemic strains. Studies conducted in

European countries reported that the prevalent serotype in sewage is serotype 4b (De Luca et al. 1998; Marakusha et al., 1996; Paillard et al., 2005). The reported prevalence of *L. monocytogenes* serotype 4b in sewage is surprising and may reflect human carriage of this serotype, subclinical infection, or resistance of these strains to environmental conditions used for the production of activated sludge. In fact, serotype 4b was found to be predominant among strains from asymptomatic human carriers (MacGowan et al., 1994). However, prevalence studies involving sewage and sewage sludge used for land application in the United States as well as human carriage studies are highly limited and hence the ecology of epidemic strains remains to be further explored.

Epidemic Clone I specific Genomic Regions. Zheng and Kathariou (1997) performed transposon mutagenesis in order to identify sequences conferring to the bacterium survival under refrigerated conditions (4 C) but dispensable at higher temperature (25 C), and identified three cold growth essential loci (*ltrA*, *ltrB*, *ltrC*). When a probe derived from *ltrB* was used to hybridize *Hind*III digested genomic DNA, epidemic clone I strains showed a distinct restriction fragment length polymorphism (RFLP) which differentiated them from other serotype 4b strains and seemed to be a stable characteristic of the epidemic clones (Zheng and Kathariou, 1995). In addition, DNA from serotype 4b clinical isolates which had the epidemic clone specific RFLP resisted Sau3AI digestion, indicating methylation of cytosine at GATC sites. The resistance against phage infection and restriction with MboI, which recognizes GATC sites but sensitive to adenine methylation at GATC, suggested the presence of a homolog of the Sau3AI restriction-modification system in the genome of epidemic clones (Zheng and Kathariou, 1997).

The discovery of additional genotypic markers has been greatly facilitated by genome sequencing of selected strains and genome subtraction of an epidemic strain genome from EGD-e (Herd and Kocks, 2001; Nelson et al., 2004). Using such genotypic markers, Yildirim and co-workers (2004) demonstrated that several DNA regions (Fig. 1) that appeared to be unique to epidemic clone I -associated strains were also present in a subpopulation of serotype 4b strains isolated from foods and food environments without known implication in illness (Yildirim et al., 2004).

Probe 144 locus. This locus was first identified by Herd and Kocks (2001) in genome subtraction studies. Comparative genomic analysis indicated that only one ORF (LMOf2365_0687) in this region was specific to F2365 and closely related strains.The genes both immediately upstream and downstream of this ORF were highly conserved among the different *L. monocytogenes* genomes and *L. innocua* (Fig. 2). The transcription orientation of LMOf2365_0687 is the same as that of the upstream region genes, one of which encodes a putative serine threonine protein phosphatase. RT-PCR analysis of this region suggested that the ECI-specific LMOf2365_0687 was co-transcribed with the upstream region ORFs LMOf2365_0685 and LMOf2365_0686 (Yue, 2005). The percent GC content of the strain specific gene is 28%, which is substantially lower than the genome average (38%), evoking the possibility of horizontal transfer.

Interestingly, *L. innocua* harbors a gene encoding an internalin-like surface protein with a LPXTG motif in this region. It is well established that the internalin family of proteins contribute to pathogenicity of the bacterium (Cabanes et al., 2002). Also, phylogenetic analysis has suggested that *L .innocua* evolved from a *L. monocytogenes* serotype 4b ancestor (Schmid et al., 2005). Thus, the fragment 144 locus is likely to contribute to pathogenicity of epidemic strains

Probe 17B locus. This region harbors three ORFs that are specific to epidemic clone I, and was also identified by the subtractive hybridization approach of Herd and Kocks (2001). The flanking regions are conserved in other *L. monocytogenes* and *L. innocua* strains (Fig. 3). Although the three ORFs were annotated as hypothetical proteins they harbor conserved domains and/or motifs that hint about their putative functional roles. ORF2798 encodes a conserved domain AbiF –abortive infection phage resistance protein. ORF2799, on the other hand, harbors the HTH motif, specific to transcription factors, which recognize the xenobiotic responsive element (XRE). The two ORFs are divergently transcribed with a1645 bp intergenic region between them. ORF2800 shows similarity to site-specific recombinases. However, ORF2800 is annotated as a pseudogene, i.e., it was found to contain premature stop codons and/or frameshifts.

Curiously, ORFs 2795-2797, upstream of the probe 17B locus, share 68-74% nucleotide sequence similarity among themselves, suggesting that they may have arisen by gene duplication. Other *L. monocytogenes* genomes also harbor one or two homologues of these sequences (Fig. 3). Downstream of the probe 17 B locus, ORFs 2801 and 2802 have been annotated as putative *gidA* and *trmE*, respectively, and encode the putative glucose-inhibited division protein GidA, which plays a role in RNA modification (Sha et al., 2004) and TrmE, involved in tRNA modification (Brown, 2005), respectively. A gene with homology to the putative *gidB* was identified further upstream of the locus, and appeared to be conserved among the different *Listeria* genomes.

Surprisingly, a recent transcriptional analysis of the genes in this locus showed that ORFs 2800 and 2799 were co-transcribed, even though ORF2800 was annotated to be a pseudogene. In addition, ORFs 2801 and 2802 appeared to be co-transcribed, whereas ORF 2798 appeared to be transcribed onto a separate mRNA molecule (Yue, 2005). In contrast, all surveyed ORFs in this region (between *trmE* and *gidB* in the genome of the serotype 1/2a strain appeared to be co-transcribed (Yue, 2005). The GC% content of the ECI-specific ORFs is 30%, lower than the genome average (Fig. 3). This, in addition to the presence of the putative mobile element- associated recombinase, point toward the possibility that the ECI specific genes in this region might have been acquired from other organisms.

Probe 85 locus. Methylation of cytosines at GATC sites in the genome of epidemic clone-I strains was first identified in 1997 by detecting resistance of genomic DNA of these organisms to digestion by Sau3AI (Zheng and Kathariou, 1997). Sequences with homology to methyltransferases were identified during subtractive hybridization of an epidemic clone I strain and the DNA fragment harboring the sequences was named “clone 85” (Herd and Kocks, 2001). The function of the gene cassette is to encode a type-II restriction endonuclease (Sau3AI) and its cognate enzyme transferring a methyl group to cytosine at the GATC recognition sequence, thereby rendering the DNA resistant to digestion by Sau3AI enzyme. The genome sequence of the ECI strain F2365 (Nelson et al., 2004) facilitated identification of the putative locus *in silico* (Yildirim et al., 2004) and a recent mutational analysis confirmed the involvement of the cassette in methylation of GATC sites (D. Elhanafi, unpublished results). Comparative genome analysis showed that the gene cassette was located in a highly conserved locus, downstream of a putative PTS system (Fig. 4). The GC% content is lower (31%) than the genome average (38%), suggesting the possibility of horizontal transfer of the gene cassette.

Probe 133 locus. This locus comprises six ECI-specific ORFs that show homology to membrane-associated proteins including two hypothetical proteins (Fig. 5). Interestingly, the six ORFs were bracketed by 42 bp palindromic direct repeats. The upstream and downstream region ORFs are all conserved in other *Listeria* strains whose genome has been sequenced. It is striking that the conserved ORFs in the flanking region include a putative PTS system IIA component and translation elongation factor Tu, respectively. DNA hybridization results indicated that although the region is present in all ECI strains screened it, can also be detected in certain lineage II strains of serotype 1/2b and 3b (Herd and Kocks, 2001; Yue, 2005). Also, the region was detected in a recently identified epidemic lineage of *L. monocytogenes* serotype 4b implicated in an outbreak in Winston Salem, North Carolina in 2001. Other serotype 4b strains, strains of serotype 1/2a, and the non-pathogenic species *L. innocua* did not harbor this locus in their genome (Yue, 2005, unpublished).

Conclusions and future perspectives. Understanding the underlying evolutionary mechanisms in the emergence of successful clonal variants is crucial for epidemiological investigations. Even though genetic diversity of clonal complexes is limited and arises through a largely asexual mode because of population bottlenecks, some variants that display increased fitness in the environment yet reduced pathogenicity may arise in the epidemic clone population (Zhu et al., 2001); still some other variants may follow the evolutionary pathway leading to increased host adaptation, transmission, and virulence, thereby causing epidemics and pandemics when a susceptible human population are exposed to a particularly invasive and fit strains (Wirth et al., 2660; Zhu et al., 2001). Thus, it is important to detect and monitor clonal groups of clinical importance and to characterize their adaptive response in relevant ecological niches. As has been noted with other pathogens, in the case of *L. monocytogenes* gene deletion and acquisition in the background of point mutations give rise to new variants (Ward et al., 2004; Nightingale et al., 2005). Even though epidemic clone I strains of *L. monocytogenes* have shown remarkable genome stability over the entire period that outbreaks of listeriosis have been monitored, genetic variability during an individual epidemic and variation during epidemic spread have been observed (Clark et al., 2000; Zhang et al., 2003). This leads us to postulate that epidemic clones of *L. monocytogenes* serotype 4b follow an evolutionary model which involves stochastic processes of clonal expansion/reduction and clonal replacement. For instance, while epidemic clone I strains that caused outbreaks in Nova Scotia (1981), Switzerland (1983-1987), Denmark (1985-1987), California (1985), and France (1992) showed indistinguishable pulsotypes, another serotype 4b strain responsible for the Massachusetts outbreak (1983) showed distinct genotype compared to ECI strains (Kathariou, 2003). In addition, ECII strains, implicated in the multistate outbreak in 1998-1999, representing a new genotype of serotype 4b, have been identified. The sequenced ECII strain from this outbreak has acquired an 80 Kb plasmid decorated with mobile genetic elements (Nelson et al., 2004) and harbored a region of divergence in a locus that was otherwise specific for serotype 4b strains, including ECI (Evans et al., 2004). Related strains were implicated in another multistate outbreak in 2002 (MMWR ref). Even though ECII was relatively recently identified and there is no report of its detection other than in the United States, current data suggest this clone seems to represent clonal expansion of a previously unusual serotype 4b strain. As a result, despite strong clonality within serotype 4b, recombination and horizontal gene transfer events do contribute to the emergence of new genetic variants of epidemic strains of *L. monocytogenes*. It is possible, as suggested by certain models, that epidemic clones have increased mutator frequency and recombination, which results in a higher level of virulence (Wirth et al., 2006).

Epidemics brought about by fit variants of clonal complexes must be constantly monitored and characterized in order to design effective detection and prevention strategies. To this end, rational design of nucleic acid-based rapid detection methods requires target loci specific to epidemic clones. Sequence-based subtyping methods (e.g. Multilocus Sequence Typing, Multi-virulence-locus sequence typing) targeting conserved genomic regions in epidemic clones as well as fast evolving virulence genes may better delineate the clonal complexes within the serotype 4b population. Marker genes with which such clones can be screened for and detected in the food environment are invaluable for effective control measures for foodborne pathogens.

The availability of genome sequence data for epidemic clones (I, II, and III) and comparative genomics has facilitated identification of strain and serotype-specific genes. Detection of the strain-specific genes in natural populations of serotype 4b strains that are genetically close to epidemic clones will clearly prove the utility of these genes as genetic markers. However, most of the strain-specific gene clusters are annotated as 'hypothetical or unknown proteins' (Neslon et al., 2004). Characterization of these genes employing postgenomic tools, in addition to sophisticated *in vivo* studies and conventional functional genomics approaches, will provide further insight into adaptation, transmission, and virulence potential of strains that show a genetic make-up similar to that of epidemic clones in natural populations of *L. monocytogenes*.

REFERENCES

1. Achtman M., 1997. Microevolution and epidemic spread of serogroup A *Neisseria meningitides* -a review. Gene. 192:135-40.

2. Autio T, Lunden J, Fredriksson-Ahomaa M, Bjorkroth J, Sjoberg AM, Korkeala H., 2002. Similar *Listeria monocytogenes* pulsotypes detected in several foods originating from different sources. Int J Food Microbiol. 77:83-90.

3. Awadalla P., 2003. The evolutionary genomics of pathogen recombination. Nat Rev Genet. 4:50-60

4. Bibb WF, Gellin BG, Weaver R, Schwartz B, Plikaytis BD, Reeves MW, Pinner RW, Broome CV. 1990. Analysis of clinical and food-borne isolates of *Listeria monocytogenes* in the United States by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. Appl Environ Microbiol. 56:2133-41

5. Bibb WF, Schwartz B, Gellin BG, Plikaytis BD, Weaver RE. 1989. Analysis of *Listeria monocytogenes* by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. Int J Food Microbiol. 8:233-9

6. Boerlin P, Boerlin-Petzold F, Bannerman E, Bille J, Jemmi T., 1997. Typing *Listeria monocytogenes* isolates from fish products and human listeriosis cases. Appl Environ Microbiol. 63:1338-43

7. Boerlin P, Piffaretti JC., Typing of human, animal, food, and environmental isolates of *Listeria monocytogenes* by multilocus enzyme electrophoresis. Appl Environ Microbiol. 57:1624-9

8. Borucki MK, Reynolds J, Gay CC, McElwain KL, Kim SH, Knowles DP, Hu J., 2004. Dairy farm reservoir of *Listeria monocytogenes* sporadic and epidemic strains. J Food Prot. 67:2496-9

9. Brosch R, Chen J, Luchansky JB., 1994. Pulsed-field fingerprinting of *Listeria*e: identification of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. Appl Environ Microbiol. 60:2584-92

10. Brown ED., 2005. Conserved P-loop GTPases of unknown function in bacteria: an emerging and vital ensemble in bacterial physiology. Biochem Cell Biol. 83:738-46

11. Buchrieser C, Brosch R, Catimel B, Rocourt J., 1993. Pulsed-field gel electrophoresis applied for comparing *Listeria monocytogenes* strains involved in outbreaks. Can J Microbiol. 39:395-401

12. Cabanes D, Dehoux P, Dussurget O, Frangeul L, Cossart P., 2002. Surface proteins and the pathogenic potential of *Listeria monocytogenes.* Trends Microbiol. 10:238-45

13. Call DR, Borucki MK, Besser TE., 2003. Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of *Listeria monocytogenes.* J Clin Microbiol. 41:632-9

14. Centers for Disease Control (1999). Multistate outbreak of listeriosis--United States, MMWR Morb Mortal Wkly Rep 47:1117-8

15. Centers for Disease Control and Prevention (1989). Epidemiological notes and reports listeriosis associated with consumption of Turkey Franks. MMWR Weekly, 38, 267-268.

16. Clark EE, Wesley I, Fiedler F, Promadej N, Kathariou S., 2000. Absence of serotype-specific surface antigen and altered teichoic acid glycosylation among epidemic-associated strains of *Listeria monocytogenes*. J Clin Microbiol. 38:3856-9

17. Cohan FM., 2002. What are bacterial species? Annu Rev Microbiol. 56:457-87

18. De Chastellier, C., and P. Berche. 1994. Fate of *Listeria monocytogenes* in murine macrophages: evidence for simultaneous killing and survival of intracellular bacteria. Infect. Immun. 62:543–553

19. Dijkshoorn L, Ursing BM, Ursing JB., 2000. Strain, clone and species: comments on three basic concepts of bacteriology. J Med Microbiol. 49:397-401

20. Doumith M, Cazalet C, Simoes N, Frangeul L, Jacquet C, Kunst F, Martin P, Cossart P,Glaser P, Buchrieser C., 2004 New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. Infect Immun. 72:1072-83

21. Eifert JD, Curtis PA, Bazaco MC, Meinersmann RJ, Berrang ME, Kernodle S, Stam C, Jaykus LA, Kathariou S., 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog Dis. 2:192-200

22. Evans MR, Swaminathan B, Graves LM, Altermann E, Klaenhammer TR, Fink RC, Kernodle S, Kathariou S., 2004. Genetic markers unique to *Listeria monocytogenes* serotype 4b differentiate epidemic clone II (hot dog outbreak strains) from other lineages. Appl Environ Microbiol. 70:2383-90

23. Farber JM, Peterkin PI., 1991. *Listeria monocytogenes*, a food-borne pathogen. Microbiol Rev. 55:476-511

24. Feil, EJ., 2004. Small change: keeping pace with microevolution. Nat Rev Microbiol. 2:483-95

25. Garcia-Alvarez M, Chaves F, Sanz F, Otero JR., 2006. Molecular epidemiology of *Listeria monocytogenes* infections in a health district of Madrid in a 3-year period (2001-2003). Enferm Infecc Microbiol Clin. 24:86-9

26. Gendel, S. M., and J. Ulaszek. 2000. Ribotype analysis of strain distribution in *Listeria monocytogenes*. J. Food Prot. 63:179–185

27. Gilbreth SE, Call JE, Wallace FM, Scott VN, Chen Y, Luchansky JB., 2005. Relatedness of *Listeria monocytogenes* Isolates recovered from selected ready-to-eat foods and listeriosis patients in the United States. Appl Environ Microbiol. 71:8115-22.

28. Graves LM, Hunter SB, Ong AR, Schoonmaker-Bopp D, Hise K, Kornstein L, DeWitt WE, Hayes PS, Dunne E, Mead P, Swaminathan B., 2005. Microbiological aspects of the investigation that traced the 1998 outbreak of listeriosis in the United States to contaminated hot dogs and establishment of molecular subtyping-based surveillance for *Listeria monocytogenes* in the PulseNet network. J Clin Microbiol. 43:2350-5

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

30. Gray MJ, Zadoks RN, Fortes ED, Dogan B, Cai S, Chen Y, Scott VN, Gombas DE, Boor KJ, Wiedmann M., 2004. *Listeria monocytogenes* isolates from foods and humans form distinct but overlapping populations. Appl Environ Microbiol. 70:5833-41

31. Grif K, Heller I, Wagner M, Dierich M, Wurzner R., 2006. A comparison of *Listeria monocytogenes* serovar 4b isolates of clinical and food origin in Austria by automated ribotyping and pulsed-field gel electrophoresis. Foodborne Pathog Dis. 3:138-41

32. Gupta S, Maiden MC., 2001. Exploring the evolution of diversity in pathogen populations. Trends Microbiol. 9:181-5

33. Herd M, Kocks C., 2001; use consistent format after year Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. Infect Immun. 69:3972-9

34. Hof H, Rocourt J., 1992. Is any strain of *Listeria monocytogenes* detected in food a health risk? Int J Food Microbiol. 16:173-82.

35. Hof, H., 2003. History and epidemiology of listeriosis. FEMS Immunol Med Microbiol. 35:199-202

36. Jacquet C, Catimel B, Brosch R, Buchrieser C, Dehaumont P, Goulet V, Lepoutre A, Veit P, Rocourt J., 1995. Investigations related to the epidemic strain involved in the French listeriosis outbreak in 1992. Appl Environ Microbiol. 61:2242-6

37. Jeffers GT, Bruce JL, McDonough PL, Scarlett J, Boor KJ, Wiedmann M., 2001. Comparative genetic characterization of *Listeria monocytogenes* isolates from human and animal listeriosis cases. Microbiology. 147:1095-104

38. Kabuki DY, Kuaye AY, Wiedmann M, Boor KJ., 2004. Kabuki DY, Kuaye AY, Wiedmann M, Boor KJ., Molecular subtyping and tracking of *Listeria monocytogenes* in latin-style fresh-cheese processing plants. Title format J Dairy Sci. 87:2803-12

39. Kathariou S., 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J Food Prot. 65:1811-29. Review.

40. Kathariou, S., 2003. Foodborne outbreaks of listeriosis and epidemic-associated lineages of *Listeria monocytogenes*. In: M. E. Torrence and R. E. Isaacson (Ed), Microbial Food Safety in Animal Agriculture, Iowa State Press.

41. Land R, Reeves PR., 2001. When does a clone deserve a name? A perspective on bacterial species based on population genetics. Trends Microbiol. 9:419-24

42. Lei XH, Fiedler F, Lan Z, Kathariou S., 2001. A novel serotype-specific gene cassette (gltA-gltB) is required for expression of teichoic acid-associated surface antigens in *Listeria monocytogenes* of serotype 4b. J Bacteriol. 183:1133-9

43. Levin BR., 1981. Periodic selection, infectious gene exchange and the genetic structure of E. coli populations. Genetics. 99:1-23.

44. MacGowan, A. P., K. Bowker, J. McLauchlin, P. M. Bennett, and D. S. Reeves. 1994. The occurrence and seasonal changes in the isolation of *Listeria* spp. in shop bought food stuffs, human faeces, sewage and soil from urban sources. Int. J. Food Microbiol. 21:325–334

45. Marakusha BI, Darwich K, Tartakovskii IS., 1996. Characteristics of *Listeria monocytogenes* strains isolated in Russia and their typing using pulse electrophoresis. Zh Mikrobiol Epidemiol Immunobiol. May-Jun;60-4

46. Mazzotta AS, Gombas DE., 2001. Heat resistance of an outbreak strain of *Listeria monocytogenes* in hot dog batter. J Food Prot. 64:321-4

47. 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

48. Mead PS, Dunne EF, Graves L, Wiedmann M, Patrick M, Hunter S, Salehi E, Mostashari F, Craig A, Mshar P, Bannerman T, Sauders BD, Hayes P, Dewitt W, Sparling P, Griffin P, Morse D, Slutsker L, Swaminathan B. 2006. Nationwide outbreak of listeriosis due to contaminated meat.. Epidemiol Infect. 1:1-8

49. Meinersmann RJ, Phillips RW, Wiedmann M, Berrang ME., 2004. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes reveals clonal and recombination histories of three lineages. Appl Environ Microbiol. 70:2193-203

50. Multistate Outbreak of Listeriosis --- United States, 2000. MMWR Weekly 49(50);1129-1130

51. Nadon CA, Woodward DL, Young C, Rodgers FG, Wiedmann M., 2001. Correlations between molecular subtyping and serotyping of *Listeria monocytogenes*. J Clin Microbiol. 39:2704-7

52. Nelson KE, Fouts DE et al., 2004. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. Nucleic Acids Res. 2004. 32:2386-95

53. Nightingale KK, Fortes ED, Ho AJ, Schukken YH, Grohn YT, Wiedmann M., 2005. Evaluation of farm management practices as risk factors for clinical listeriosis and fecal shedding of *Listeria monocytogenes* in ruminants. J Am Vet Med Assoc. 227:1808-14

54. Nightingale KK, Windham K, Wiedmann M., 2005. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. J Bacteriol. 187:5537-51

55. Okwumabua O, O'Connor M, Shull E, Strelow K, Hamacher M, Kurzynski T, Warshauer D., 2005. Characterization of *Listeria monocytogenes* isolates from food animal clinical cases: PFGE pattern similarity to strains from human listeriosis cases. FEMS Microbiol Lett. 249:275-81

56. Orskov F, Orskov I. 1983. From the national institutes of health. Summary of a workshop on the clone concept in the epidemiology, taxonomy, and evolution of the enterobacteriaceae and other bacteria. J Infect Dis. 148:346-57

57. Paillard D, Dubois V, Thiebaut R, Nathier F, Hoogland E, Caumette P, Quentin C., 2005. Occurrence of *Listeria* spp. in effluents of French urban wastewater treatment plants. Appl Environ Microbiol. 71:7562-6

58. Piffaretti JC, Kressebuch H, Aeschbacher M, Bille J, Bannerman E, Musser JM, Selander RK, Rocourt J., 1989. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. Proc Natl Acad Sci; 86:3818-22

59. Promadej N, Fiedler F, Cossart P, Dramsi S, Kathariou S., 1999. Cell wall teichoic acid glycosylation in *Listeria monocytogenes* serotype 4b requires gtcA, a novel, serogroup-specific gene. J Bacteriol. 181:418-25

60. Rasmussen, O. F., P. Skouboe, L. Dons, L. Rossen, and J. E. Olsen. 1995. *Listeria monocytogenes* exists in at least three evolutionary lines: evidence from flagellin, invasive associated protein and listeriolysin O genes. Microbiology 141:2053–2061.

61. Revazishvili T, Kotetishvili M, Stine OC, Kreger AS, Morris JG Jr, Sulakvelidze A. 2004. Comparative analysis of multilocus sequence typing and pulsed-field gel

electrophoresis for characterizing *Listeria monocytogenes* strains isolated from environmental and clinical sources. J Clin Microbiol. 42:276-85.

62. Roberts A, Nightingale K, Jeffers G, Fortes E, Kongo JM, Wiedmann M.2006. Genetic and phenotypic characterization of *Listeria monocytogenes* lineage III. Microbiology. 152:685-93

63. Rocourt J, Jacquet C, Reilly A. 2000. Epidemiology of human listeriosis and seafoods. Int J Food Microbiol. 62:197-209

64. Rocourt, J., Jacquet, Ch., Bille, J., 1997. Human listeriosis, 1991/1992. WHO/FNU/FOS/ 97.1. World Health Organization, Geneva, Switzerland.

65. Rocourt,J., 1999. in *Listeria*, listeriosis and food safety, The Genus *Listeria* and *Listeria monocytogenes*: Phylogenetic position, taxonomy, and identification. pg 1-21. Edited by Ryser and Marth. Marcel Dekker, Inc. New York

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

67. Sauders BD, Mangione K, Vincent C, Schermerhorn J, Farchione CM, Dumas NB, Bopp D, Kornstein L, Fortes ED, Windham K, Wiedmann M., 2004. Distribution of *Listeria monocytogenes* molecular subtypes among human and food isolates from New York State shows persistence of human disease--

associated *Listeria monocytogenes* strains in retail environments. J Food Prot. 67:1417-28

68. Schmid MW, Ng EY, Lampidis R, Emmerth M, Walcher M, Kreft J, Goebel W, Wagner M, Schleifer KH., 2005. Evolutionary history of the genus *Listeria* and its virulence genes. Syst Appl Microbiol. 28:1-18

69. Seeliger, H.P.R. 1961. Listeriosis. Basel:Karger.

70. Sha J, Kozlova EV, Fadl AA, Olano JP, Houston CW, Peterson JW, Chopra AK., 2004. Molecular characterization of a glucose-inhibited division gene, *gidA,* that regulates cytotoxic enterotoxin of *Aeromonas hydrophila.* Infect Immun. 72:1084-95

71. Spratt BG, Maiden MC., 1999. Bacterial population genetics, evolution and epidemiology. Philos Trans R Soc Lond B Biol Sci. 354:701-10.

72. Stone, S. C., and J. Shoenberger. 2001. Update: Multistate outbreak of listeriosis--United States, 2000. Ann Emerg Med 38:339-41.

73. Tibayrenc, M., 2005. Bridging the gap between molecular epidemiologists and evolutionists. Trends Microbiol. 13:575-80

74. Tran HL, Kathariou S., 2002. Restriction fragment length polymorphisms detected with novel DNA probes differentiate among diverse lineages of

serogroup 4 *Listeria monocytogenes* and identify four distinct lineages in serotype 4b. Appl Environ Microbiol. 68:59-64

75. Van Kessel JS, Karns JS, Gorski L, McCluskey BJ, Perdue MLg, 2004. Prevalence of Salmonellae, *Listeria monocytogenes*, and fecal coliforms in bulk tank milk on US dairies. J Dairy Sci. 87:2822-30

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

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

78. Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MC, Ochman H, Achtman M.,g 2006. Sex and virulence in *Escherichia coli*: an evolutionary perspective. Mol Microbiol. 60:1136-1151

79. Yildirim S, Lin W, Hitchins AD, Jaykus LA, Altermann E, Klaenhammer TR, Kathariou S. 2004. Epidemic clone I-specific genetic markers in strains of *Listeria monocytogenes* serotype 4b from foods. Appl Environ Microbiol. 70:4158-64

80. Yue, L. 2005. Master thesis. Gene cassettes unique to epidemic-associated lineages of *Listeria monocytogenes* serotype 4b. North Carolina State University

81. Zheng W, Kathariou S., 1995. Differentiation of epidemic-associated strains of *Listeria monocytogenes* by restriction fragment length polymorphism in a gene region essential for growth at low temperatures (4 degrees C). Appl Environ Microbiol. 61:4310-4

82. Zheng W, Kathariou S., 1997. Host-mediated modification of Sau3AI restriction in *Listeria monocytogenes*: prevalence in epidemic-associated strains. Appl Environ Microbiol. 63:3085-9

83. Zhu P, van der Ende A, Falush D, Brieske N, Morelli G, Linz B, Popovic T, Schuurman IG, Adegbola RA, Zurth K, Gagneux S, Platonov AE, Riou JY, Caugant DA, Nicolas P, Achtman M., 2001. Fit genotypes and escape variants of subgroup III *Neisseria meningitidis* during three pandemics of epidemic meningitis. Proc Natl Acad Sci U S A. 98:5234-9

Table 1. Epidemic Clones of *Listeria monocytogenes*

Epidemic Clone (Serotype)	Location (Year)	Source	Reference
ECI (4b)	Nova Scotia (1981)	Coleslaw	Schlech et al. 1983
	Switzerland (1983-1987)	Cheese	Bille, 1989
	California (1985)	Mexican style cheese	Linnan et al., 1988
	Denmark (1985-1987)	Unidentified source	Samuelsson et al, 1990
	France (1992)	Jellied pork tongue	Goulet et al. 1995
	Texas (2003)	Mexican style cheese	Bordoni, P., 2004
ECIa (4b)	Massachusetts (1983)	Pasteurized milk	Fleming et al., 1985
ECII (4b)	Multistate, USA (1998-1999)	Hotdog	CDC, 1998
ECIII (1/2a)	Multistate, USA (2000)	Turkey deli meats	CDC, 2001a

Table 2. Outbreaks caused by serotype 4b strains and occurred in Europe or North America

Year	Location	No of cases	Mortality rate %	Source	Reference
1981	Nova Scotia	41	41.5	Coleslaw	Schlech et al., 1983
1983	Massachusetts	49	29.0	Pasteurized milk	Fleming et al., 1985
1983-1987	Switzerland	122	25.0	Vacherin Mont d'Or cheese	Bula et al., 1995
1985	California	142	34.0	Mexican style cheese	Linnan et al., 1985
1986-1987	Pennsylvania	36	44.4	Brie cheese, salami	Schwartz, et al., 1989
1987-1989	England	355	26.5	Pâté and meat spreads	McLaughlin et al., 1991
1989-1990	Denmark	23	0.0	Blue cheese	Jensen et al.,
1992	France	280	22.5	Pork tongue in jelly	Jacquet et al.1995
1993	France	38	28.9	Rillettes	Goulet, 1998
1994-1995	Sweden	9	22.2	Seafood	Ericsson et al. 1997
1995	France	33	20.0	Cheese	Jacquet et al. 1995
1997	France	14	NR	Cheese	Ryser, 1999a
1997	Italy	1566	NR	Corn salad	Aureli et al., 2000
1998-1999	Multistate, US	101	20.8	Hot dogs, deli meets	Mead et al. 1999
2000-2001	NC	12	41.7	Mexican style cheese	CDC, 2001
2002	8 North E. states	63	11.1	Deli turkey meat	CDC, 2002b
2003	TX	12	NR	Mexican style cheese	CDC

Adapted from FDA/USDA risk assessment study, 2003. http://www.cfsan.fda.gov/~dms/listrisk.html
NR. Not Reported

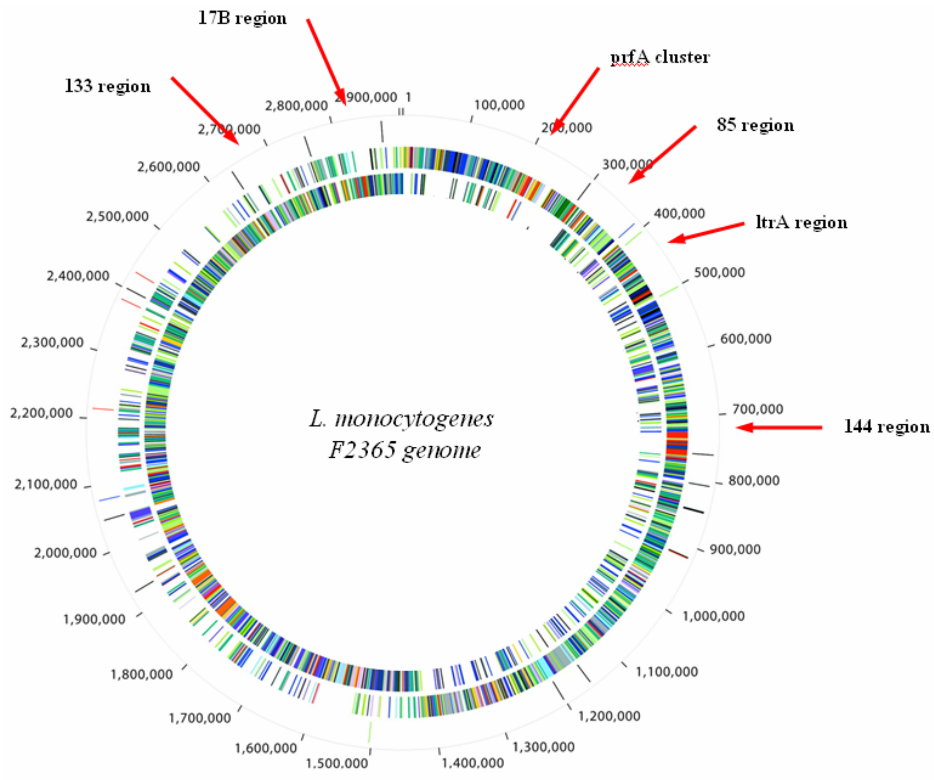


Fig. 1. Respective distribution of ECI specific DNA fragment in the genome of F2365 (The outermost circle indicates the AscI (black), NotI (red), SfiI (blue) and SrfI (green) restriction map of the closed *L. monocytogenes* serotype 4b F2365 strain. **(The figure was reproduced with the permission from Nelson et al., 2004, Nucleic Acid Research, Oxford University Press)**

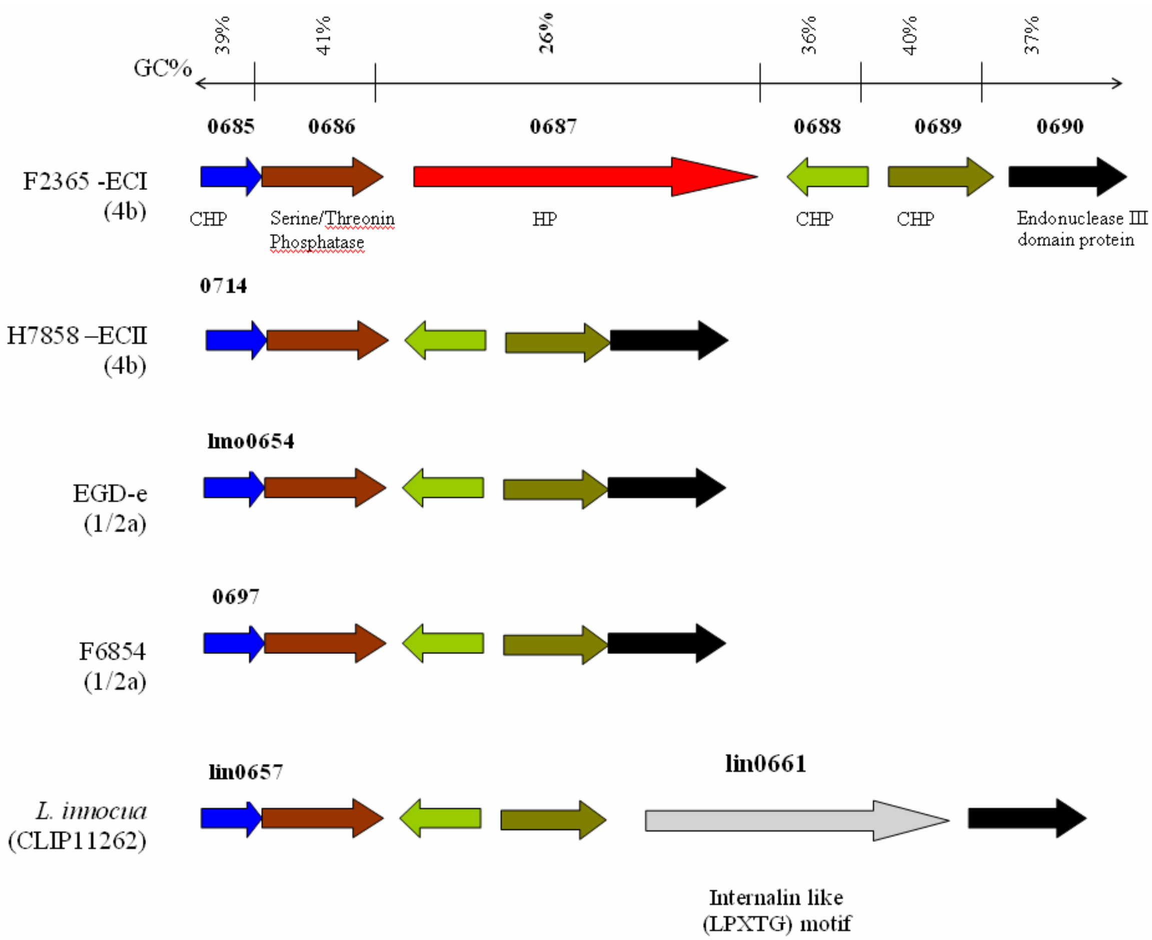


Figure 2. Comparative genomic analysis of probe 144 region.
(HP: hypothetical protein, CHP: conserved hypothetical protein)

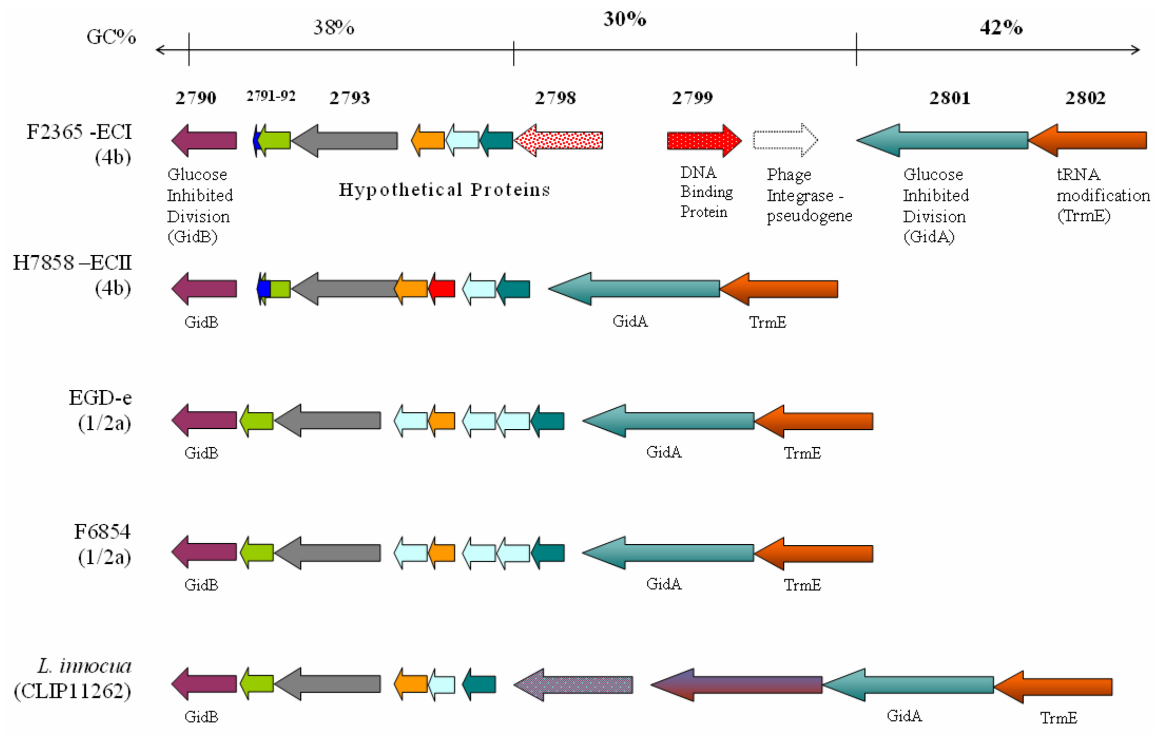


Figure 3. Comparative genomic analysis of probe 17B region

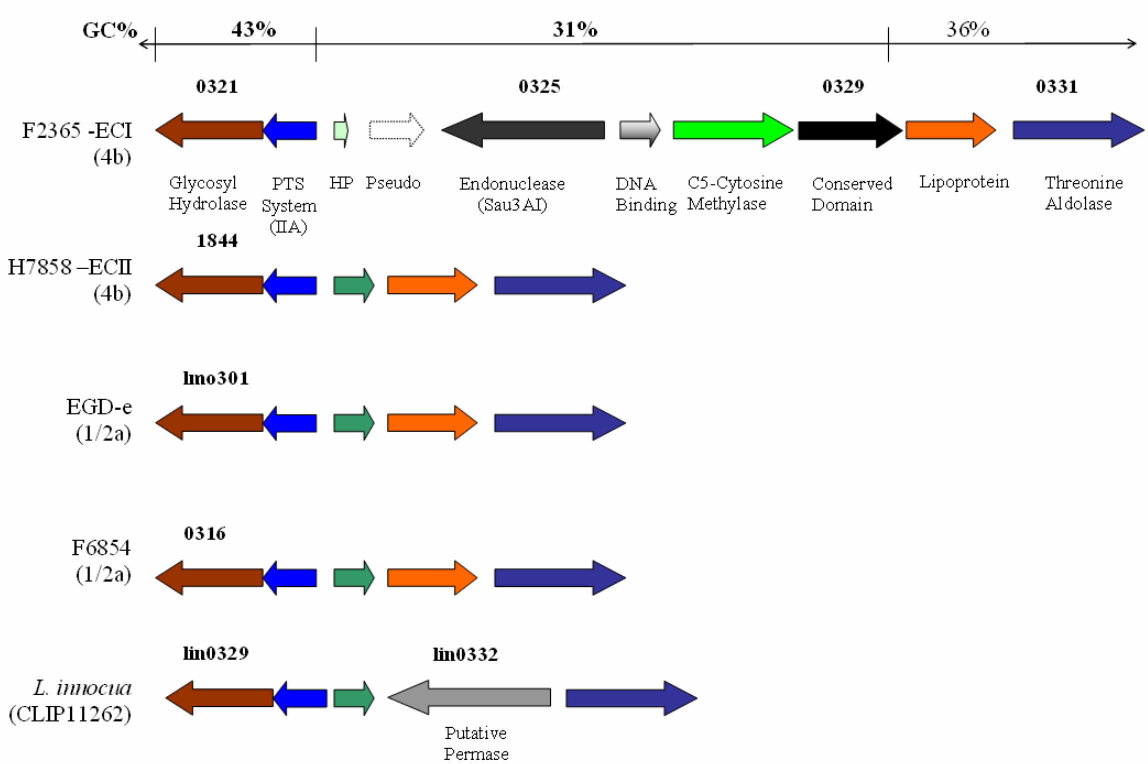


Figure 4. Comparative genomic analysis of probe 85 region.
(HP: hypothetical protein, CHP: conserved hypothetical protein)

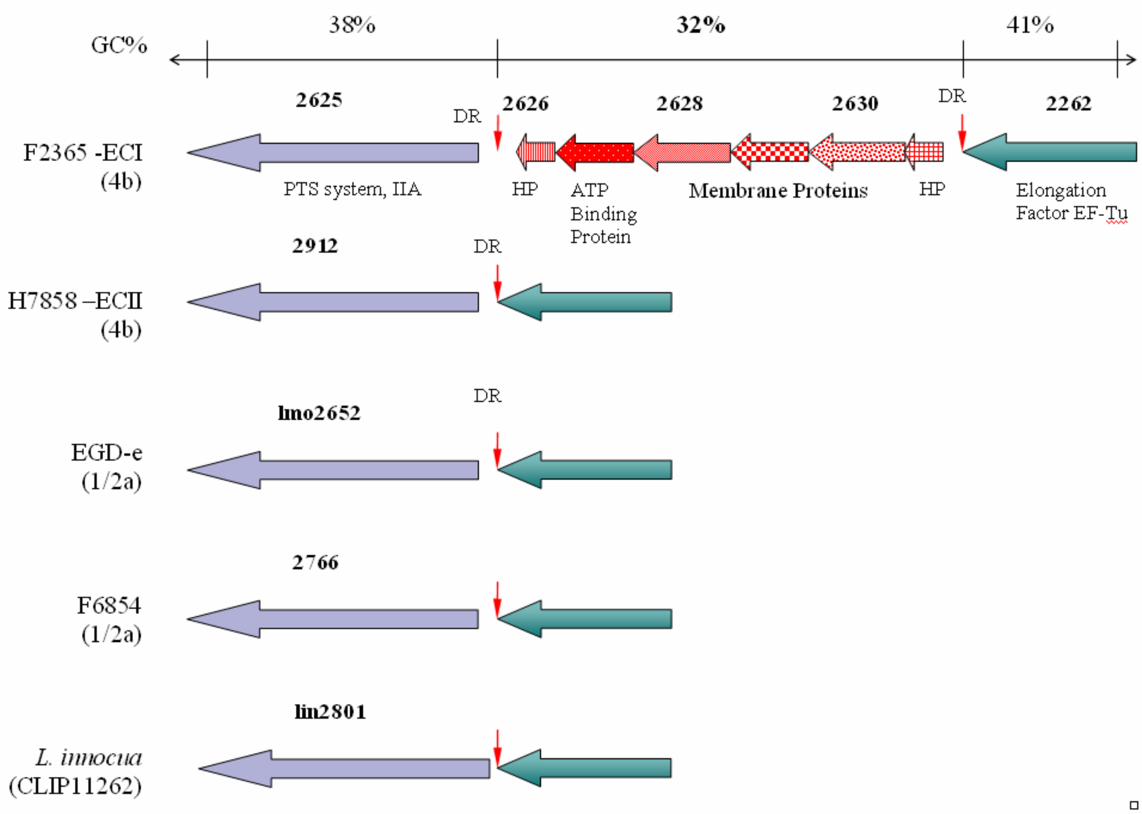


Figure 5. Comparative genomic analysis of probe 133region.
(HP: hypothetical protein)

CHAPTER III

EPIDEMIC CLONE I-SPECIFIC GENETIC MARKERS IN STRAINS OF LISTERIA MONOCYTOGENES SEROTYPE 4B FROM FOODS

Suleyman Yildirim,[1] Wen Lin,[2] Anthony D. Hitchins,[3] Lee-Ann Jaykus,[1] Eric Altermann,[1] Todd R. Klaenhammer,[1] And Sophia Kathariou[1*]
Department Of Food Science, North Carolina State University, Raleigh, North Carolina,[1] Food And Drug Administration, Irvine, California,[2] Food And Drug Administration, Washington, D.C.[3]

Published in: Applied and Environmental Microbiology **(the American Society for Microbiology), July 2004.**

Corresponding author: **Sophia Kathariou**
North Carolina State University
Department of Food Science
Campus Box 7624
Raleigh, NC 27695-7624
Phone: (919) 513-2075
Fax: (919) 515-7124
E-mail: skathar@unity.ncsu.edu

ABSTRACT

Listeria monocytogenes contamination of ready-to-eat foods has been implicated in numerous outbreaks of food-borne listeriosis. However, the health hazards posed by L. monocytogenes detected in foods may vary, and speculations exist that strains actually implicated in illness may constitute only a fraction of those that contaminate foods. In this study, examination of 34 serogroup 4 (putative or confirmed serotype 4b) isolates of L. monocytogenes obtained from various foods and food-processing environments, without known implication in illness, revealed that many of these strains had methylation of cytosines at GATC sites in the genome, rendering their DNA resistant to digestion by the restriction endonuclease Sau3AI. These strains also harbored a gene cassette with putative restriction-modification system genes as well as other, genomically unlinked genetic markers characteristic of the major epidemic-associated lineage of L. monocytogenes (epidemic clone I), implicated in numerous outbreaks in Europe and North America. This may reflect a relatively high fitness of strains with these genetic markers in foods and food-related environments relative to other serotype 4b strains and may partially account for the repeated involvement of such strains in human food-borne listeriosis.

INTRODUCTION

Food contamination by *Listeria monocytogenes* has been implicated in numerous outbreaks and sporadic cases of human illness. Most commonly implicated in listeriosis are highly processed, ready-to-eat (RTE) foods that are kept refrigerated for various periods of time. At risk for listeriosis are people in the extremes of age, pregnant women and their fetuses, cancer patients, and others experiencing immunosuppression (13, 24, 35, 38).

Listeriosis can have severe symptoms (septicemia, meningitis, and stillbirths) and a high mortality rate (20 to 30%). Hence, regulations exist in numerous nations concerning the density (e.g., <1 CFU/25 g) of cells of the etiologic agent permissible in RTE foods. Such regulations are based on the hypothesis that any L. monocytogenes strain that can be detected in RTE foods has the potential to pose serious hazards to human health.

The potential hazard posed by *Listeria*l contamination of RTE foods can be influenced by the number of cells at the point of consumption, which would depend on conditions of storage, type of food matrix and its impact on growth, presence of competing microflora and antimicrobial agents, etc. In addition, the strain type of L. monocytogenes involved may be of importance. It is likely, based on studies with other bacterial pathogens, that some strains and strain clusters (clonal groups) within the species might be more pathogenic than others. Speculations have been formulated that only a fraction of the strains of L. monocytogenes found in foods may be capable of causing human illness (20).

There is indeed evidence that the repertoire of strains capable of contaminating food is wider than that of strains recovered from patients with listeriosis. Specifically, even though foods may be contaminated by strains of various serotypes, more than 95% of human listeriosis cases involve just three serotypes, 1/2a, 1/2b, and 4b. In addition, most outbreaks have involved a small number of genetically related strains of serotype 4b (6, 24, 31, 35). Furthermore, strains of serotype 1/2c are relatively rare among human clinical isolates but are frequently isolated from foods (23).

In spite of the incongruities in the observed distribution of clinical and food-derived strains, significant differences in virulence characteristics between human clinical and food-derived strains have not been consistently detected (5, 9, 11, 28, 32). However, recent findings suggest that food-derived and clinical isolates tend to differ, at the population level, in parameters that may be of relevance during infection. These include the readiness to resume growth at 37°C following starvation at 4°C (1) and the apparent molecular weight of the virulence determinant ActA (21). In addition, the protein profiles of three food-derived isolates of serotype 1/2a were found to have significant differences from those of the serotype 1/2a strain EGDe, the genome of which has been sequenced, suggesting differences either in gene content or in expression (34). Overall, however, limited information currently exists concerning the genomic diversity and virulence of L. monocytogenes strains isolated from routine surveys of food with no known implication in illness. It is expected that such studies will be greatly facilitated by the recent availability of the genome sequence of L. monocytogenes, including the sequences of strain EGDe of serotype 1/2a (15) and strain F2365 of serotype 4b (www.tigr.org). The latter strain has been implicated in a food-borne outbreak of listeriosis (California outbreak) involving contaminated Mexican-style cheese (38).

Serotype 4b strains closely related to F2365 have been found responsible for a number of temporally and geographically distinct outbreaks in North America and

Europe and represent a "clone" in the epidemiological sense (29), referred to as epidemic clone I (23, 24). Epidemic clone I strains share certain special characteristics: their genomic DNA is resistant to digestion by the enzyme Sau3AI, suggesting methylation of cytosines at GATC sites (18, 44); they harbor restriction fragment length polymorphisms that differentiate them from other serotype 4b strains (41, 43); and they harbor a number of unique genomic fragments and gene clusters commonly absent from other serotype 4b strains (18; S. Yildirim and S. Kathariou, unpublished results).

In this study, we characterized several strains of L. monocytogenes of serotype 4b obtained from routine surveys of foods and food-processing environments. Focusing on genotypes at selected loci, it was found that a substantial fraction of these strains had Sau3AI-resistant genomic DNA and harbored additional, genomically unlinked genetic markers specific to epidemic clone I.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The serogroup 4 L. monocytogenes strains of food and environmental origin that were used in this study are listed in Table 1. With the exception of strain NCSK02-16 (provided to us by K. Boor, Cornell University) and NCSK02-3, (provided by L. Wolf, North Carolina Department of Health Epidemiology Laboratory), all other strains were from the strain collections of the Food and Drug Administration. The strains had been isolated during routine surveys of various foods and food-processing environments in the time period between 1985 and 2003, as indicated in Table 1, and were chosen for inclusion in this study based strictly on their serotype (determined to be serogroup 4 or, for certain strains, serotype 4b). In addition, we included as reference strains three known epidemic clone I strains (all of serotype 4b) that were isolated from foods implicated in outbreaks. These included NCLG-5 (cheese isolate, Swiss outbreak of 1983 to 1987), NCLG-8 (Mexican-style cheese isolate, California outbreak of 1985), and NCLG-11 (coleslaw isolate, Nova Scotia outbreak of 1981). These reference strains were provided to us by the Centers for Disease Control and Prevention. Growth and preservation of the bacteria were as previously described (26). Serotype designations were confirmed by reactivity with monoclonal antibodies highly specific for serotypes 4b, 4d, and 4e (25) and with Southern blotting employing probes derived from the gene gltA, found in all screened serotype 4b, 4d, and 4e strains but not in other serotypes of L. monocytogenes (26).

Isolation of genomic DNA and restriction enzyme analysis. The DNeasy tissue kit (QIAGEN, Valencia, Calif.) was used to extract and isolate DNA from *Listeria* cultures grown for 24 h according to the method supplied by the manufacturer. Determination of DNA resistance or sensitivity to Sau3AI or MboI digestion was performed as previously described (44). Restriction enzymes were purchased from New England Biolabs (Beverly, Mass.) or from Promega (Madison, Wis.) and were used according to the manufacturers' instructions.

PCR and Southern blots. Primers used in this study are listed in Table 2. Some primers were designed based on preliminary sequence data of L. monocytogenes serotype 4b, which were obtained from The Institute for Genomic Research through the website at http://www.tigr.org. Primers were designed using the software Primer3 (http://www.genome.wi.mit.edu/genome_software/other/primer3.html) and were purchased from Bio-synthesis, Inc. (Lewisville, Tex.) or from QIAGEN. PCR employed Ex Taq DNA polymerase (Fisher). The reaction mixtures were subjected to a hot start (95°C for 5 min) prior to 30 cycles of amplification (95°C for 1 min, 50°C for 1 min, 72°C for 2 min, and a final extension at 72°C for 20 min) in a Progene thermocycler. Amplified products were separated by electrophoresis through a 1.0% (wt/vol) agarose gel with 1x Tris-borate-EDTA running buffer. DNA probes for Southern blotting were obtained by using the PCR products amplified from genomic DNA of the serotype 4b strain F2381, implicated in the same outbreak as strain F2365, employed in the serotype 4b genome sequencing (www.tigr.org). The PCR products were excised from the gel (0.8% agarose gel with 1x Tris-borate-EDTA running buffer) and purified with a QIAquick gel extraction kit (QIAGEN). DNA probes were labeled with digoxigenin (Genius kit; Roche) following the instructions provided by the manufacturer. Southern blotting was performed as described previously (26).

Nucleotide sequence analysis. For DNA and protein database searches and analyses, we used FASTA (University of Wisconsin GCG package; Genetics Computer Group, Madison, Wis.), BLAST algorithms (http://www.ncbi.nlm.nih.gov/BLAST/), and ARTEMIS software (36; http://www.sanger.ac.uk/Software/Artemis). Preliminary sequence data for L. monocytogenes serotype 4b were obtained from The Institute for Genomic Research through the website at http://www.tigr.org. GAMOLA (Global Annotation of Multiplexed On-site Blasted DNA sequences) (www.cals.ncsu.edu:8050/food_science/KlaenhammerLab//GAMOLA) was used to automatically predict coding open reading frames and annotate the draft sequence in the genomic region encompassing DNA fragment 85 (18). GAMOLA relies on the available software Glimmer2 (10), the NCBI tool kit (National Center for Biotechnology Information, Bethesda, Md.), and HMMER2.2g (12) and combines the results into functionally annotated DNA sequences in GenBank format. The annotated sequences of the serotype 4b genome in the region encompassing DNA fragment 85 (18) were visualized by using ARTEMIS (36). Restriction map analysis of specific genomic regions employed the program Webcutter, version 2.0 (Max Heiman, Yale University, 1997), (http://www.firstmarket.com/cutter/cut2.html), and the coordinates were verified with ARTEMIS.

RESULTS

Unexpectedly high prevalence of strains with Sau3AI-resistant DNA among serotype 4b strains of food origin. Genomic DNAs of all strains listed in Table 1 were screened for sensitivity to digestion by the GATC-recognizing enzymes MboI and Sau3AI. Methylation of the cytosine at GATC sites inhibits cleavage by Sau3AI but has no effect on digestion by MboI, which, in contrast, is inhibited by methylation of the adenine at the same site (27). All DNAs could be readily digested by MboI. In contrast, digestion with Sau3AI revealed that a substantial portion of the isolates had DNA that resisted digestion (Table 1). Since the DNA of these isolates could be cut by MboI (Table 1), it can be concluded that the reason for the resistance to digestion by Sau3AI was not the quality of the DNA but instead the methylation of the cytosines at the GATC sites in the genome of these isolates. All of these Sau3AI-resistant isolates appeared to be typical serotype 4b strains on the basis of their reactivities with monoclonal antibodies c74.22 and c74.33, specific for serotype 4b, and the rare but closely related serotypes 4d and 4e (25). These strains were also positive by PCR and Southern blotting for the serotype-specific gene gtcA, which is specific to serogroup 4 (33), and for the serotype 4b-, 4d-, and 4e-specific genes gltA and gltB, involved in teichoic acid glycosylation (26). The three reference epidemic clone I strains that we employed (NCLG-5, NCLG-8, and NCLG-11, from food implicated in the Swiss, California, and Nova Scotia outbreaks of listeriosis, respectively) had DNA that could be readily digested by MboI but resisted digestion by Sau3AI, as previously described (44; data not shown).

Serotype 4b (or serogroup 4) strains found to have methylated (Sau3AI-resistant) DNA included isolates from various commodities and food-processing environments, including processed avocado, dairy products such as ice cream bars and soft cheese, and the dairy plant environment (4 of 11 screened strains); seafood such as frozen lobster, frozen crabmeat, and cooked shrimp (6 of the 10 screened strains); and other food products (beef franks and hummus) (Table 1).

Identification of an epidemic clone I-specific restriction-modification system (genes 85M, 85R, and 85S) in the genome of food-derived strains with Sau3AI-resistant DNA. In 2001, Herd and Kocks showed that DNA fragment 85 was specific to epidemic clone I and was internal to a gene with putative involvement in cytosine methylation at GATC sites (18). In the current study, the GAMOLA-based annotation of the serotype 4b genomic region encompassing fragment 85 suggested that this fragment was localized in contig 760 and was internal to a gene (designated 85M) with significant homology to genes encoding DNA methyltransferases. The deduced polypeptide product of 85M (471 amino acids) had 56% identity (over 434 amino acids) with GATC-specific methyltransferases of Streptococcus thermophilus (7) and Lactococcus lactis (42), as well as lower degrees of similarity with several other methyltransferases produced by other bacteria. The putative methyltransferase gene 85M was immediately upstream of, and transcribed in the same orientation as, a putative restriction endonuclease gene (designated 85R). Another gene with putative Sau3AI activity, designated 85S, was found adjacent to 85M and was transcribed in a divergent orientation (Fig. 1). The deduced polypeptide product of 85S (555 amino acids) had 34% identity (over its entire length) with the Sau3AI restriction endonuclease of Staphylococcus aureus (39) and 25% identity (over 495 amino acids) with the restriction component in the L. lactis LlaKR2I restriction-modification system (42) mentioned above. The deduced product of 85R (292 amino acids) had only a relatively small domain (ca. 90 amino acids) with similarity to restriction endonucleases produced by several other bacteria (identity values of 33 to 38%). The results from this analysis suggest that the genomes of epidemic clone I strains harbor a gene cassette (85M-85R-85S) with GATC-specific restriction-modification functions. Modification (methylation) genes are commonly in the immediate vicinity of genes encoding restriction enzymes with the same site specificity (cognate restriction endonucleases) (3). The G+C contents of the putative restriction-modification system genes 85M, 85R, and 85S were 33, 28, and 31%, respectively, noticeably lower than is typical for the L. monocytogenes genome (38%). The lower overall G+C content of the

85M-85R-85S genomic region is also readily detected in the ARTEMIS-based visualization of the region shown in Fig. 1.

All strains with Sau3AI-resistant DNA harbored the 85M sequences in their genomic DNA, yielding a single hybridizing fragment of ca. 5.8 kb in Southern blots of EcoRI-digested DNA (Fig. 2A). The size of this fragment was in agreement with that predicted on the basis of the genome sequence of strain F2365, which also revealed one EcoRI site in 85M, located near the 3' end of the coding sequence. The hybridizing band seemed to be of the same size in all isolates, suggesting lack of polymorphism in the EcoRI sites flanking the fragment. None of the serogroup 4 or serotype 4b strains with DNA that could be cut with Sau3AI yielded a signal with the 85M probe (Fig. 2A), even under conditions of low stringency (data not shown). These data suggest that only isolates with Sau3AI-resistant DNA harbored sequences that hybridized with the 85M sequences and that these sequences were conserved among isolates with Sau3AI-resistant DNA.

Southern blotting with the 85R and 85S probes showed that all strains that were 85M positive were also positive with 85R and 85S and, conversely, none of the 85M-negative isolates hybridized with either 85R or 85S (data not shown). The same hybridizing band was observed for all isolates with the 85M and 85R probes, suggesting that the probed fragments were on the same EcoRI fragment in the genome of these strains. This was also in agreement with the sequence analysis of this region, which indicated that these probes were on the same EcoRI genomic fragment in strain F2365. When 85S was used as the probe a band of a different size (ca. 2.7 kb), but identical for all strains, was produced (data not shown), again suggesting lack of polymorphisms in the corresponding EcoRI sites. The hybridization results for all strains with the 85M, 85R, and 85S probes are summarized in Table 1.

Serotype 4b strains with Sau3AI-resistant DNA also harbored additional epidemic clone I-specific genetic markers, unlinked on the serotype 4b genome. To determine whether the strains with Sau3AI-resistant DNA might also harbor additional genes typical of epidemic clone I, we searched for other sequences reported to be specific to this clone (18). PCR and Southern blots indicated that two of these fragments, 17B and 133, were indeed conserved among all known epidemic clone I strains that we screened, whereas they could not detected in the genomes of other strains of serotype 4b (data not shown). These fragments were unlinked with 85M-85R-85S and with each other on the genome of the epidemic clone I strain F2365 (distance in each case exceeding 170 kb).

The Southern blot results revealed that all food-derived serotype 4b strains that had Sau3AI-resistant DNA and hybridized with 85M, 85R, and 85S also hybridized with the 17B probe (Fig. 2B). Qualitatively identical results were obtained in Southern blots with the other epidemic clone I-specific probe, 133, except that the size of the hybridizing band (3.1 kb) was different from that obtained with the 17B probe (data not shown). Strains with DNA that could be cut with Sau3AI did not yield hybridizing bands with either 133 or 17B (Fig. 2B and Table 1).

In summary, these findings suggest that food-derived strains that had Sau3AI-resistant DNA harbored the putative restriction-modification system genes 85M, 85R, and 85S, as well as additional, genomically unlinked sequences known to be unique to epidemic clone I.

DISCUSSION

In this study, the examination of strains of serotype 4b isolated from routine surveys of food revealed a high incidence of isolates with DNA that appeared to be methylated at the cytosines of GATC sites and thus resisted digestion by Sau3AI, a phenotype associated with strains of a major clonal group (epidemic clone I) implicated in numerous outbreaks (18, 44). These strains harbored genes of a putative restriction-modification system (85M, 85R, and 85S), as well as other, genomically unlinked DNA sequences (such as 17B and 133) characteristic of epidemic clone I. All strains with this combination of genetic markers appeared to be typical serotype 4b isolates in terms of their recognition by the serotype-specific monoclonal antibodies c74.22 and c74.33. None had the distinctive c74.22-negative, c74.33-positive phenotype encountered among certain clinical isolates from outbreaks of listeriosis (8), supporting the earlier speculation that the c74.22-negative phenotype of those clinical isolates was selected in the course of the infection of the human host (8).

The putative involvement of the 85M-85S-85R gene cassette in methylation of cytosines at GATC sites and in a cognate restriction endonuclease activity (specific for GATC sites lacking cytosine methylation) is indicated by the similarities of the deduced polypeptides with GATC-specific methyltransferases and restriction enzymes from other bacteria. Further studies would be needed to clarify the role of each of these genes in restriction-modification of DNA.

The lower than average G+C content of this cassette suggests that the genes may have been introduced into the genome of epidemic clone I bacteria by lateral transfer from a currently unidentified source. Even though restriction-modification genes are widely disseminated among bacteria (3), our results indicate that the food-derived isolates with the Sau3AI-resistant DNA and the 85M-85R-85S gene cassette do not represent random acquisitions of these genes in genetic backgrounds otherwise unrelated to epidemic clone I. Besides the restriction-modification system genes, these isolates harbored additional, genomically unlinked genes specific for epidemic clone I, such as the genes corresponding to fragments 17B and 133, the functions of which are currently unknown. It remains to be determined whether the 85M-85R-85S restriction-modification system may have mediated the acquisition of some of these genes, as has been postulated for restriction-modification systems in other bacterial pathogens (14).

The G+C contents of fragments 17B and 133 are ca. 30%, noticeably lower than the average for the L. monocytogenes genome (38%), suggesting that these genes also were acquired by horizontal transfer from another source. However, even though fragment 133 is internal to a gene that is harbored by epidemic clone I and not other serotype 4b strains, it is also harbored by strains of serotype 1/2b and 3b (18; S. Kathariou and S. Kernodle, unpublished results). This distribution is of interest, since serotypes 1/2b, 4b, and 3b constitute a well-defined genomic division in L. monocytogenes (2, 23, 31). One can speculate that the gene harboring fragment 133 was acquired by a strain ancestral to that division by horizontal transfer and that it was lost by many strains of serotype 4b but maintained by epidemic clone I. These results suggest that the restriction-modification gene cassette (85M-85R-85S) was acquired by a lineage that already harbored some of the genes which are characteristic of epidemic clone I and absent from most other strains of serotype 4b. Regardless of the evolutionary pathways that led to the unique markers of epidemic clone I, our data indicate that these markers are stable genetic attributes of many serotype 4b strains isolated from routine surveys of foods.

In an earlier study of L. monocytogenes strains from animals, foods, and the environment epidemic clone I strains were found relatively frequently among animal isolates but were rarely isolated from foods (4). Epidemic clone I strains were also not identified in a survey of raw, refrigerated meat and poultry (37). However, these studies examined strains from foods regardless of serotype. In general, and for reasons which remain unclear, serotype 4b strains are not predominant in foods, which tend to be contaminated by serotype 1/2 strains, especially 1/2a and 1/2c (23). Our detection of strains with epidemic clone I genetic markers was likely facilitated by our inclusion of a sufficient number of serotype 4b isolates and by the employment of specific genetic markers. Although accurate estimates of the prevalence of such isolates among serotype 4b strains from processed and RTE foods in the United States and elsewhere remain unavailable, the fact that such strains were identified in foods of different types suggests that they may be widely distributed, possibly due to genetic attributes that confer a competitive edge over other strains of serotype 4b in foods or in the food-processing environment.

In bacteria, restriction-modification systems are thought to have evolved partly as a defense against phage invasion, bringing about digestion of unmodified (unmethylated) DNA injected into the cell by the phage (3). It is of interest that GATC sequences are underrepresented in phage genomes (22), possibly to counteract the production of GATC-cleaving endonucleases by potential hosts. Methylation at frequent sites in the genome (such as GATC sites) may also affect other molecular processes. In Salmonella, Dam methylation plays important roles in expression of virulence genes and impacts several key virulence attributes (16). In Escherichia coli, Dam methylation of adenines at GATC sites is important in the regulation of several genes (19, 40) and is impaired by the presence of methylated cytosines at these sites (27). The distribution of GATC sites is not random in the E. coli genome (17, 22, 30), and certain classes of GATC motifs may be involved in transcriptional responses during transitions to decreased temperature and increased oxygen (17, 30). Although similar studies on the distribution of GATC sites in the L. monocytogenes genome have not yet been reported, one may speculate that methylation of such sites in promoter regions might affect the accessibility of the latter to transcriptional activators and influence transcription.

Serotype 4b strains generally tend to be underrepresented in foods but are an important contributor to human listeriosis. It is conceivable that the overall low prevalence of serotype 4b in foods may be counterbalanced by the relatively high incidence of strains with the epidemic clone I markers that we observed in this study. Further studies will be needed to elucidate the genetic and virulence attributes of such strains and to adequately monitor their survival and growth in foods and food-processing environments.

In summary, our results suggest that when foods are contaminated by L. monocytogenes of serotype 4b, the bacteria have a substantial likelihood of harboring genetic markers specific to epidemic clone I, a clonal group implicated in numerous food-borne outbreaks of listeriosis. Employment of the genetic markers described here would greatly facilitate the identification of such serotype 4b strains and could be incorporated in risk assessments of the hazard posed by *Listeria* contamination of foods.

ACKNOWLEDGMENTS

This work was partially funded in the form of a National Alliance for Food Safety cooperative agreement between the USDA-Agricultural Research Service, North Carolina State University, and Cornell University, through the National Alliance for Food Safety, and by USDA NRI grant 2001-0969. Sequencing of the genome of L. monocytogenes serotype 4b by The Institute for Genomic Research was accomplished with support from the USDA. We acknowledge Irene Wesley (USDA-ARS, National Animal Disease Center, Ames, Iowa) for valuable feedback and support. We thank B. Swaminathan and L. M. Graves (Centers for Disease Control and Prevention) for the availability of the reference strains NCLG-5, NCLG-8, and NCLG-11. Kathryn Boor (Cornell University) and Leslie Wolf (North Carolina Department of Health) are thanked for strains NCSK02-16 and NCSK02-3, respectively. Preliminary sequence data of L. monocytogenes serotype 4b were obtained from The Institute for Genomic Research through the website at http://www.tigr.org. E. Lanwermeyer is thanked for editorial assistance and support.

REFERENCES

1. Avery, S. M., and S. Buncic. 1997. Differences in pathogenicity for chick embryos and growth kinetics at 37 degrees C between clinical and meat isolates of *Listeria monocytogenes* previously stored at 4 degrees C. Int. J. Food Microbiol. 34:319-327.

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

3. Bickle, T. A., and D. H. Kruger. 1993. Biology of DNA restriction. Microbiol. Rev. 57:434-450.

4. Boerlin, P., and J. C. Piffaretti. 1991. Typing of human, animal, food, and environmental isolates of *Listeria monocytogenes* by multilocus enzyme electrophoresis. Appl. Environ. Microbiol. 57:1624-1629.

5. Brosch, R., B. Catimel, G. Millon, C. Buchrieser, E. Vindel, and J. Rocourt. 1993. Virulence heterogeneity of *Listeria monocytogenes* strains from various sources (food, human, animal) in immunocompetent mice and its association with typing characteristics. J. Food Prot. 56:296-301.

6. Buchrieser, C., R. Brosch, B. Catimel, and J. Rocourt. 1993. Pulsed-field gel electrophoresis applied for comparing *Listeria monocytogenes* strains involved in outbreaks. Can. J. Microbiol. 39:395-401.

7. Burrus, V., C. Bontemps, B. Decaris, and G. Guedon. 2001. Characterization of a novel type II restriction-modification system, Sth368I, encoded by the integrative element ICESt1 of Streptococcus thermophilus CNRZ368. Appl. Environ. Microbiol. 67:1522-1528.

8. Clark, E. E., I. Wesley, F. Fiedler, N. Promadej, and S. Kathariou. 2000. Absence of serotype-specific surface antigen and altered teichoic acid glycosylation among epidemic-associated strains of *Listeria monocytogenes*. J. Clin. Microbiol. 38:3856-3859

9. Conner, D. E., V. N. Scott, S. S. Summer, and D. T. Bernard. 1989. Pathogenicity of foodborne, environmental and clinical isolates of *Listeria monocytogenes* in mice. J. Food Sci. 54:1553-1556.

10. Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg. 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. 27:4636-4641.

11. Del Corral, F., R. L. Buchanan, M. M. Bencivengo, and P. H. Cooke. 1990. Quantitative comparison of selected virulence associated characteristics in food and clinical isolates of *Listeria*. J. Food Prot. 53:1003-1009.

12. Eddy, S. R. 1998. Profile hidden Markov models. Bioinformatics 14:755-763

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

14. Fitzgerald, J. R., S. D. Reid, E. Ruotsalainen, T. J. Tripp, M. Y. Liu, R. Cole, P. Kuusela, P. M. Schlievert, A. Jarvinen, and J. M. Musser. 2003. Genome diversification in Staphylococcus aureus: molecular evolution of a highly variable chromosomal region encoding the staphylococcal exotoxin-like family of proteins. Infect. Immun. 71:2827-2838.

15. Glaser, P., L. Frangeul, C. Buchrieser, C. Rusniok, A. Amend, F. Baquero, et al. 2001. Comparative genomics of *Listeria* species. Science 294:849-852.

16. Heithoff, D. M., R. L. Sinsheimer, D. A. Low, and M. J. Mahan. 1999. An essential role for DNA adenine methylation in bacterial virulence. Science 284:967-970.

17. Henaut, A., T. Rouxel, A. Gleizes, I. Moszer, and A. Danchin. 1996. Uneven distribution of GATC motifs in the *Escherichia coli* chromosome, its plasmids and its phages. J. Mol. Biol. 257:574-585.

18. Herd, M., and C. Kocks. 2001. Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. Infect. Immun. 69:3972-3979.

19. Hernday, A., M. Krabbe, B. Braaten, and D. Low. 2002. Self-perpetuating epigenetic pili switches in bacteria. Proc. Natl. Acad. Sci. USA 99(Suppl. 4):16470-16476.

20. Hof, H., and J. Rocourt. 1992. Is any strain of *Listeria monocytogenes* detected in food a health risk? Int. J. Food Microbiol. 16:173-182.

21. Jacquet, C., E. Gouin, D. Jeannel, P. Cossart, and J. Rocourt. 2002. Expression of ActA, Ami, InlB, and listeriolysin O in *Listeria monocytogenes* of human and food origin. Appl. Environ. Microbiol. 68:616-622.

22. Karlin, S., I. Ladunga, and B. E. Blaisdell. 1994. Heterogeneity of genomes: measures and values. Proc. Natl. Acad. Sci. USA 91:12837-12841.

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

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

25. Kathariou, S., C. Mizumoto, R. D. Allen, A. K. Fok, and A. A. Benedict. 1994. Monoclonal antibodies with a high degree of specificity for *Listeria monocytogenes* serotype 4b. Appl. Environ. Microbiol. 60:3548-3552.

26. Lei, X. H., F. Fiedler, Z. Lan, and S. Kathariou. 2001. A novel serotype-specific gene cassette (*gltA-gltB*) is required for expression of teichoic acid-associated surface antigens in *Listeria monocytogenes* of serotype 4b. J. Bacteriol. 183:1133-1139.

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

28. Notermans, S., J. Dufrenne, T. Chakraborty, S. Steinmeyer, and G. Terplan. 1991. The chick embryo test agrees with the mouse bio-assay for assessment of the pathogenicity of *Listeria* species. Lett. Appl. Microbiol. 13:161-164.

29. Ørskov, F., and I. Ørskov. 1983. Summary of a workshop on the clone concept in epidemiology, taxonomy, and evolution of *Enterobacteriaceae* and other bacteria. J. Infect. Dis. 148:346-357.

30. Oshima, T., C. Wada, Y. Kawagoe, T. Ara, M. Maeda, Y. Masuda, S. Hiraga, and H. Mori. 2002. Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli*. Mol. Microbiol. 45:673-695.

31. Piffaretti, J. C., H. Kressebuch, M. Aeschbacher, J. Bille, E. Bannerman, J. M. Musser, R. K. Selander, and J. Rocourt. 1989. Genetic characterization of clones

of the bacterium *Listeria monocytogenes* causing epidemic disease. Proc. Natl. Acad. Sci. USA 86:3818-3822.

32. Pine, L., S. Kathariou, F. Quinn, V. George, J. D. Wenger, and R. E. Weaver. 1991. Cytopathogenic effects in enterocytelike Caco-2 cells differentiate virulent from avirulent *Listeria* strains. J. Clin. Microbiol. 29:990-996.

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

34. Ramnath, M., K. B. Rechninger, L. Jansch, J. W. Hastings, S. Knochel and A. Gravesen. Development of a *Listeria monocytogenes* EGDe partial proteome reference map and comparison with the protein profiles of food isolates. Appl. Environ. Microbiol. 69:3368-3376.

35. Rocourt, J., C. Jacquet, and J. Bille. 1997. Human listeriosis. Publication W.H.O./FNU/FOS/97.1. World Health Organization, Geneva, Switzerland.

36. Rutherford, K., J. Parkhill, J. Crook, T. Horsnell, P. Rice, M-A. Rajandream, and B. Barrell. 2000. Artemis: sequence visualization and annotation. Bioinformatics 16:944-945.

37. Ryser, E. T., S. M. Arimi, M. M.-C. Bunduki, and C. W. Donnelly. 1996. Recovery of different *Listeria* ribotypes from naturally contaminated, raw

refrigerated meat and poultry products with two primary enrichment media. Appl. Environ. Microbiol. 62:1781-1787.

38. Schuchat, A., B. Swaminathan, and C. V. Broome. 1991. Epidemiology of human listeriosis. Clin. Microbiol. Rev. 4:169-183.

39. Seeber, S., C. Kessler, and F. Gotz. 1990. Cloning, expression and characterization of the Sau3AI restriction and modification genes in *Staphylococcus carnosus* TM300. Gene 94:37-43.

40. Slominska, M., G. Konopa, J. Ostrowska, B. Kedzierska, G. Wegrzyn, and A. Wegrzyn. 2003. SeqA-mediated stimulation of a promoter activity by facilitating functions of a transcription activator. Mol. Microbiol. 47:1669-1679.

41. Tran, H. L., and S. Kathariou. 2002. Restriction fragment length polymorphisms detected with novel DNA probes differentiate among diverse lineages of serogroup 4 *Listeria monocytogenes* and identify four distinct lineages in serotype 4b. Appl. Environ. Microbiol. 68:59-64.

42. Twomey, D. P., L. L. McKay, and D. J. O'Sullivan. 1998. Molecular characterization of the *Lactococcus lactis* LlaKR2I restriction-modification system and effect of the IS982 element positioned between the restriction and modification genes. J. Bacteriol. 180:5844-5854.

43. Zheng, W., and S. Kathariou. 1995. Differentiation of epidemic-associated strains of *Listeria monocytogenes* by restriction fragment length polymorphism in a gene

region essential for growth at low temperatures (4°C). Appl. Environ. Microbiol. 61:4310-4314.

44. Zheng, W., and S. Kathariou. 1997. Host-mediated modification of Sau3AI restriction in *Listeria monocytogenes*: prevalence in epidemic-associated strains. Appl. Environ. Microbiol. 63:3085-3089.

Table 1. Food and environmental strains of *L. monocytogenes* serotype 4b used in this study

			Digestion by		Reactivity with			
Strain (serotype)	Yr isolated	Origin	Sau3AI	MboI	*gltA*	85 (M,S, and R)	133	17B
Sau3AI-resistant strains								
NCSK03-61 (4)[a]	2003	Cooked Dung. crab	−	+	+	+	+	+
NCSK01-13 (4)[a]	2001	Frozen breaded shrimp	−	+	+	+	+	+
NCSK01-32 (4)[a]	2001	Guacamole	−	+	+	+	+	+
NCSK01-45 (4)[a]	2001	Frozen avocado pulp	−	+	+	+	+	+
NCSK01-46 (4)[a]	2001	Frozen avocado pulp	−	+	+	+	+	+
NCSK01-58 (4)[a]	2001	Virginia ham	−	+	+	+	+	+
NCSK99-1 (4b)	1999	Processed avocado[b]	−	+	+	+	+	+
NCSK99-4 (4b)	1999	Processed avocado[b]	−	+	+	+	+	+
NCSK99-6 (4b)	1999	Processed avocado[b]	−	+	+	+	+	+
NCSK97-1 (4)[a]	1997	Hummus	−	+	+	+	+	+
NCSK92-3 (4)[a]	1992	Beef frankfurters	−	+	+	+	+	+
NCSK88-4 (4b)	1988	Crabmeat, frozen	−	+	+	+	+	+
NCSK88-5 (4b)	1988	Shrimp, cooked	−	+	+	+	+	+
NCSK88-12 (4b)	1988	Lobster, frozen	−	+	+	+	+	+
NCSK88-16 (4)[a]	1988	Red snapper	−	+	+	+	+	+
NCSK87-11 (4b)	1987	Soft cheese	−	+	+	+	+	+
NCSK86-1 (4b)	1986	Ice cream bar	−	+	+	+	+	+
NCSK86-9 (4b)	1986	Dairy, environment	−	+	+	+	+	+
NCSK86-13 (4b)	1986	Ice cream bars	−	+	+	+	+	+
NCSK03-77 (4)[a]	NK[c]	Food	−	+	+	+	+	+
Sau3AI-sensitive strains								
NCSK98-31 (4b)	1998	Hot dog	+	+	+	ND[d]	ND	−
NCSK87-3 (4)[a]	1987	Ice cream	+	+	+	−	−	−
NCSK88-2 (4)[a]	1988	Cheese	+	+	+	−	−	−
NCSK88-6 (4b)	1988	Shrimp plant environment	+	+	+	−	−	−
NCSK88-10 (4b)	1988	Blue cheese	+	+	+	−	−	−
NCSK88-17 (4b)	1988	Scallops	+	+	+	−	−	−
NCSK88-18 (4b)	1988	Crabmeat, cooked	+	+	+	−	−	−
NCSK87-8 (4)[a]	1987	Milk, pasteurized	+	+	+	−	−	−
NCSK87-16 (4b)	1987	Raw shrimp	+	+	+	−	−	−
NCSK86-7 (4)[a]	1986	Soft cheese	+	+	+	−	−	−
NCSK86-14 (4)[a]	1986	Dairy plant, environment	+	+	+	−	−	−
NCSK85-15 (4)[a]	1985	Ice cream plant	+	+	+	−	−	−
NCSK02-16 (4)[a]	NK	Food	+	+	+	−	−	−
NCSK02-3 (4b)	NK	Food	+	+	+	ND	−	−

[a] Strains are positive with *gltA* and serotype 4b-specific MAbs, hence likely to be of serotype 4b, 4d, or 4e.
[b] Products yielding each of these strains were from different producers.
[c] NK, exact year of isolation not known.
[d] ND, not determined.

Table 2. Primer sequences used in this study

Gene or fragment	Forward primer (5' to 3')	Reverse primer (5' to 3')
85M [1]	AATATATTTTCAATGTTTGATGGT	GCTAATTCAATCCCTATTCT
85R[2]	CTGATGATGCTGGAAAAGCA	TGATGTCCTTCCACATAAGGC
85S[2]	ACCACGGGCATACTCGTTTA	GGGCCTCAAATGAAGAACAT
gltA[3]	TCATCGTATCGCTTCTGTG	GTGCCATTACTACAGGTGCA
gltB[3]	TTCAAATAAGCCGTTCCAAA	AAAGGCAGAGATTACGCA
gtcA[4]	TGGGTTACTACAAGAAGAG	AGTACTGATGCGATAAAAGCA
133[5]	TCAACATCAGCGTAATTATGAG	AGTTCTTGTAGGTGTTTTACAG
17B[6]	TCCCTAACGTATTTTTTCAAATC	TGAGTTAGGTAATTTTATAAGTTG

[1]Accession no. AJ410373
[2] Constructed on the basis of based on preliminary sequence data of *L. monocytogenes* serotype 4b, which were obtained from the Institute for Genomic Research through the website at http://www.tigr.org
[3]Accession no. AFO33015
[4]Accession no. AFO72894
[5]Accession no. J410381
[6] Accession no. AJ410357

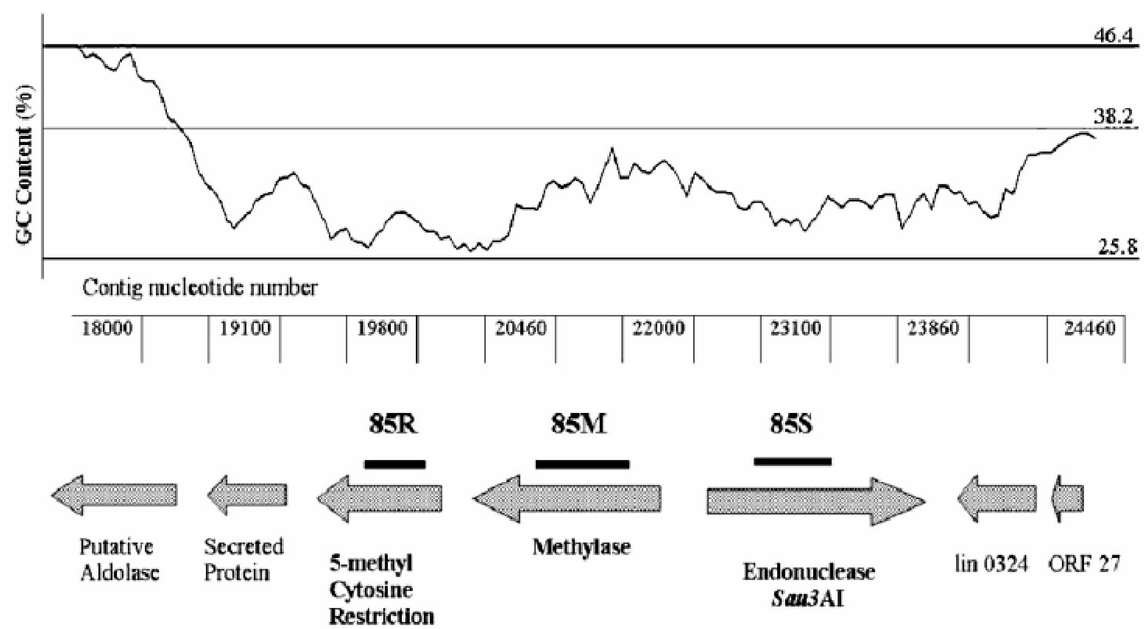


Fig.1. Restriction-modification system genes in the *L. monocytogenes* serotype 4b genome. Annotation and visualization of the region were done using GAMOLA and ARTEMIS as described in Materials and Methods. Putative open reading frames designated 5-methyl cytosine restriction, methylase, and putative Sau3AI correspond to 85R, 85M, and 85S, respectively, described in the text. Black bars indicate probes used for Southern blot-based detection of 85R, 85M, and 85S. Arrows indicate direction of transcription. The deduced G+C contents of the genes are shown at the top.

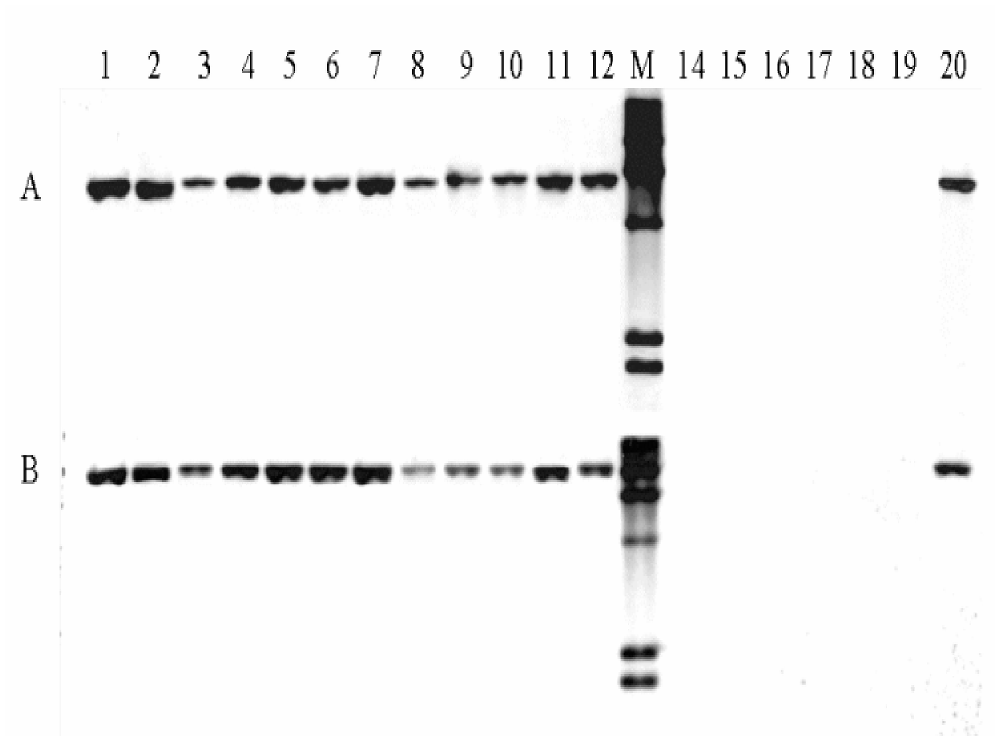


Fig.2. Southern blotting with a digoxigenin-labeled probe internal to the putative methylase 85M (A) and with epidemic clone I-specific probe

-

CHAPTER IV

CLONAL ASSOCIATION BETWEEN *LISTERIA MONOCYTOGENES* SEROTYPE 4B EPIDEMIC CLONE-I AND SEROTYPE 4B STRAINS IN NATURAL POPULATIONS ISOLATED FROM READY-TO-EAT FOODS (RTE)

ABSTRACT

Listeria monocytogenes is a foodborne bacterial pathogen associated with significant morbidity and mortality in humans and animals. Over the past two decades epidemic clones (EC) within *L. monocytogenes* serotype 4b population (ECI and ECII) have caused large outbreaks which resulted in heavy public health burden and substantial economic losses. However, characterization of strains harboring genetic markers typical of ECs but with no epidemiological links to outbreaks has not been adequately addressed. Here we report on the clonal association of eighteen strains with no epidemiological links and recovered from foods and food environments, which were previously reported to harbor ECI specific genetic markers (Yildirim et al., 2004; AEM, 70: 7581). Sixty seven serotype 4b isolates collected from food and food environments were characterized. Pulsed Field Gel Electrophoresis showed that ECI clinical strains and ECI food strains harboring ECI-unique marker genes were clustered together and separated from other serotype 4b strains without the markers. All eighteen strains shared either genetic identity or high genetic relatedness with ECI clinical strains. Further, hybridizations of genomic DNA of ECI food strains with probes derived from strain-specific genes of the ECI clinical strain (F2365) implicated in the 1985 California Jalisco cheese outbreak showed that the strain specific genes of F2365 were present in the genome of ECI food strains, while absent in other serotype 4b strains. The ECI food strains showed adaptive resistance to quaternary ammonium compounds (benzalkonium) and 12 isolates (67%) grew well in the presence of inhibitory levels of cadmium, while 8 isolates (44%) showed resistance to both arsenic and cadmium indicating multiple resistance mechanism. The

presence of such ECI-like clinical strains in food and food environments with such resistance attributes highlights the genomic stability and dynamic nature of this clone.

INTRODUCTION

Listeria monocytogenes is a foodborne bacterial pathogen that can be isolated from various environmental niches such as soil, silage, decaying plants and vegetation, and the intestine of healthy humans and animals (Vazquez-Boland et al, 2001). Because of its notorious adaptation capacity, the bacterium is capable of surviving in adverse food processing plant environments and may grow in cold stored ready-to-eat foods. Even though most *L. monocytogenes* infections are thought to be sporadic (Pinner et al., 1992; Schuchat et al., 1992) more than 30 large outbreaks have occurred in Europe and North America ever since listeriosis was established to be food-borne in 1981 (Swaminathan et al., 2001; FDA risk assessment, 2003). This is illustrated by multi-state outbreaks that occurred in the United States (August 1998- March, 1999 and 2002) which were due to the consumption of contaminated hotdogs and turkey deli meats. The implicated bacteria were new genetic variants of serotype 4b strains, named epidemic clone II (ECII). These outbreaks were associated with a heavy public health burden (28 deaths in total) and severe economic consequences (>300 million dollars) to the food industry (MMWR, CDC). Thus, there is a clear need for continuous screening and monitoring of the epidemic clones so that emergence and spread of new clones with increased fitness, antimicrobial resistance, and enhanced pathogenicity, as well as their associations with certain environmental niches, can be predicted. Nevertheless, the identification of *L. monocytogenes* epidemic clones is based on a limited number of genetic markers and characterization of strains harboring epidemic clone specific markers in natural populations have not been adequately studied.

Here, we report on genetic relatedness of strains harboring ECI-specific genetic markers and isolated from foods with no known involvement in listeriosis outbreaks (Yildirim et al., 2004). We show that such strains are indeed members of the same clonal group as epidemic-associated ECI strains.

MATERIALS AND METHODS

Details of the strain collection and handling were previously described (Yildirim et al., 2004). The same handling procedure was followed for the additional strains chracterized. A summary of strain information is given in Table 1. The MIC for cadmium resistance was determined as described in McLauchlin et al., 1997. Briefly, cultures were grown in nutrient broth overnight at 37 C and 3 μl of cultures were inoculated onto iso-sensitest agar plates (Oxoid, UK) containing 75 μl of cadmium chloride monohydrate (Fisher, PA) using a multichannel pipette (eppendorf, Westbury, NY). A control plate with no cadmium was used to compare growth. The procedure for testing resistance to benzalkonium (BC) was described elsewhere (Soumet et al., 2005). In this study strains were grown only at one level (10 $\mu g ml^{-1}$) at the highest end of benzalkonium testing range described in the procedure.

Pulsed-Field Gel Electrophoresis (PFGE) was performed by following the standardized procedure of Graves and Swaminathan (2001) and plugs were digested with AscI and ApaI restriction enzymes (Promega, Madison, WI) . The PFGE patterns were analyzed by using Bionumerics software version 3.5 (Applied Maths, Inc., TX) and the patterns were clustered based on Unweighted Pair Group Method with Arithmetic Mean (UPGMA) method and Dice coefficient.

DNA hybridizations with strain-specific probes. DNA array hybridizations were performed as described previously (Herd and Kocks, 2001). Primers targeting strain-specific genes of *L. monocytogenes* serotype 4b strains F2365 and H7550 (Nelson et al., 2004) were designed using Primer3 (Rozen et al., 2000) by targeting approximately 500 bp amplicons. The PCR products were amplified on a 96-well microtiter plate in the thermocyler (Biometra, Germany) using the following conditions: 5 min-95 C, 1 min-95 C, 1 min-50 C, 2 min-72 C, and a final 10 min extension at the end of 35 cycles. PCR employed Ex Taq DNA polymerase (Fisher). The amplified products were then purified using 96-well PCR purification kit (Fisher) before denaturing and suspending in DMSO (1:1 ratio). The PCR products were transferred onto nylon membrane using Multi-Blot replicator (VP408, V&P Scientific, Inc., San Diego, CA). Approximately 0.5 g of genomic DNAs were digested with EcoRI and purified on a QIAGEN (Valencia, CA) column. The digested DNA was labeled using the digoxigenin-labelling system following the manufacturer's recommendations (Genius kit; Roche, Indianapolis, IN). Hybridizations were carried out at high stringency conditions as described (Yildirim et al., 2004).

RESULTS

A total of 67 isolates collected in the period from 1985 to 2006 were characterized in this study (Table 1). The isolates were randomly chosen based solely on their serotype (serotype 4b). Isolates that caused major outbreaks and isolates with ECI markers that caused sporadic cases were also included in the study for comparison. We detected 18 isolates harboring epidemic clone-I specific genetic markers, including the previously reported isolates (Yildirim et al., 2004).

Even though heavy metal resistance was previously reported to be uncommon among serotype 4b strains (Lebrun et al., 1992; McLauchlin et al., 1997) we unexpectedly found that the growth of 12 ECI food strains (66%) with ECI genetic markers were not inhibited in the presence of 75 μg/ml cadmium in the medium. Interestingly, eight ECI clones showed resistance to both arsenic and cadmium, suggesting the possibility of multiple mechanism acquired by these strains (Table 5). Further, all of these strains were resistant to quaternary ammonium compounds (benzalkonium chloride) albeit some showed weak growth (adaptive). This observation suggests adaptive evolution of serotype 4b population colonizing food environments.

Clonal association between non-clinical isolates with ECI genetic markers and confirmed epidemic-associated strains. The results of PFGE banding patterns were interpreted in accordance with the criteria described in Tenover et al, 1995. All eighteen isolates harboring ECI specific markers were analyzed by PFGE with AscI and ApaI, along with ECI strains that caused outbreaks in California (1985), Switzerland (1983-1987) and Nova Scotia (1981). These latter two strains were chosen for comparison because they represent two variants within the clonal complex of epidemic clone I strains. Table 2 lists the number of shared fragments in PFGE patterns of the eighteen non-clinical harboring ECI specific genetic markers, obtained with AscI and ApaI. According to patterns obtained with both enzymes, four isolates consistently produced unique pulsotypes, even though they harbored ECI specific markers. The two reference strains chosen for comparison differed from each other only by two bands using restriction enzyme ApaI, whereas four different bands were observed in the AscI pattern. Five isolates showed pulsotypes with AscI indistinguishable from the Nova Scotia outbreak strain, while three of these isolates showed a two-band difference in ApaI patterns, suggesting strong genome stability of Nova Scotia clone for over two decades since it caused an outbreak in 1981. On the other hand, only one strain showed a pulsotype indistinguishable from that of the Swiss outbreak strain.

Interestingly, from the dendogram analysis employing Dice coefficient and AscI patterns (Fig.1) we observed that the pulsotypes of all food isolates harboring the genetic markers clustered together, clearly separating them from other serotype 4b strains without the markers. Similarly, ApaI patterns of the food and clinical isolates were clustered together, separating them from other serotype 4b isolates without ECI specific genetic markers (Fig. 2). Consistent with the observations in Table 2 three non-clinical strains with ECI markers formed a subcluster, suggesting they may represent clonal expansion of the ECI.

In addition to the PFGE data presented here, genomic DNA of the eighteen non-clinical isolates positive with ECI markers were hybridized with probes derived from genes specific to the ECI strain F2365, and to the ECII strain H7858. These strain-specific genes were identified in a comparative genomic analysis (Nelson et al., 2004). Consistent with the PFGE data, these hybridizations showed that F2365 strain- specific genes were present in the genome of all eighteen strains, and absent in other serotype 4b strains unrelated to ECI clones with few exceptions. These exceptions involved F2365 strain-specific genes that were actually found to be present in the genome of a few other serotype 4b strains, and thus apparently were not ECI-specific. Probes that hybridized only with ECI clones are listed in Table 3. In addition, the hybridization patterns with few exceptions confirmed the comparative genome analysis in that ECII strain (H7858) specific genes are absent in the genome of ECI food strains and the strain specific probes clearly distinguish between the strains that are descendants of two epidemic clones (Table4).

DISCUSSION

Numerous listeriosis outbreaks in Europe and North America have been caused by closely related clones, designated as epidemic clone I, for over two decades (Kathariou, 2002). Even though clinical isolates of these clones were included in several studies, characterization of strains that are related to epidemic clone I but are isolated from foods and food environments with no known implication in illness has not been adequately addressed. Thus, the data presented here add to our understanding of the population structure and relevant phenotypic characteristics of ECI-related strains.

PFGE analysis and DNA-DNA hybridizations suggested that the eighteen isolates from food and food environments had a clonal association to ECI clinical strains. The presence of genes specific to F2365 (Nelson et al., 2004) in the genome of ECI related non-clinical isolates and the absence of them in the genome of other serotype 4b strains unrelated to ECI indicates that these genes are in linkage disequilibrium (LD) with F2365 specific genes. Such LD in bacterial pathogenic strains can suggest that the organisms are descendants of a successful clone, in some cases selected by the host's immune system (Gupta and Maiden, 2001). Even though a parametric method to calculate LD, such as described in Awadalla, 2003, was not employed in this study, the fact that the strain-specific genes were absent in other serotype 4b food-derived strains, devoid of the ECI-specific genetic markers, suggests non-random association of the strain specific genes. Furthermore, the close genetic relationship of these strains with ECI clinical strains was demonstrated by PFGE analysis, through which the two groups were clustered together. Taken together, our data strongly suggest the clonal relationship of the eighteen isolates described above with ECI clinical strains.

ECI strains were previously reported to be heterogeneous despite clonality (Buchrieser 1993; Brosch et al., 1996). In addition, ECI and ECII-related strains originating from food and food environments showed diversity in terms of pulsotypes (Eifert et al., 2005). In parallel to these observations we detected several PFGE profile variants in food strains harboring ECI genetic markers, suggesting a range of genetic relatedness with respect to the confirmed epidemic strains. The PFGE profiles ranged from being indistinguishable to those of epidemic strains, to virtually unrelated pulsotypes despite harboring ECI strain specific genes. One can speculate that this may be the result of clonal expansion of the ECI strains, while ECII strains, which acquired an 80 kb conjugative plasmid, may be the clonal replacement of ECI. Thus, we hypothesize that strains derived from epidemic clones in the serotype 4b population are evolving into new variants which are capable of better colonizing the food and food environments.

A pertinent question would be whether or not these strains compare with the clinical strains in terms of their capacity to cause human illness, and foodborne outbreaks. Further studies are needed to elucidate pathogenic potential of the strains with one important caveat that immune selection does not necessarily promote the emergence of more virulent variants. In fact, some clones may represent 'escape variants' (Zhu et al., 2001), i.e. variants with increased fitness to escape from the host immune attack but reduced virulence. By contrast, they may steadily evolve into a 'superbug' as in the example of methycilline resistant S*taphylococcus areus* (MRSA), that is, clones that can better colonize food processing plants or retail outlets and acquire genetic elements conferring to the clone increased virulence. Consequently, these clones must be constantly monitored and the probes that we used in this study could be included in detection and epidemiological studies.

The emergence of antibiotic and/or antimicrobial resistant variants and the appearance of variants which can effectively avoid immune surveillance are not uncommon in clonal populations of bacterial pathogens. Nevertheless, understanding the dynamics of these populations will facilitate the design of effective public health interventions. Thus, it will be interesting to further investigate the extent to which the strains investigated here have diversified their genomes, employing comparative genomic hybridizations based on total genome sequencing.

REFERENCES

1. Awadalla P., 2003. The evolutionary genomics of pathogen recombination. Nat Rev Genet. 4:50-60

2. Boerlin P, Piffaretti JC., 1991. Typing of human, animal, food, and environmental isolates of *Listeria monocytogenes* by multilocus enzyme electrophoresis. Appl Environ Microbiol. 57:1624-9

3. Brosch R, Brett M, Catimel B, Luchansky JB, Ojeniyi B, Rocourt J. 1996. Genomic fingerprinting of 80 strains from the WHO multicenter international typing study of *Listeria monocytogenes* via pulsed-field gel electrophoresis (PFGE). Int J. Food Microbiol. 32:343-55

4. Buchrieser C, Brosch R, Catimel B, Rocourt J., 1993. Pulsed-field gel electrophoresis applied for comparing *Listeria monocytogenes* strains involved in outbreaks. Can J Microbiol. 39:395-401

5. Eifert JD, Curtis PA, Bazaco MC, Meinersmann RJ, Berrang ME, Kernodle S, Stam C, Jaykus LA, Kathariou S., 2005. Molecular characterization of *Listeria monocytogenes* of the serotype 4b complex (4b, 4d, 4e) from two turkey processing plants. Foodborne Pathog Dis. 2:192-200

6. FDA risk assessment, 2003. www.cfsan.fda.gov

7. Gupta S, Maiden MC., 2001. Exploring the evolution of diversity in pathogen populations. Trends Microbiol. 9:181-5

8. Herd M, Kocks C., 2001; Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. Infect Immun. 69:3972-9

9. Kathariou S., 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J Food Prot. 65:1811-29. Review.

10. McLauchlin J, Hampton MD, Shah S, Threlfall EJ, Wieneke AA, Curtis GD., 1997. Subtyping of *Listeria monocytogenes* on the basis of plasmid profiles and arsenic and cadmium susceptibility. J Appl Microbiol. 83:381-8

11. Nelson KE, Fouts DE et al., 2004. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. Nucleic Acids Res. 2004. 32:2386-95

12. Okwumabua O, O'Connor M, Shull E, Strelow K, Hamacher M, Kurzynski T, Warshauer D., 2005. Characterization of *Listeria monocytogenes* isolates from food animal clinical cases: PFGE pattern similarity to strains from human listeriosis cases. FEMS Microbiol Lett. 249:275-81

13. Pinner, R. W., A. Schuchat, B. Swaminathan, P. S. Hayes, K. D. Deaver,R. E. Weaver, B. D. Plikaytis, M. Reeves, C. V. Broome, and J. D. Wenger.1992. Role of foods in sporadic listeriosis II. Microbiologic and epidemiologic investigation. JAMA 267:2046–2050

14. Rozen, S. and Skaletsky, HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386

15. Schuchat, A., K. A. Deaver, J. D. Wenger, B. D. Plikaytis, L. Mascola, R. W. Pinner, A. L. Reingold, C. V. Broome, and the *Listeria* Study Group. 1992. Role of foods in sporadic listeriosis. I. Case-control study of dietary risk factors. JAMA 267:2041–2045.

16. Soumet C, Ragimbeau C, Maris P., 2005. Screening of benzalkonium chloride resistance in *Listeria monocytogenes* strains isolated during cold smoked fish production. Lett Appl Microbiol. 41:291-6

17. Swaminathan, B., T. J. Barrett, S. B. Hunter, and R. V. Tauxe. 2001. PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. Emerg. Infect. Dis. 7:382–389.

18. Tenover, F. C., R. D. Arbeit, R. V. Goering, P. A. Mickelsen, B. E. Murray, D. H. Persing, and B. Swaminathan. 1995. Interpreting chromosomal DNA restriction

patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33:2233-9.

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

20. Yildirim S, Lin W, Hitchins AD, Jaykus LA, Altermann E, Klaenhammer TR, Kathariou S. 2004. Epidemic clone I-specific genetic markers in strains of *Listeria monocytogenes* serotype 4b from foods. Appl Environ Microbiol. 70:4158-64

21. Zhu P, van der Ende A, Falush D, Brieske N, Morelli G, Linz B, Popovic T, Schuurman IG, Adegbola RA, Zurth K, Gagneux S, Platonov AE, Riou JY, Caugant DA, Nicolas P, Achtman M., 2001. Fit genotypes and escape variants of subgroup III *Neisseria meningitidis* during three pandemics of epidemic meningitis. Proc Natl Acad Sci U S A. 98:5234-9

Table 1. Distribution of 67 isolates of serotype 4b by ECI markers

Year	No. of Isolates (n =67)	ECI and related strains from sporadic and epidemic outbreaks	Sau3AI Resistant+ (from foods and food environments (n =18)	Source for the Strains with ECI Genetic Markers
1985-1995	17	7	9	Ice cream bar (2), soft cheese (1), dairy environment (1); crabmeat (1), shrimp (1), lobster (1), red snapper (1), smoked salmon (1)
1996-2006	50	8	9	Avocado pulp (3), guacamole (1), humus (1), environmental swab (1), turkey drumstick wash water (2), unknown (1)

+The genome of ECI and the genetically related strains is methylated at GATC sites (Yildirim et al., 2004)

Table 2. Comparison of pulsotypes of eighteen ECI-like isolates with reference strains

Restriction Enzyme	Swiss Isolate (Number of different fragments)			Nova Scotia Isolate (Number of different fragments)		
	(≤3)	(4-6)	(≥7)	(≤3)	(4-6)	(≥7)
AscI	6	8	4	8	4	4
ApaI	11	3	4	10	5	3

Table 3. EC-I specific genes absent in other serotype 4b strains

ORF ID	Putative Function	GC%
LMOf2365_0082	Conserved hypothetical protein	43.2
LMOf2365_0687	Conserved hypothetical protein	26.2
LMOf2365_2347	Conserved domain protein	26.7
LMOf2365_2630	Putative membrane protein	34.0
LMOf2365_2701	Conserved hypothetical protein	28.0
LMOf2365_2703	Conserved hypothetical protein	27.3
LMOf2365_2705	Hypothetical protein	34.3
LMOf2365_2707	Putative membrane protein	32.4
LMOf2365_2798	Hypothetical protein	28.6
LMOf2365_2799	DNA-binding protein	30.3
LMOf2365_2800	Site specific recombinase	29.9
LMOf2365_0325	Type II restriction enzyme (Sau3AI)	30.4
LMOf2365_0326	DNA binding protein	32.9
LMOf2365_0327	C-5 Cytosine specific methylase	33.0
LMOf2365_0328	Conserved domain protein	27.8

Table 4. EC-II* specific genes absent in other serotype 4b strains

ORF ID	Putative Function	GC%
LMOh7858_0076	Conserved hypothetical protein	43.2
LMOh7858_0865	Hypothetical protein	28.6
LMOh7858_0866	Transcriptional regulator	30.3
LMOh7858_0867	RNA-directed DNA polymerase from retron	25.7
LMOh7858_1174	Conserved hypothetical protein	30.0
LMOh7858_1169	Hypothetical protein	31.7
LMOh7858_1168	AAA superfamily ATPase	30.8
LMOh7858_1167	Rlx like protein	32.8
LMOh7858_1163	Phage integrase	34.2
LMOh7858_1161	HTH-domain protein	30.6
LMOh7858_2445	Phage conserved hypothetical protein	39.6
LMOh7858_2474	Putative prophage lambdaW5	31.3
LMOh7858_2753	Hypothetical protein	31.7
LMOh7858_2754	DNA methyltransferase	31.7
LMOh7858_2756	Conserved hypothetical protein	36.9
LMOh7858_2759	Hypothetical protein	32.8
LMOh7858_2762	Hypothetical protein	34.9
LMOh7858_2764	Phage integrase	34.1

*All plasmid (pLM80) genes except pLM80_0082 (transposase) and pLM80_0086 (cadmium efflux pump) are not present in ECI

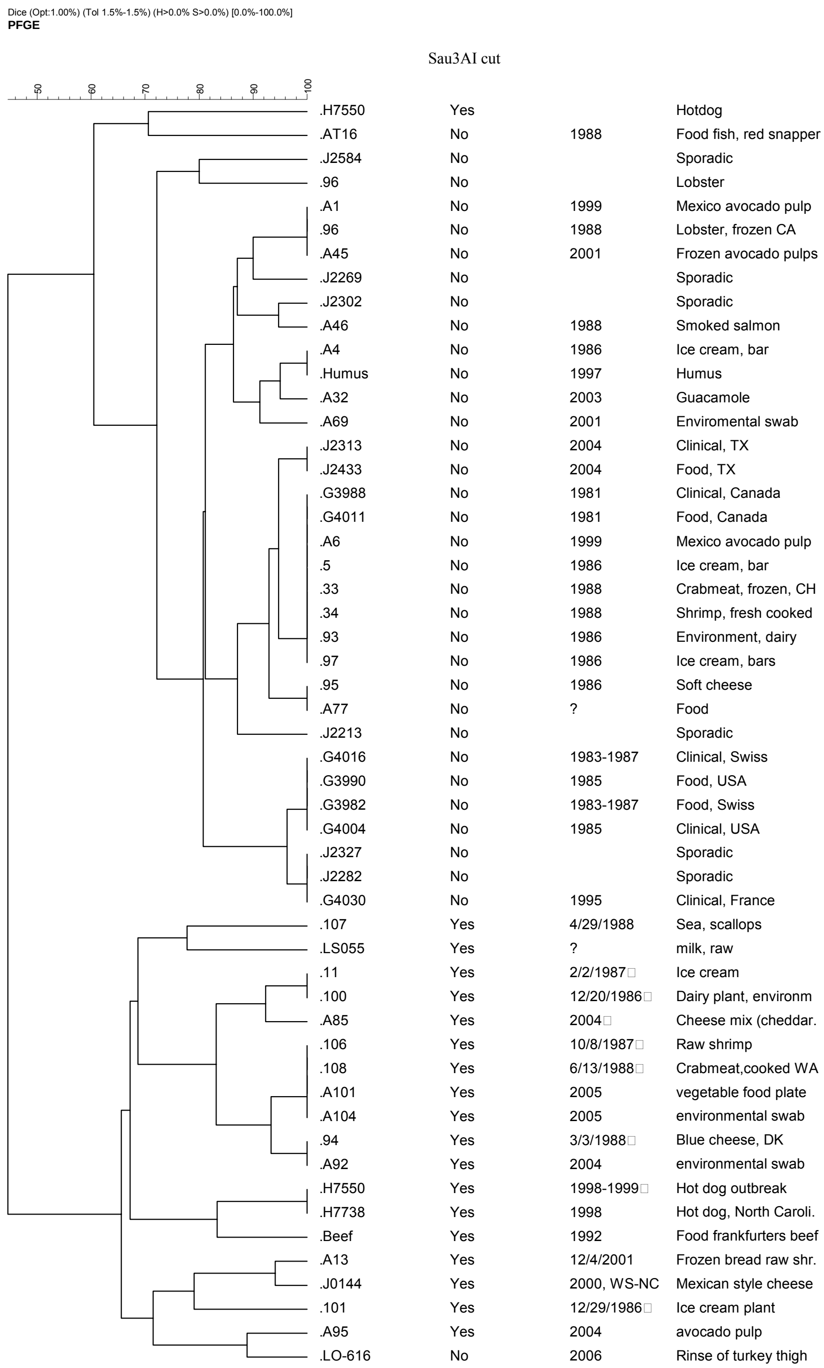


Fig. 1. Dendogram obtained with AscI pattern

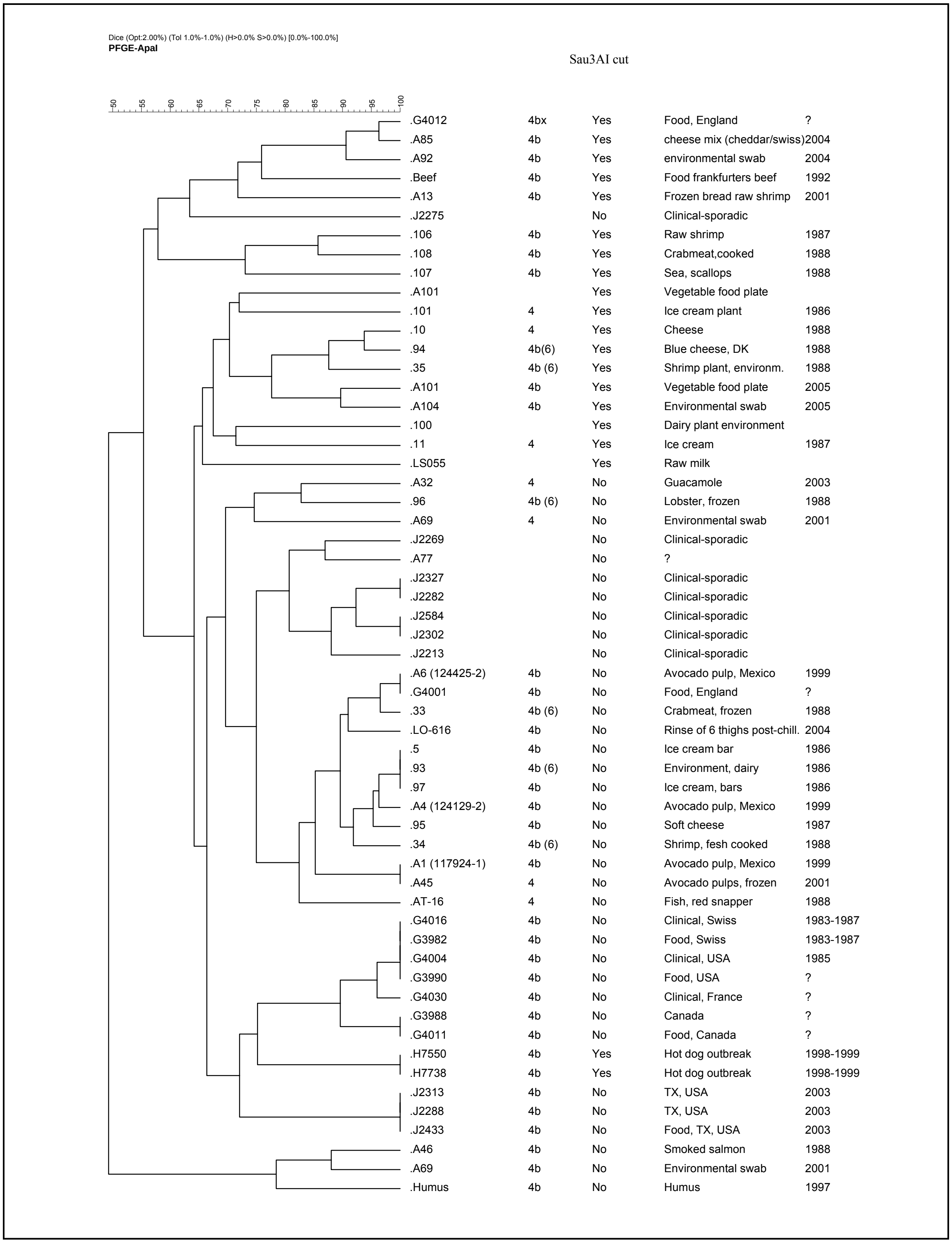


Fig. 2. Dendogram obtained with ApaI pattern

Table 5. Heavy metal resistance and transposons associated with resistance phenotype

Year	Strain ID	Source	Serotype	Sau3AI cut*	Arsenite Resistance	Cadmium Resistance	Probe cadA	PCR cadA	Transposase. pLM80-0086	cadA (Tn5422 associated)
1999	A1	Mexico avocado pulp	4b	No	+	+	−	−	−	−
1999	A4	Mexico avocado pulp	4b	No	−	+	−	−	−	+
1999	A6	Mexico avocado pulp	4b	No	−	−	−	−	−	−
2001	A45	Frozen avocado pulps	4b	No	+	+	NP	−	−	−
2003	A32	Guacamole	4b	No	−	−	−	−	−	−
1986	5	Ice cream, bar	4b	No	−	+	NP	+	+	−
1986	93	Environment, dairy	4b (6)	No	−	+	+	+	+	−
1986	97	Ice cream, bars	4b	No	−	+	NP	+	+	−
1986	95	Soft cheese	4b	No	−	−	−	−	−	−
1988	33	Crabmeat, frozen, CH	4b(6)	No	−	−	−	−	−	−
1988	34	Shrimp, fresh cooked	4b(6)	No	+	+	−	−	−	−
1988	96	Lobster, frozen CA	4b(6)	No	+	+	NP	−	−	−
1988	A46	Smoked salmon	4b	No	+	+	−	−	−	−
1988	AT-16	Food fish, red snapper	4b	No	+	+	NP	+	+	−
2001	A77	Unknown food source	4b	No	−	−	NP	−	−	−
1997	Hummus	Food hummus	4b	No	+	+	−	−	−	−
2001	A69	Environmental swab	4b	No	+	+	NP	+	+	−
1992	Beef	Frankfurters beef	4b	Yes	−	−	−	−	−	−
1988	107	Sea, scallops	4b	Yes	−	−	−	−	−	−
2001	A13	Frozen bread raw shrimp	4b	Yes	−	+	+	+	+	−
1987	11	Ice cream	4b	Yes	+	+	−	−	−	−
1988	35	Shrimp plant, environment	4b	Yes	+	+	−	−	−	−
2005	A104	Environmental swab	4b	Yes	+	+	−	−	−	−

NP –Not probed

*ECI like strains genomic DNA does not cut with Sau3AI

CHAPTER V

IDENTIFICATION AND CHARACTERIZATION OF A TYPE-II RESTRICTION-MODIFICATION SYSTEM SPECIFIC TO *LISTERIA MONOCYTOGENES* EPIDEMIC CLONE SEROTYPE 4B IN THE GENOMES OF NON-4B STRAINS ISOLATED FROM FOODS AND FOOD ENVIRONMENTS.

ABSTRACT

Listeria monocytogenes is a foodborne pathogen which has a clonal population structure with distinct phylogenetic lineages, and gene flow between the strains of different lineages is limited. Restriction-modification systems in *Listeria* genomes have been identified before only in the genome of a clonal group of strains of serotype 4b. Specifically, strains of epidemic clone I harbor a type-II restriction-modification system with specificity for 5'-GATC-3' sites, and which results in methylation of cytosines at these sites. In this study, we describe the presence of a homologous gene cassette in the chromosome of certain strains of serotype 1/2a and 1/2b. The lack of signals in DNA-DNA hybridizations with other epidemic strain-specific probes suggested that only the restriction-modification gene cassette was acquired by these strains. The gene cassette was found to have been transferred *en bloc* downstream of a highly conserved putative PTS operon. The position was in the same genomic location in these strains and in epidemic clone I. The cassette was completely sequenced in one strain (A7) of serotype 1/2a, and pairwise alignment showed that the sequence of the cassette was remarkably conserved at the nucleotide level (97%) between the two genomes. The average G+C content of the cassette (31%) and the use of codons hinted the possibility of horizontal transfer. Genetic organization analysis indicated that the gene annotated in the genome of ECI strain caused Jalisco cheese outbreak in California (F2365) as a pseudogene was part of the cassette and the stop codon of this gene was bracketed by a 17-bp perfect inverted repeat. The presence of a putative recombinase gene downstream of the methylase gene suggested that the gene cassette was inserted via site-specific recombination. The

complete deletion of the cassette from the genome of A7 and F2365 rendered the chromosomes of both strains sensitive to Sau3AI digestion. Pulsed field gel electrophoresis patterns of the serotype 1/2a strains harboring the gene cassette indicated clonality of the strains. Strains of serotype 1/2a with alternative methylase(s) with the same target specificity were also identified.

INTRODUCTION

Restriction–modification systems consist of two or more enzymes, namely, a restriction endonuclease that recognizes a specific DNA sequence and hydrolyzes phosphodiester bonds making up the backbone of DNA, and a cognate methyltransferase that can methylate the same sequence and thereby render it resistant to the cleavage. There are three main types (I, II, and III) of restriction enzymes, which have distinct characteristics in terms of cofactor requirements, subunit composition, and mode of action (Wilson and Murray, 1991).

Traditionally, it has been thought that the primary biological function of restriction-modification genes is to protect the host genome against invasion of phage DNA (Bickle et al., 1993). On the other hand, the genetic barriers provided by RM systems are limited and short-lived due to mechanisms phages have developed to circumvent them (Matic et al., 1996). Therefore, alternative hypotheses have been formulated recently with regard to the presence and maintenance of RM systems in bacterial genomes. RM systems can modulate the frequency of gene exchange by limiting acquisition of foreign DNA. Conversely, they can also facilitate integration of foreign DNA into the chromosome by creating smaller pieces of the incoming DNA such that the acquired genes would contribute to the mosaic structure of the host genome (McKane & Milkman, 1995; Milkman et al., 1999; Kobayashi, 1998; Arber, 2005). In contrast, they may be instrumental to the spread of some genes by separating them from linked alleles (Arber, 2000; Milkman, 1999). Thus, RM systems may play a role in emergence of novel genetic variants of bacteria. In parallel to these observations, Jeltsch (2003) postulated that the presence of strain-specific RM systems in different lines of the same species divides the species into different biotypes which restricts the gene flow between them. Such division of one species into different biotypes provides an opportunity for genetic isolation and hence emergence of new lineages that are capable of adapting to new ecological niches. From this perspective, RM systems act for the benefit of the biological evolution of the population by promoting a rich diversity of genomes. On the other hand, there is increasing line of evidence that RM complexes act like mobile genetic elements that can promote their own horizontal transmission into new hosts and cause genome rearrangements (Kobayashi, 1999). According to the 'selfish gene' hypothesis proposed by Kobayashi (1998, 2001), RM systems are genomic parasites akin to viruses; they can increase their frequency in the genome and population and assure their maintenance by post-segregational killing or by causing cell death if challenged by a homologous stretch of DNA to replace the resident RM system (Sadykov et al., 1998; Kobayashi, 2004).

Even though RM genes, per se, are not mobile, they are often found linked to one of the well known mobile or mobility related genetic elements such as plasmids, prophage in the chromosomes, transposons, conjugative transposons, and integrons (Brassard et al, 1995; Kita et al., 1999; Rochepeau et al., 1997; Twomey et al., 1998;

Burrus et al., 2001; Rowe-Magnus et al. 2001). RM systems have undergone extensive horizontal transfer (Xu et al., 2000; Sekizaki et al., 2001; Mruk and Kaczorowski, 2003, Lin et al., 2004) to the extent that close homologues occur in distantly related organisms as manifested by sequence and phylogenetic analysis. Indeed, the codon usage and the GC content of their genes are often different from that of the host genome, and they are known to appear in distinct genomic context in different strains of the same organism. (Kobayashi, 2004; Jeltsch and Pingoud, 1996; Nobusato et al., 2000a, 2000b; Chinen et al., 2000, Bujnicki, 2001).

Listeria monocytogenes is a Gram positive, low-GC content, facultative intracellular foodborne pathogen, which is capable of causing septicemia, meningitis, and stillbirth in susceptible animal and human hosts (Farber and Peterkin, 1990; Kathariou, 2002). Molecular and other typing methods showed that the organism has a clonal population structure and has consistently classified the species of *L. monocytogenes* into distinct phylogenetic lineages: serotypes 4b, 3b and 1/2b belong to lineage I, while 1/2a, 3a, 3c and 1/2c are clustered in lineage II. However, only three serotypes (4b, 1/2b, and 1/2a) are responsible for more than 95% of the cases of illness (Rocourt et al., 1997) and strains of serotype 4b are responsible for most of the outbreaks (Kathariou, 2003). Recently, several genomic loci specific for epidemic strains of serotype 4b have been identified (Herd and Kocks, 2001; Yildirim, et al, 2004; Nelson et al., 2004). One of these loci included a gene cassette, encoding a type II restriction-modification system which renders the chromosome of the epidemic strains methylated at the cytosine of 5'-GATC-3' sites (Nelsonm et al, 2004; Yildirim et al., 2004).

Sau3AI endonuclease, which recognizes 5'-GATC-3' sequences and cleaves its substrate if cytosine at these sequences is unmethylated, was first characterized in *Staphylococcus aureus* (Seeber, 1990). Several isochizomers have recently been identified in other bacterial genomes such as AspMD1 in *Alcaligenes* sp MD1 (Chiura et al., 1992), LlacKR2I in *Lactococcus lactis* (Twomey et al, 1998), Sth368I in *Streptococcus thermophilus* (Burrus et al., 2001), and SsuDAT1I in *Streptococcus suis* (Sekizaki et al., 2001).

L. monocytogenes is not known to be naturally competent and gene flow between strains of separate lineages is limited, as suggested by the clonal nature of the organism and significant genetic distance between the two lineages (Meinersmann, et al., 2004; Nightingale et al., 2005). However, studies explaining evolutionary forces driving the population diversity of this bacterium are limited.

Here, we present evidence for the presence on the chromosome of serotype 1/2a and 1/2b strains of *L. monocytogenes* of a restriction-modification gene cassette previously identified only in the genome of epidemic strains of serotype 4b. The findings suggest the potential of horizontal transfer into the genome of serotype 1/2a or 1/2b strains.

MATERIALS AND METHODS

Bacterial strains, vectors, media and enzymes. Strains of *L. monocytogenes* used in this work are listed in Table 1.. *Escherichia coli* strain S-17 and pCON-1 as plasmid vector were used for cloning. *E. coli* was grown at 37 °C on Luria-Bertani medium supplemented with ampicillin, as needed. *L. monocytogenes* strains were grown in either Tryptic Soy Broth-Yeast Extract or Brain Heart Infusion at 37 °C supplemented with nalidixic acid and chloramphenicol. Isolation of genomic DNA and restriction enzyme analysis were described in a previous work (Yildirim et al, 2004).

DNA methods. The construction of deletion mutants lacking the restriction-modification cassette was performed in collaboration with Dr. Driss Elhanafi in our laboratory, as follows: Two ~700-bp fragments flanking the gene cassette were amplified by PCR from F2365 (4b) and A7 (1/2a) with primers located inside the cassette and in the flanking ORFs (LMO3321 and LMO324, the sequences of which are listed in Table 2). The amplified PCR products on both sides of the cassette were digested with XbaI, BamHI and BamHI , EcoRI, respectively, and the fragments were ligated together in the thermosensitive plasmid pCON-1 digested by XbaI and EcoRI. The plasmid was then electroporated into *E. coli* S17-1 and after selection with ampicillin the successful colonies were conjugated with *L. monocytogenes* strains. The subsequent selection was performed on BHI medium, supplemented with nalidixic acid (30 µg/ml) and chloramphenicol (5 µg/ml) and the resulting strain was grown at 30 °C successively with a final passage at 42 °C, followed by subsequent selection and PCR confirmation of the colony with the deletion.

DNA sequencing and analysis. Automated DNA sequencing was performed on double stranded-templates from PCR products at the University of North Carolina Chapell Hill Genome analysis facility (http://152.19.68.152), which utilizes Applied Biosystems machines and chemistries. Pairwise alignment was done using BLAST2 (http://www.ncbi.nlm.nih.gov/) and Align (http://www.ebi.ac.uk/). Multiple sequence alignments were performed with CLUSTAL-W. Related sequences were detected in the NCBI databases by using BLASTX, BLASTP, and PSI-BLAST search tools.

PCR and Southern Blots. Primers used in this study are listed in Table 2. Primers were designed using a web based program, (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and were purchased from QIAGEN. PCR employed Ex Taq DNA polymerase (Fisher Scientific, PA). The reaction mixtures were prepared on ice and subjected to a hot start (95°C for 5 min) prior to 30 cycles of amplification (95°C for 1 min, 50°C for 1 min, 72°C for 2 min, and a final extension at 72°C for 10 min) in a thermocycler (Biometra, T-1 thermoblock, Germany). Amplified products were separated by electrophoresis through a 1.0% (wt/vol) agarose gel with 1M Tris-borate-EDTA running buffer.

To prepare DNA probes for Southern blotting, PCR products were directly purified from the PCR solution with a QIAquick PCR purification kit (QIAGEN, CA), and labeled with digoxigenin (Genius kit; Roche) according to manufacturer's instructions. Southern blotting was performed according to procedure described in the DIG application manual (Roche).

Pulsed-field gel electrophoresis (PFGE) analysis. Preparation of genomic DNAs and their restriction digests was performed according to the standardized PulseNet protocol (Graves and Swaminathan, 2001) with slight modifications; *L. monocytogenes* strains were grown on TSB-YE overnight and cell pellets from 1.0 mL liquid culture were suspended in 10 mM Tris-HCl and 2 mM EDTA buffer [pH 8.0]. The cell suspension was then mixed with lysozyme (Sigma-Aldrich CO, St. Louis, MO). The mixture was incubated for 15 min. at 37 °C before adding an equal volume of 1.2% PFGE grade agarose (Amresco, Solon, OH), proteinase-K (QIAGEN) and 10% SDS mixture to obtain the agarose plugs. Bacterial cells were lysed by incubating in lysis buffer (1% N-lauroyl sarcosine, 50 mM Tris-HCl, EDTA) for at least three hours at 55 °C. The plugs were then washed several times using sterile water and Tris-EDTA buffer before incubating with either AscI or ApaI enzyme (New England BioLabs Inc., Beverly, MA) solution overnight. The plugs were sized and loaded on an agarose gel (1.0% [wt/vol]), which was subsequently electrophoresed in 0.5X TBE buffer (Tris base, Boric acid, EDTA ([pH 8.0])) in a Bio-Rad CHEF-DR® III system (Bio-Rad Laboratories, Inc, Hercules, CA) for 22 hrs at 6 V and 14 °C. The pulse times were 4 to 40 s.

Nucleotide sequence accession numbers. The nucleotide sequences presented in this study have been submitted to the GenBank database (accession no. DQ003209).

RESULTS

Detection and prevalence of RM cassette among *L. monocytogenes* serotype 1/2a strains. Ongoing efforts to characterize food and clinical isolates of *L. monocytogenes* for susceptibility to Sau3AI and MboI digestion is routinely done. Twelve serotype 1/2a strains (n=12) and three serotype 1/2b strains (n=3) resist Sau3AI and sensitive to MboI digestion (Table 1). Methylation of the cytosine at 5'-GATC-3' sites inhibits cleavage by Sau3AI but has no effect on digestion by MboI, which, in contrast, is inhibited by methylation of the adenine at the same site (McClelland et al., 1994). Therefore, these results suggested that C5 of 5'-GATC-3' sites in the genomes of these strains were methylated. Of the 12 serotype 1/2a strains, 11 were isolated from foods (avocado or related products and seafood) and one (strain 184R) was from a human clinical sample (Table 1). All three serotype 1/2b strains, however, were isolated from the food processing plant environments.

Interestingly, the digoxigenin labeled probes (85S, M, and R) (Yildirim et al., 2005) specific for the type II restriction-modification gene cassette unique to epidemic clone-I strains (serotype 4b), reacted with the EcoRI digested genomic DNAs of several of the strains (eight of the 12 strains of serotype 1/2a, and all three strains of serotype 1/2b) with Sau3AI-resistant DNA, resulting in a band corresponding to a 5.3 kb EcoRI fragment. Four serotype 1/2a strains (A35, LA255, MI18, and the clinical isolate 184R) did not react with the probes even under low stringency, suggesting methylation at 5'-GATC-3' sites by yet unknown methylase(s) or RM system(s) in their genomes. Other strains with unmethylated DNA did not produce signal.

Absence of other marker genes specific to epidemic clone I strains: The probes that were specific to epidemic clone I strains (Yildirim et al, 2005; Herd and Kocks, 2001) did not react with the genomic DNA of any of the serotype 1/2a or 1/2b strains with Sau3AI-resistant DNA, indicating that only the restriction-modification gene cassette was acquired by these strains. The probes did not produce detectable signals even under low stringency conditions (data not shown).

When present, the RM cassette was in a genomically equivalent location in serotype 1/2a, 1/2b, and epidemic clone I (serotype 4b) strains. Genome sequences of the serotype 1/2a strain EGD-e (Glaser et al., 2001) and the epidemic clone I (serotype 4b) strain F2365 (Nelson et al., 2004) were compared by using the Artemis comparison tool, ACT, (http://www.sanger.ac.uk/Software/ACT/). Alignment of the two genome sequences revealed that ORFs LMOf2365_322 and LMOf2365_330 flanking the RM gene cassette in F2365 were highly conserved in the EGDe genome (lmo0301-0303) (Fig. 1). PCR amplifications utilizing a primer inside the cassette and a primer in lmo0301 yielded a PCR product of the expected size from all eight strains. Sequencing of the PCR product obtained from A7 (serotype 1/2a) confirmed that the RM cassette was indeed located immediately adjacent to the counterpart of LMOf2365_322 in this strain. Similar results were obtained with strain (J22), serotype 1/2b.

Nucleotide sequence organization of the gene cassette and sequence content are highly conserved between Epidemic Clone I (serotype 4b) and serotype 1/2a (strain A7). The entire RM cassette was sequenced from strain A7 (serotype 1/2a). The sequence organization of the gene cassette in A7 was found to be conserved with respect to that of F2365 (serotype 4b, epidemic clone I). Pairwise alignment of the sequence and the homologous sequence from F2365 also revealed high conservation at the nucleotide level (97.4%). The deduced amino acid sequence of the Sau3AI restriction endonuclease and its cognate methylase showed 97% and 98% identity to restriction enzyme and methylase, respectively, in the genome of serotype 4b. Remarkably, the gene encoding a putative transcription factor, located between the endonuclease and the methylase gene, was also found to be conserved, and the deduced amino acid sequence was identical in all (69) amino acid residues to that of F2365. The sequence has a DNA binding domain with a helix-turn-helix motif and therefore is reminiscent of the so-called control (C) protein present in several restriction-modification gene complexes, which regulates either one or both of the RM-system genes (Kobayashi, 2004).

Conserved motifs for other genes of the cassette were also identified in the sequences from A7 and F2365: A PSI-BLAST search of the amino acid sequence of conserved domain protein (Fig. 1) showed that it harbored an HNH conserved motif unique to nucleases, homing and restriction endonucleases, and bacteriocins such as colicin and pyrocin. At the putative catalytic center of the restriction endonuclease the consensus residue (P)DX…D..E..K) was identified. This consensus sequence is highly conserved in all isoschizomers, including MutH (the mismatch repair protein) (Friedhoff et al., 2001; Kosinski et al., 2005). On the other hand, the ORF which is convergently located at the 5' of the endonuclease and was annotated as containing a genuine frameshift (pseudogene) in the genome of F2365 was also found in the homologous location of A7, suggesting that it may constitute part of the gene cassette. The residues up to and including the frameshift were conserved. In addition, a 17-bp perfect inverted repeat (AGCAAAATGCTTCATCT) was present in the 3' end of the pseudogene, including the stop codon, both in F2365 and in A7, suggesting that this was also conserved as part of the cassette. The putative PTS operon upstream of the gene cassette was highly conserved among virtually all serotypes of *L. monocytogenes* as primers targeting the operon readily amplified the expected size of products from strains of various serotypes (data not shown). Likewise, the putative lipoprotein downstream of the cassette was found to be conserved, albeit with relatively less homology (Fig. 1)

The average G+C content of the entire cassette in both genomes were noticeably lower (31%) than the average for the genome of *L. monocytogenes* (38%). Conversely, the G+C content of the pseudogene (37%) is about the same as the genome average. The codon usage pattern of the gene cassette was also different than that of *L. monocytogenes* genome. The most frequent codons in the *L. monocytogenes* genome (F2365), in decreasing order, are :GAA/AAA/AUU/GAU/CAA; in the gene cassette, the most frequent were UUA/AUA/AUU/UUU/UCU (http://www.kazusa.or.jp/codon/) (Nakamura et al., 2000)

Genetic characterization of the gene cassette. In order to demonstrate that the observed methylation phenotype in serotype 1/2a (A7) and serotype 4b (F2365) was indeed due to the restriction-modification system, the entire gene cassette was deleted from their respective chromosome. Homologous recombination and selection with antibiotics, performed in collaboration with Dr. Driss Elhanafi, resulted in deletion mutants of F2365 and A7, designated DS1 and DS14, respectively. Genomic DNA from DS and DS14 could be readily cleaved by Sau3AI, and the deletion was confirmed by PCR and southern blots (Fig. 2). However, attempts to clone the entire RM cassette in *E. coli* for complementation were unsuccessful. When the methylase gene alone was cloned in *E. coli* (Dam+), under the IPTG induced promoter, the host chromosome was rendered resistant to Sau3AI digestion (data not shown) but the colonies showed limited growth, even in the absence of the inducer.

Alternative methylation of GATC sites. Of the 12 serotype 1/2a strains with resistance to Sau3AI digestion, four failed to yield hybridization signals with probes derived from the RM cassette, even under low stringency conditions. These findings suggested that the 5'-GATC-3' sites of these strains were methylated by alternative system(s), referred to as xRM. Amplification of the conserved locus (LMO322 and LMO0329), where the gene cassette was inserted in other strains, failed to yield any PCR proiduct with two of the strains. When the genomic DNA of the other two strains (MI-18 and A35) was used as template with these primers, products of 0.6 kb and 1.2 kb, respectively, were obtained. Sequence analysis of these PCR products revealed that the correct fragment had been amplified, but no open reading frame could be identified between the conserved genes.

Clonality of the serotype 1/2a strains with the gene cassette. Genomic DNAs of the strains listed in the Table 1 were digested with AscI and ApaI and separated by PFGE. Cluster analysis based on Dice coefficient and Ward algorithm (Fig. 3) indicated that strains that harbored the gene cassette in their genome were discernible from those that did not, in terms of PFGE pattern. Indeed, with the exception of SE076, strains that had Sau3AI-resistant DNA harbored the RM cassette separately than those that had Sau3AI resistant DNA but harbored alternative methylase(s), or those the DNA of which could be digested by Sau3AI (Fig.2). The PFGE profile of strain SE076, obtained with both enzymes, showed a very specific pattern and was invariably placed separately from other strains with Sau3AI-resistant DNA.

Two other related but clearly distinguishable pulsotypes were identified among the serotype 1/2a strains whose genome is methylated by the gene cassette described here. One was found in strain A11 and the other in a cluster of six strains (A7, A8, A29, A33, A36, A95). Moreover, this cluster is likely to represent a so-called 'persistent strain' type of *L. monocytogenes* (Autio et al., 2002) because the pulsotypes obtained with both enzymes were the same and they were all isolated from the same food or food products.

Interestingly, AscI patterns of the four strains (A35, 184R, LA255, and MI-18) whose genome were methylated by an alternative methylase clustered together and separately than the rest of the strain collection. AscI and ApaI patterns of LA-255 and MI-18 strains were the same, and were in a cluster distinct from A35 and 184R. The PFGE patterns of serotype 1/2b strains also showed that two of the strains shared indistinguishable pulsotypes and the third was closely related (data not shown).

DISCUSSION

In this study, we have demonstrated that a restriction-modification gene cassette which was identified as a genetic marker for epidemic strains of serotype 4b (lineage I) (Zhang and Kathariou,1997: Herd and Kocks, 2001; Yildirim et al, 2004), was identified en bloc by strains of *L. monocytogenes* serotypes 1/2a (lineage II) and 1/2b (lineage I). Only one type of RM system has been reported for *L monocytogenes* and before this system was only detected in serotype 4b (Herd and Kocks, 2001; Yildirim et al, 2004). This is the first report to date to the best of our knowledge about the spread of a homologue of this RM system among other lineages of the bacterium.

The RM system that we characterized in this study contained several similar but also novel features compared to previously reported type II RM systems. It is conceivable that the conserved domain protein is in effect a novel mobility-related genetic element, possibly originating in association with such as an insertion sequence or phage, because it was present in the genome of A7 as part of the gene cassette and it harbored a motif conserved in homing and restriction endonucleases, and in certain bacteriocins. Most surprising, however, was the presence of a gene, annotated as pseudogene in the serotype 4b genome, located upstream of the endonuclease as part of the gene cassette, with the same frameshift in A7 and F2365. Pseudogenes usually arise from inactivation of paralogs in a genome, deleterious mutations in single-copy host genes, and unsuccessful horizontal transfers (Liu et al., 2004). However, we found that a 17-bp perfect inverted repeat at the 3' end of this pseudogene was also conserved as part of the cassette. Therefore, it is tempting to speculate that the putative recombinase may recognize the inverted repeat and use it as scaffold to initiate excision/insertion of the cassette, including the pseudogene, in the conserved region via site-specific recombination.

Furthermore, the DNA binding protein located between the endonuclease and the methylase is reminiscent of a transcriptional regulator, more specifically 'the control protein, due to presence of the Helix-Turn-Helix (HTH) motif which is the hallmark of most activators and repressors (Aravind et al., 2005). In contrast to the C proteins characterized before (e.g. PvuIIC, BamHIC, EcoRVC, SmaIC), which have several features in common with each other such as sequence similarity, transcriptional orientation with respect to endonuclease, and a conserved DNA sequence element termed "C-box" immediately upstream of the gene (Tao et al., 1991; Vijesurier et al., 1999; Kita et al., 2002; Zheleznaya et al., 2003; Streeter et al.,2004; Knowle et al., 2005), the putative C protein described in this work did not have homology to other C proteins submitted to GenBank. Its orientation was the same as the methylase encoding gene, and lacked a consensus C-box upstream of the gene. Therefore, this protein appears to be a novel control protein and needs to be further characterized.

There are striking similarities between the gene cassette characterized in this work and some other isoschizomeric type II RM-systems such as Sau3AI, LlacKR2I, Sth368I, and SsuDAT1I, in terms of the functional catalytic center located in the N terminus of the protein, the bidirectional organization of the gene cassette, and the presence of a mobility- related gene. Nevertheless, dissemination of an isoschizomer of this RM system among different strains of a bacterial pathogen has only been described for *S. suis.* (Sekizaki et al (2001a; 2001b).

We were unable to clone and express the entire RM cassette in *E coli* and our observations with cloning the RM systems with specificity to 5'-GATC-3' sites were in line with previous reports (Seeber et al., 1990; Guthrie et al., 1996). Over-expression of methylase completely abolished growth of *E. coli*, which is consistent with previous reports with respect to the presence of unmethylated GATC sites in the *E. coli* genome and the 'GATC regulatory network' under the control of the Dam methylase (Riva et al., 2004; Oshima et al., 2002).

Insertion of restriction-modification systems into an operon-like locus conserved within the same species is not uncommon (Stein et al., 1998; Claus et al., 2000; Sekizaki at al., 2001; Saunders and Snyder, 2002). The GC content of the flanking regions bordering the gene cassette (41-45%) was remarkably higher than that of the genome average (38%), which brings the mind a decaying prophage or a genomic island that might have been acquired by an ancestral *L. monocytogenes* lineage. The mechanism driving insertion of the gene cassette in this locus remains poorly understood even though it is possible that conserved sequences provide an opportunity for recombination. We amplified this region from four strains with an alternative methylase and obtained variable size of PCR products. Hence it is possible that the conserved locus may be facilitating incorporation of various "migratory genes" as described by Sibley and Raleigh (2004).

It has been speculated that an operon in the vicinity of the RM system is likely to acquire stability in persistence and expression, especially when it is in competition with an operon of a similar function in the genome (Saunders and Snyder, 2002; Kobayashi, 2004). The putative operon upstream of the gene cassette may play a role in glucose uptake. Although we grew both wild type and RM mutant in a synthetic minimal medium, HTM (Tsai and Hodgson, 2003), containing either glucose or fructose, we did not observe any apparent differences between the two with respect to growth rates (data not shown). However, growth was highly limited (OD600 less than 0.2) in this growth medium and further experiments need to be conducted to see whether the presence of an RM system may indeed have an impact on the fitness of its host.

The data presented here add to an emerging theme of RM-systems as genetic elements that are horizontally transferred within and between species; in this case between the lineages of *L. monocytogenes*. While our description of the spread of the RM cassette among lineages of this bacterium is consistent with the "selfish behavior" of these genes, it is also plausible that the cassette was horizontally transferred and contributed to emergence of novel strains that later acquired additional DNA fragments, as in the example of epidemic clone I strains of serotype 4b. In fact, it is interesting to note that several of the serotype 1/2a strains that acquired the gene cassette share similar pulsotypes despite having been recovered from foods several years apart. They were isolated from similar foods (avocado and its products) or food processing environments (floor drain), and therefore may represent a persistent strain type of this bacterium. Consequently, one could speculate that the RM gene cassette may have facilitated acquisition of DNA fragments which conferred a competitive advantage to the strains in adapting to food processing environments. Comparative genomic hybridizations will shed more light on the question of whether the gene contents of these strains are also similar.

REFERENCES

1. Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM. 2005. The many faces of the helix-turn-helix domain: transcription regulation and beyond. FEMS Microbiol Rev. 29:231-62.

2. Arber W., 2005. Gene products with evolutionary functions. Proteomics. 5:2280-4.

3. Arber, W. 2000. Genetic variation: molecular mechanisms and impact on microbial evolution. FEMS Microbiol. Rev. 24:1–7.

4. Bickle, T. A., and D. H. Kruger. 1993. Biology of DNA restriction. Microbiol. Rev. 57:434–450.

5. Brassard S, Paquet H, Roy PH. 1995. A transposon-like sequence adjacent to the AccI restriction-modification operon. Gene. 157:69-72.

6. Bujnicki JM. 2001. Understanding the evolution of restriction-modification systems: clues from sequence and structure comparisons. Acta Biochim Pol. 48:935-67

7. Burrus V, Bontemps C, Decaris B, Guedon G. 2001. Characterization of a novel type II restriction-modification system, Sth368I, encoded by the integrative

element ICESt1 of *Streptococcus thermophilus* CNRZ368. Appl Environ Microbiol. 67:1522-8.

8. Chinen A, Uchiyama I, Kobayashi I. 2000. Comparison between *Pyrococcus horikoshii* and *Pyrococcus abyssi* genome sequences reveals linkage of restriction-modification genes with large genome polymorphisms. Gene. 259:109-21.

9. Chiura HX, Kamiyama T, Hirano H, Futagami M, Watahiki M, Kobayashi K, Simidu U, Takagi J. 1992. Purification and characterization of AspMD1, an isoschizomer of Sau3AI, from a marine bacterium, *Alcaligenes* sp MD1. Nucleic Acids Res. 1992 Apr 25;20(8):1996.

10. Claus H, Friedrich A, Frosch M, Vogel U. 2000. Differential distribution of novel restriction-modification systems in clonal lineages of *Neisseria meningitidis*. J Bacteriol. 182:1296-303.

11. Farber JM, Peterkin PI. 1990. *Listeria monocytogenes,* a food-borne pathogen. Microbiol Rev. 55:476-511.

12. Friedhoff P, Lurz R, Luder G, Pingoud A. 2001. Sau3AI, a monomeric type II restriction endonuclease that dimerizes on the DNA and thereby induces DNA loops. J Biol Chem. 276:23581-8.

13. Glaser P, Frangeul L, Buchrieser C, Rusniok C, Amend A, Baquero F, Berche P, Bloecker H, Brandt P, Chakraborty T, Charbit A, Chetouani F, Couve E, de

Daruvar A, Dehoux P, Domann E, Dominguez-Bernal G, Duchaud E, Durant L, Dussurget O, Entian KD, Fsihi H, Garcia-del Portillo F, Garrido P, Gautier L, Goebel W, Gomez-Lopez N, Hain T, Hauf J, Jackson D, Jones LM, Kaerst U, Kreft J, Kuhn M, Kunst F, Kurapkat G, Madueno E, Maitournam A, Vicente JM, Ng E, Nedjari H, Nordsiek G, Novella S, de Pablos B, Perez-Diaz JC, Purcell R, Remmel B, Rose M, Schlueter T, Simoes N, Tierrez A, Vazquez-Boland JA, Voss H, Wehland J, Cossart P. 2001. Comparative genomics of *Listeria* species. Science. 294:849-52

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

15. Guthrie EP, Quinton-Jager T, Moran LS, Slatko BE, Kucera RB, Benner JS, Wilson GG, Brooks JE., 1996. Cloning, expression and sequence analysis of the SphI restriction-modification system. Gene. 180:107-12

16. Herd and Kocks, 2001. Gene fragments distinguishing an epidemic-associated strain from a virulent prototype strain of *Listeria monocytogenes* belong to a distinct functional subset of genes and partially cross-hybridize with other *Listeria* species. Infect Immun. 69:3972-9

17. Jeltsch, A. 2003. Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems? Gene. 317:13-6.

18. Jeltsch, A., and A. Pingoud. 1996. Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. J. Mol. Evol. 42:91–96.

19. Kathariou S. 2002. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. J Food Prot. 65:1811-29.

20. Kathariou, S., 2003. Foodborne outbreaks of listeriosis and epidemic-associated lineages of *Listeria monocytogenes*. In: M. E. Torrence and R. E. Isaacson (Ed), Microbial Food Safety in Animal Agriculture, Iowa State Press.

21. Kita, K., J. Tsuda, T. Kato, K. Okamoto, H. Yanase, and M. Tanaka. 1999. Evidence of horizontal transfer of the EcoO1091 restriction-modification gene to *Escherichia coli* chromosomal DNA. J. Bacteriol. 181:6822–6827.

22. Knowle D, Lintner RE, Touma YM, Blumenthal RM. 2005. Nature of the promoter activated by C.PvuII, an unusual regulatory protein conserved among restriction-modification systems. J Bacteriol. 187:488-97.

23. Kobayashi I., 1998. Selfishness and death: raison d'etre of restriction, recombination and mitochondria. Trends Genet. 14:368-74.

24. Kobayashi, I. 2004. Restriction-Modification Systems as Minimal Forms of Life. p. 19-62. *In* H. J. Gross (Series ed) and A. Pingoud (ed), Restriction Endonucleases. Springer-Verlag, Germany.

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

26. Kobayashi, I., Nobusato, A., Kobayashi- Takahashi, N. & Uchiyama, I. (1999) Shaping the genome-restriction-modification systems as mobile genetic elements. Curr. Opin. Genet. Dev. 9, 649–656.

27. Kosinski J, Feder M, Bujnicki JM., 2005. The PD-(D/E)XK superfamily revisited: identification of new members among proteins involved in DNA metabolism and functional predictions for domains of (hitherto) unknown function. BMC Bioinformatics. 6:172.

28. Lin TL, Shun CT, Chang KC, Wang JT. 2004. Isolation and characterization of a HpyC1I restriction-modification system in *Helicobacter pylori*. J Biol Chem. 279:11156-62.

29. Liu, Y., Harrison, PM., Kunin, V., Gerstein, M. 2004. Comprehensive analysis of pseudogenes in prokaryotes: widespread gene decay and failure of putative horizontally transferred genes. Genome Biol 5: R64

30. Matic I, Taddei F, Radman M. 1996. Genetic barriers among bacteria. Trends Microbiol. 4:69-72.

31. McClelland M, Nelson M, Raschke E. 1994. Effect of site-specific modification on restriction endonucleases and DNA modification methyltransferases. Nucleic Acids Res. 22:3640-59.

32. McKane M, Milkman R. 1995. Transduction, restriction and recombination patterns in *Escherichia coli*. Genetics. 139:35-43.

33. Meinersmann RJ, Phillips RW, Wiedmann M, Berrang ME. 2004. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes reveals clonal and recombination histories of three lineages. Appl Environ Microbiol. 70:2193-203

34. Milkman R, Raleigh EA, McKane M, Cryderman D, Bilodeau P, McWeeny K. 1999. Molecular evolution of the *Escherichia coli* chromosome. V. Recombination patterns among strains of diverse origin. Genetics. 153:539-54

35. Milkman, R. 1999. Gene transfer in *E coli* p. 291-309. *In* R. L. Charlebois (ed.), Organization of the prokaryotic genome. ASM Press, Washington, D.C.

36. Mruk I, Kaczorowski T. 2003. Genetic organization and molecular analysis of the EcoVIII restriction-modification system of *Escherichia coli* E1585-68 and its comparison with isospecific homologs. Appl Environ Microbiol. 69:2638-50

37. Murray NE. 2001. 2001 Fred Griffith review lecture. Immigration control of DNA in bacteria: self versus non-self. Microbiology. 148:3-20.

38. Nakamura, Y., Gojobori, T. and Ikemura, T. 2000. Codon usage tabulated from international DNA sequence databases: status for the year Nucl. Acids Res. 28:292

39. Nakayama Y, Kobayashi I. 1998. Restriction-modification gene complexes as selfish gene entities: roles of a regulatory system in their establishment, maintenance, and apoptotic mutual exclusion. Proc Natl Acad Sci U S A. 95:6442-7.

40. Nelson KE, Fouts DE, Mongodin EF, Ravel J, DeBoy RT, Kolonay JF, Rasko DA, Angiuoli SV, Gill SR, Paulsen IT, Peterson J, White O, Nelson WC, Nierman W, Beanan MJ, Brinkac LM, Daugherty SC, Dodson RJ, Durkin AS, Madupu R, Haft DH, Selengut J, Van Aken S, Khouri H, Fedorova N, Forberger H, Tran B, Kathariou S, Wonderling LD, Uhlich GA, Bayles DO, Luchansky JB, Fraser CM. 2004. Whole genome comparisons of serotype 4b and 1/2a strains of the food-borne pathogen *Listeria monocytogenes* reveal new insights into the core genome components of this species. Nucleic Acids Res. 32:2386-95

41. Nightingale KK, Windham K, Wiedmann M. 2005. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and foods. J Bacteriol. 187:5537-51.

42. Nobusato A, Uchiyama I, Kobayashi I. 2000a. Diversity of restriction-modification gene homologues in *Helicobacter pylori*. Gene. 259:89-98.

43. Nobusato A, Uchiyama I, Ohashi S, Kobayashi I. 2000b. Insertion with long target duplication: a mechanism for gene mobility suggested from comparison of two related bacterial genomes. Gene. 259:99-108.

44. Oshima T, Wada C, Kawagoe Y, Ara T, Maeda M, Masuda Y, Hiraga S, Mori H. 2002. Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in *Escherichia coli.* Mol Microbiol. 45:673-95.

45. Riva A, Delorme MO, Chevalier T, Guilhot N, Henaut C, Henaut A. 2004. Characterization of the GATC regulatory network in *E. coli.* BMC Genomics. 5:48.

46. Rochepeau, P., L. B. Selinger, and M. F. Hynes. 1997. Transposon-like structure of a new plasmid-encoded restriction-modification system in *Rhizobium leguminosarum* VF39SM. Mol. Gen. Genet. 256:387–396.

47. Rocourt J and Bille J. 1997. Foodborne listeriosis. World Health Stat Q. 50:67-73

48. Rowe-Magnus DA, Guerout AM, Ploncard P, Dychinco B, Davies J, Mazel D. 2001. The evolutionary history of chromosomal super-integrons provides an ancestry for multiresistant integrons. Proc Natl Acad Sci U S A. 98:652-7

49. Sadykov M, Asami Y, Niki H, Handa N, Itaya M, Tanokura M, Kobayashi I. 2003. Multiplication of a restriction-modification gene complex. Mol Microbiol. 48:417-27.

50. Saunders NJ, Snyder LA. 2002. The minimal mobile element. Microbiology. 148:3756-60.

51. Seeber, S., C. Kessler, and F. Gotz. 1990. Cloning, expression and characterization of the Sau3AI restriction and modification genes in *Staphylococcus carnosus* TM300. Gene 94:37–43.

52. Sekizaki T, Osaki M, Takamatsu D, Shimoji Y. 2001. Distribution of the SsuDAT1I restriction-modification system among different serotypes of *Streptococcus suis*. J Bacteriol. 183:5436-40.

53. Sekizaki T, Otani Y, Osaki M, Takamatsu D, Shimoji Y. 2001. Evidence for horizontal transfer of SsuDAT1I restriction-modification genes to the *Streptococcus suis* genome. J Bacteriol. 183:500-11

54. Sibley MH, Raleigh EA. 2004. Cassette-like variation of restriction enzyme genes in *Escherichia coli* C and relatives. Nucleic Acids Res. 32:522-34.

55. Stein DC, Gunn JS, Piekarowicz A. 1998. Sequence similarities between the genes encoding the S.NgoI and HaeII restriction/modification systems. Biol Chem. 379:575-8.

56. Streeter SD, Papapanagiotou I, McGeehan JE, Kneale GG. 2004. DNA footprinting and biophysical characterization of the controller protein C.AhdI suggests the basis of a genetic switch. Nucleic Acids Res. 32:6445-53.

57. Tao, T., and R. M. Blumenthal. 1992. Sequence and characterization of pvuIIR, the PvuII endonuclease gene, and of PvuIIC, its regulatory gene. J. Bacteriol. 174:3395–3398.

58. Tao, T., J. C. Bourne, and R. M. Blumenthal. 1991. A family of regulatory genes associated with type II restriction-modification systems. J. Bacteriol. 173:1367–1375.

59. Tsai HN, Hodgson DA. 2003. Development of a synthetic minimal medium for *Listeria monocytogenes*. Appl Environ Microbiol. 69:6943-5.

60. Twomey, D. P., L. L. McKay, and D. J. O'Sullivan. 1998. Molecular characterization of the *Lactococcus lactis* LlaKR2I restriction-modification system and effect of an IS982 element positioned between the restriction and modification genes. J. Bacteriol. 180:5844–5854.

61. Vijesurier RM, Carlock L, Blumenthal RM, Dunbar JC. 2000. Role and mechanism of action of C. PvuII, a regulatory protein conserved among restriction-modification systems. J Bacteriol. 182:477-87

62. Wilson, G.G. & Murray, N.E. (1991) Restriction and modification systems. Annu. Rev. Genet. 25, 585–627.

63. Xu Q, Stickel S, Roberts RJ, Blaser MJ, Morgan RD. 2000. Purification of the novel endonuclease, Hpy188I, and cloning of its restriction-modification genes

reveal evidence of its horizontal transfer to the *Helicobacter pylori* genome. J Biol Chem. 275:17086-93.

64. Yildirim S, Lin W, Hitchins AD, Jaykus LA, Altermann E, Klaenhammer TR, Kathariou S. 2004. Epidemic clone I-specific genetic markers in strains of *Listeria monocytogenes* serotype 4b from foods. Appl Environ Microbiol. 70:4158-64

65. Zheleznaya LA, Kainov DE, Yunusova AK, Matvienko NI. 2003. Regulatory C protein of the EcoRV modification-restriction system. Biochemistry (Mosc). 68:105-10

66. Zheng W, Kathariou S. 1998. Host-mediated modification of Sau3AI restriction in *Listeria monocytogenes*: prevalence in epidemic-associated strains. Appl Environ Microbiol. 63:3085-9

Table 1. Strains used in this study

Strain ID	Year Isolated	Source	Serotype	Sau3AI cut	MboI cut	85M and 85S probe Reactivity	Reference or source
A7	1999	Guacamole	1/2a	No	Yes	+	FDA
A8	1999	Mexico avocado pulp	1/2a	No	Yes	+	FDA
A11	2001	Avocado	1/2a	No	Yes	+	FDA
A29	2002	Avocado	1/2a	No	Yes	+	FDA
A33	2003	Avocado	1/2a	No	Yes	+	FDA
A36	2003	Guacamole	1/2a	No	Yes	+	FDA
SE076	1986	Crab meat, frozen	1/2a	No	Yes	+	FDA
A95	2004	Avocado pulp	1/2a	No	Yes	+	FDA
A35	2003	Guacamole	1/2a	No	Yes	−	FDA
LA255	?	Food Ag Lab QC-Lab strain	1/2a	No	Yes	−	NCDOH
#1-MI-18	2005	Processing plant environment	1/2a	No	Yes	−	This study
184R	2003?	Clinical	1/2a	No	Yes	−	NCDOH
A22	2001	Sliced turkey	1/2a	Yes	Yes	−	FDA
A34	2003	Cuttlefish	1/2a	Yes	Yes	−	FDA
A38	2003	Frozen snow crab clusters	1/2a	Yes	Yes	−	FDA
A39	2003	Cooked baby clam meat	1/2a	Yes	Yes	−	FDA
A42	2001	Whole milk mozzarella curd	1/2a	Yes	Yes	−	FDA
A67	2001	Frozen capelin roe	1/2a	Yes	Yes	−	FDA
A74	2003	Avocado scooped for ice cream	1/2a	Yes	Yes	−	FDA
A75	2003	Smoked salmon	1/2a	Yes	Yes	−	FDA
A94	2004	Cheese chile relleno	1/2a	Yes	Yes	−	FDA
A96	2004	Breaded cat fish	1/2a	Yes	Yes	−	FDA
A97	2004	Avocado puree	1/2a	Yes	Yes	−	FDA
AT-05	1987	Raw milk	1/2a	Yes	Yes	−	FDA
AT-01	1986	Cheese brie	1/2a	Yes	Yes	−	FDA
J22	2004	Floor drain	1/2b	No	Yes	+	This study
82-6a	2004	Floor drain	1/2b	No	Yes	+	This study
128a-5	2004	Floor drain	1/2b	No	Yes	+	This study

Table 2. Primer sequences and target loci

Primer Name	Sequence (5'-3')	Description
LMO0321RMDELF	GCATTCTAGACATCCGAATGTCTTCCTC	Put. glycosyl hyrolase 5' end
LMO323RMDELR	GCTCGGATCCCACTTCCAAAGTGGCAAATAA	Hypot. protein
LMO0328RMDELF	CACTGGATCCAACATGGTATAGAGATTG	Put. recombinase 3' end
LMO0329RMDELR	ATACGAATTCTTTTTCATCTGCTTTGACATC	Put lipoprotein 3' end
LMO_0322R	AAGTGACCTTCCGCATGAAC	Put. PTS component
85S_endR	GTCCATGTAAAGCCGTTTGG	Endonuclease 3' end
LMO329_408F	TTGGTTGTGGTAAAGAGGTTGTT	Putative lipoprotein 5' end
LMO329_408R	CGAAATCTTTTTCATCTGCTTTG	Putative lipoprotein 3' end
LMO330_578	CGAAAATTCTCGGGAAACAA	Put. aldolase
LMO330_578	ATCCGCGGACAAAATATACG	Put. aldolase
LMO0328L	CATCTTATACCTATGGCAGTACAGGA	Put. recombinase, 3' end
LMO0324F	CTTGCTGTCCTAAGCTGTTTTT	Pseudogene 5' end
LMO0324R	GATGAAGCATTTTGCTGATTTT	Pseudogene 3' end
LMO0325F	ACCACGGGCATACTCGTTTA	Endonuclease 5' end
LMO0325R	GGGCCTCAAATGAAGAACAT	Endonuclease 3' end
M1379F	AATATATTTTCAATGTTTGATGGT	Methylase 5' end
M1379R	GCTAATTCAATCCCTATTCT	Methylase 3' end
LMO321F	ACGTTTAACGAGCCGAAATG	Put. glycosyl hyrolase 5' end
LMO321R	CGATTCGGTAGGCATCATCT	Put. glycosyl hyrolase 3' end

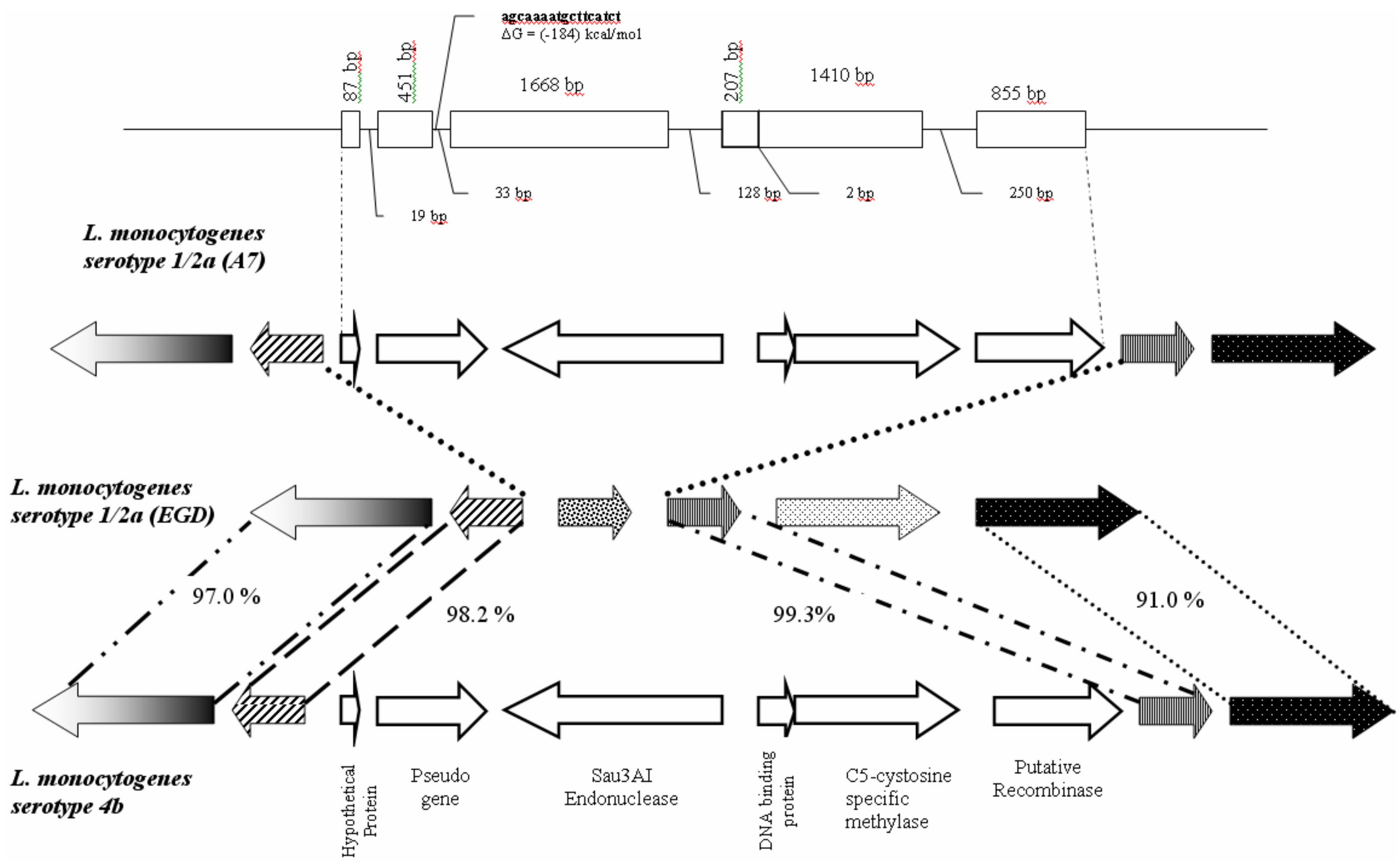


Figure 1. Comparison of genomic regions between strains EGD (serotype 1/2a), F2365 (serotype 4b), and A7 (serotype 1/2a)

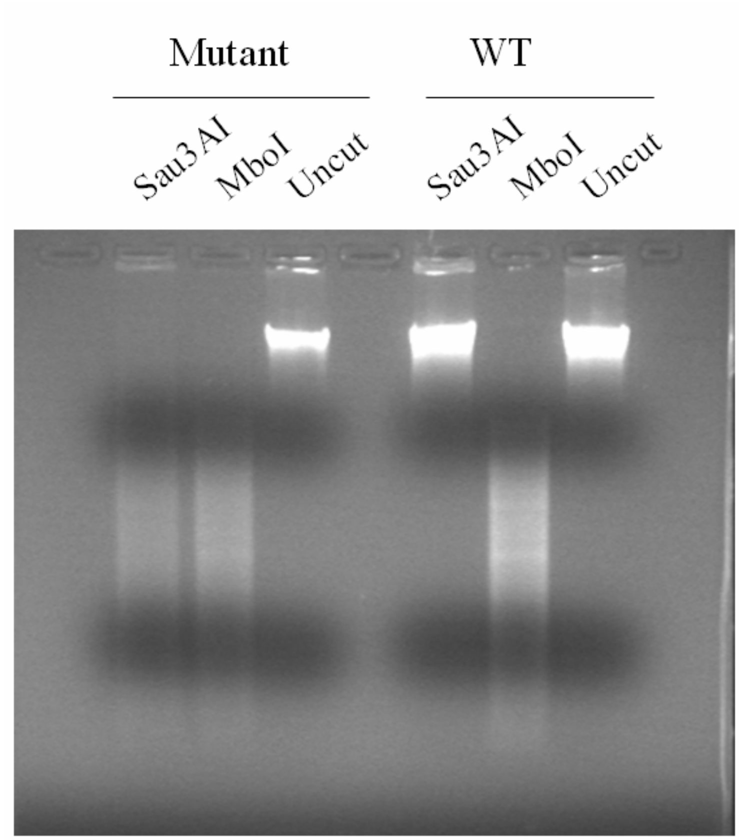


Figure 2. Restriction digest of the chromosome of A7 strain
and its isogenic mutant

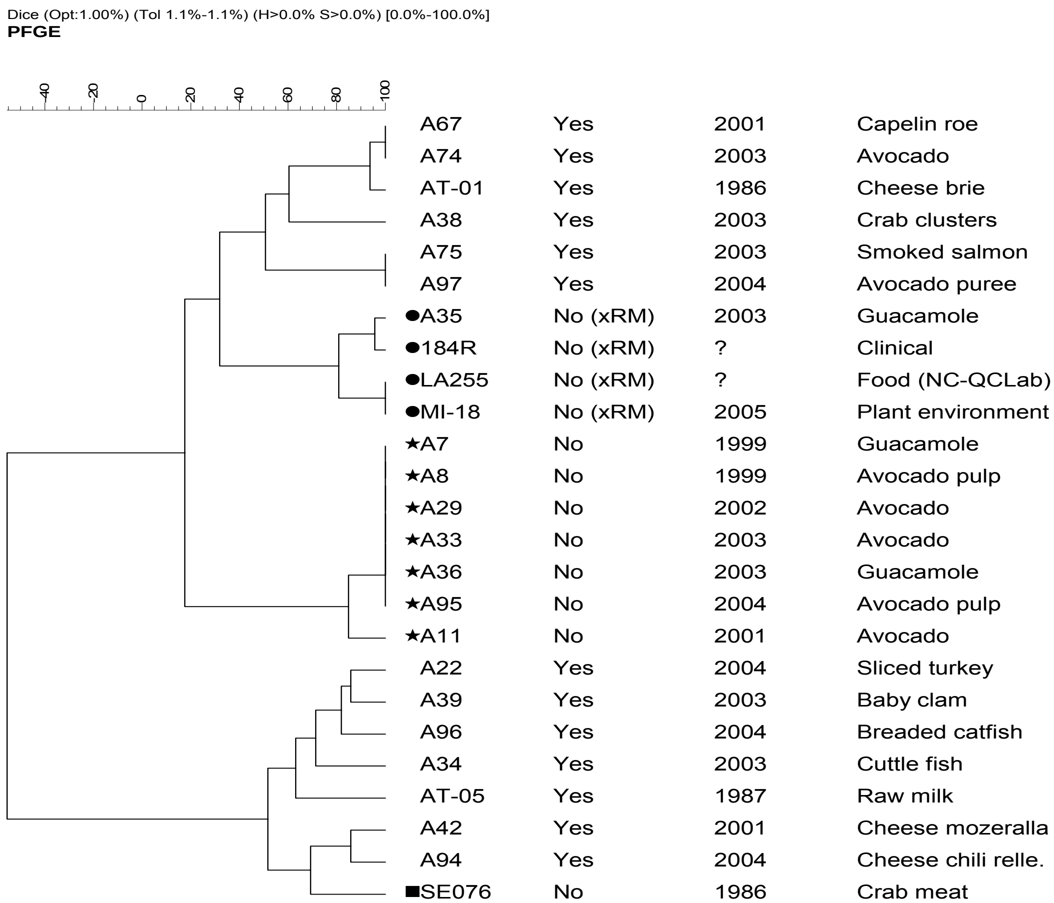
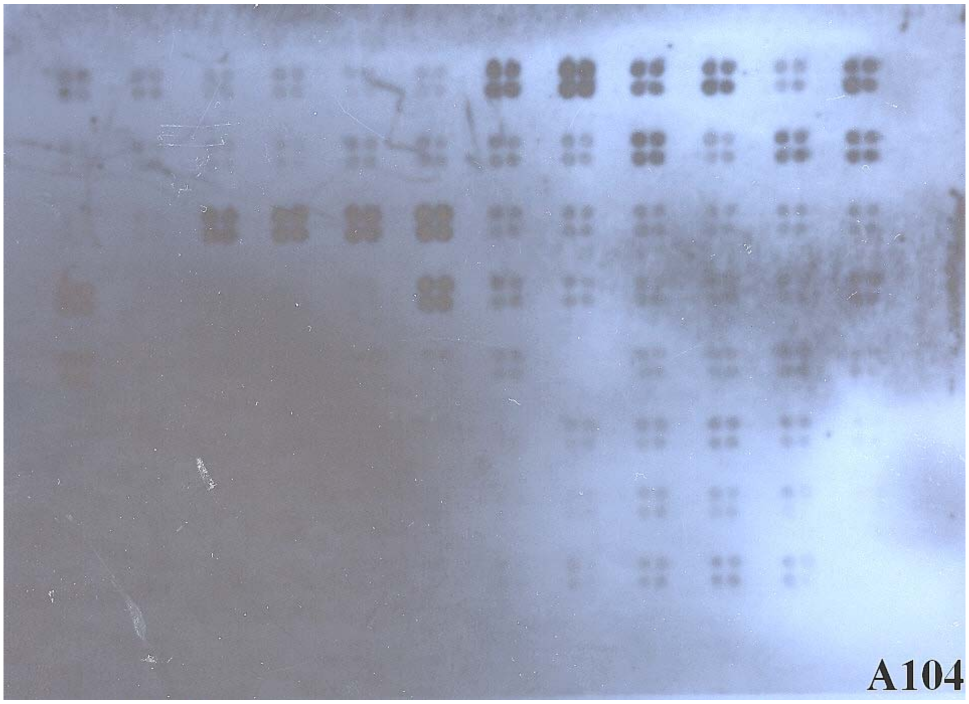
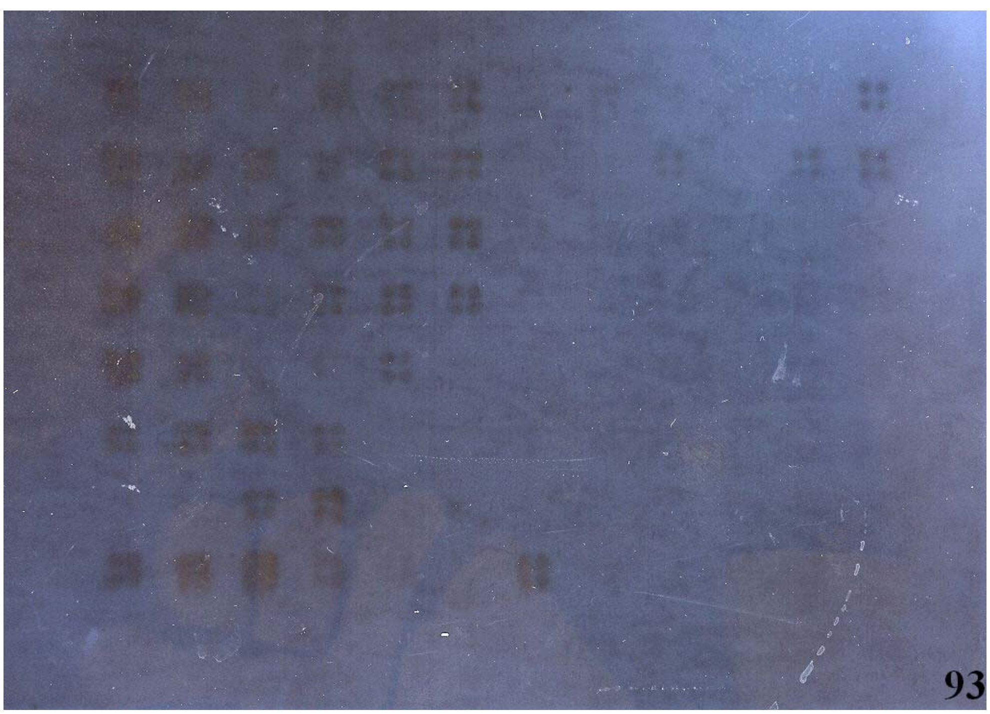


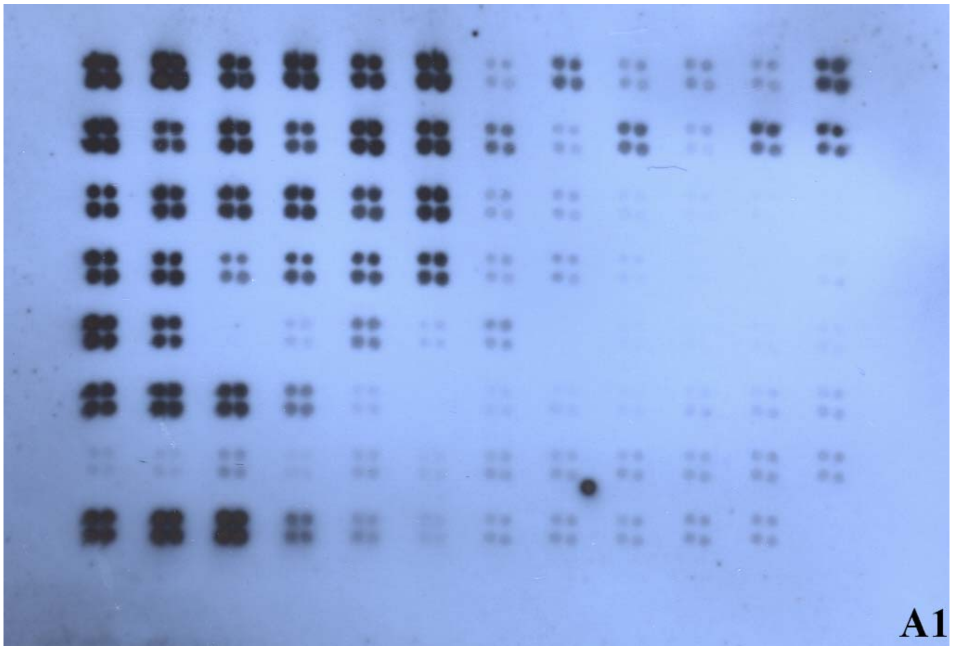
Figure 3. Dendogram analysis of the pulsotypes of strains obtained with AscI (all serotype 1/2a only; “Yes” or “No” shows whether the chromosome can be restricted by Sau3AI enzyme or not; “xRM” is to denote strains with alternative methylation system, whose recognzition site is GATC)

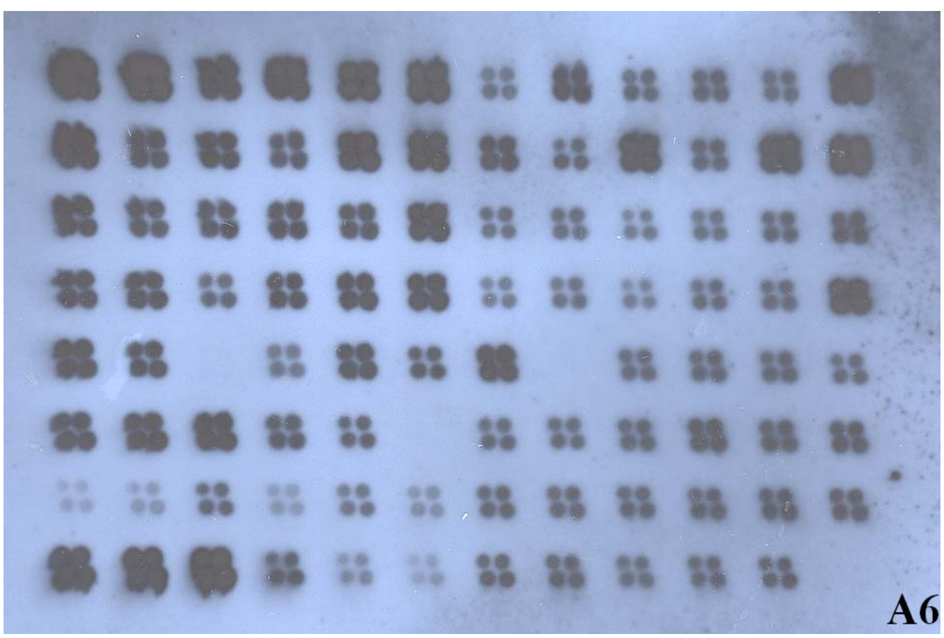
Appendix A

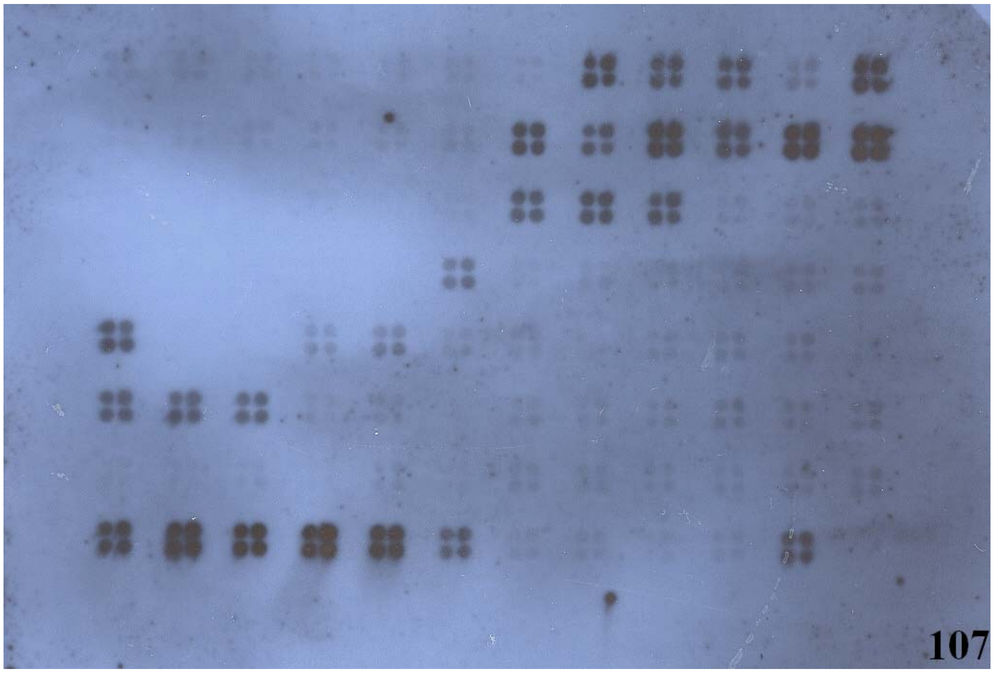
Macro-array blot pictures

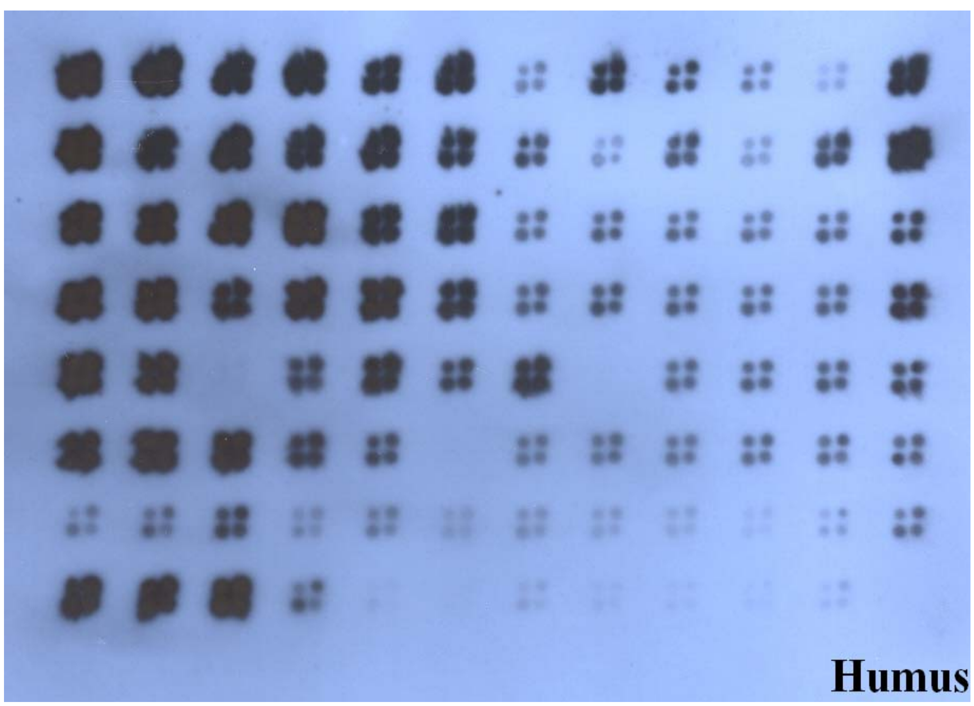

A104

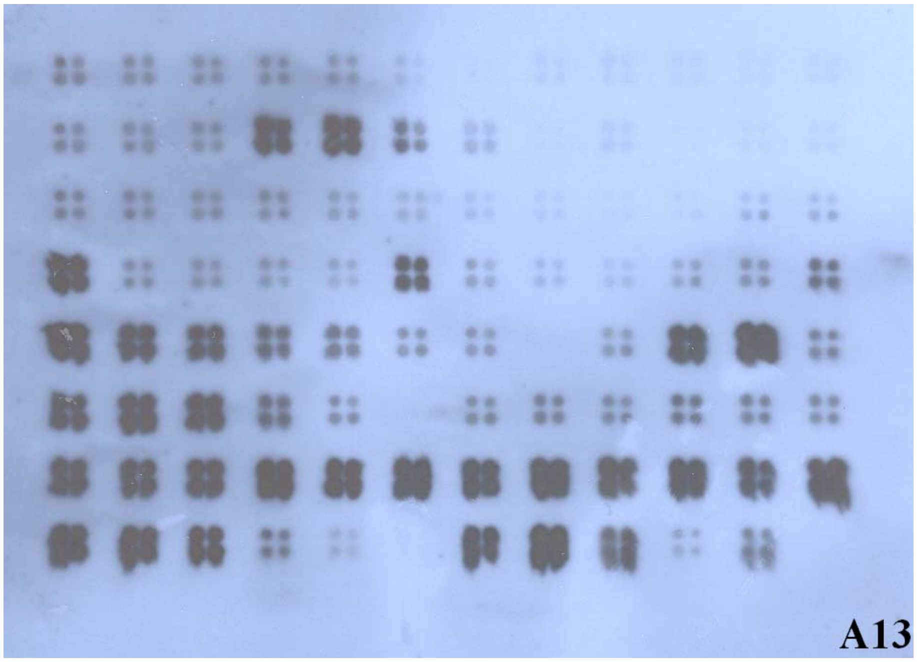

93

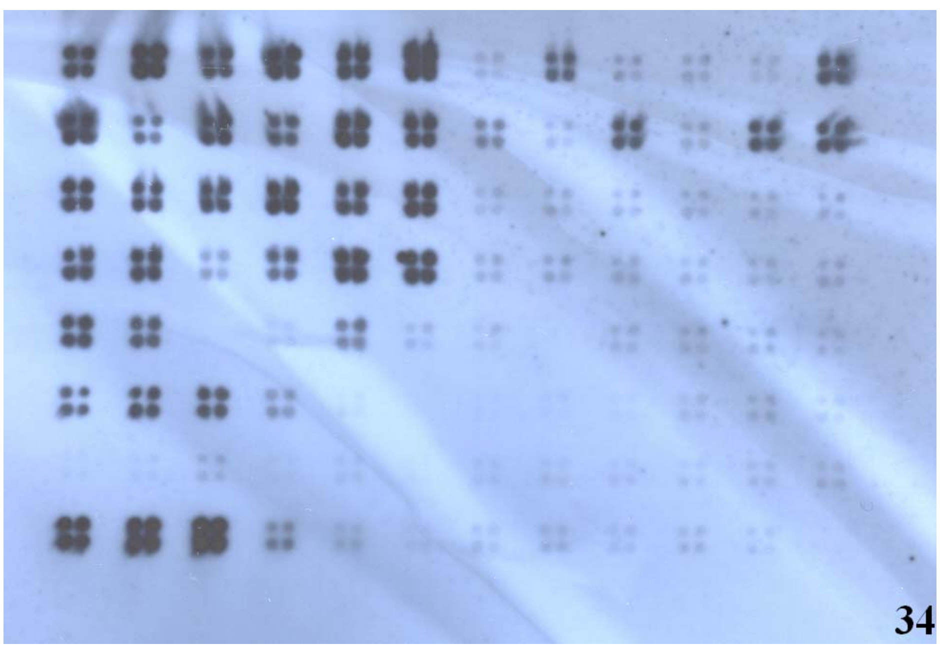

A1

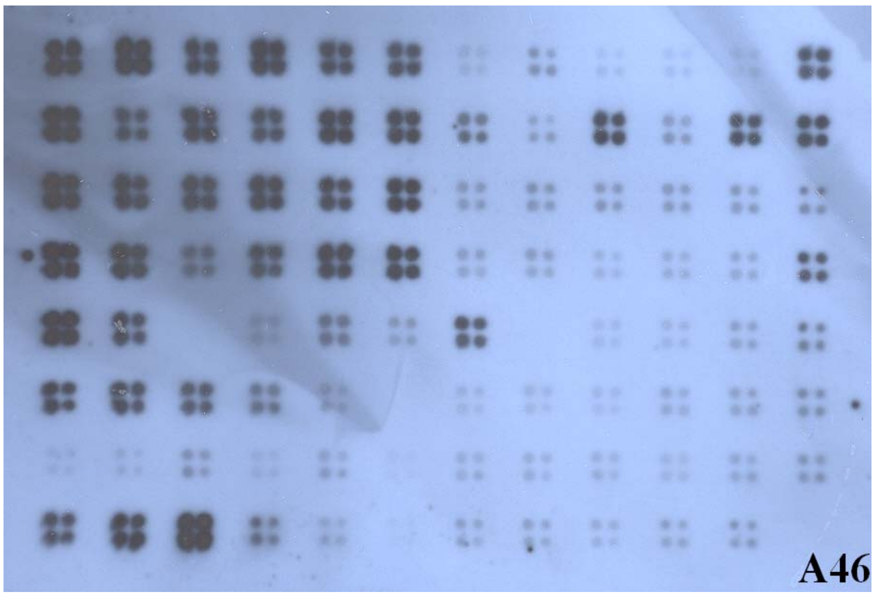

A6

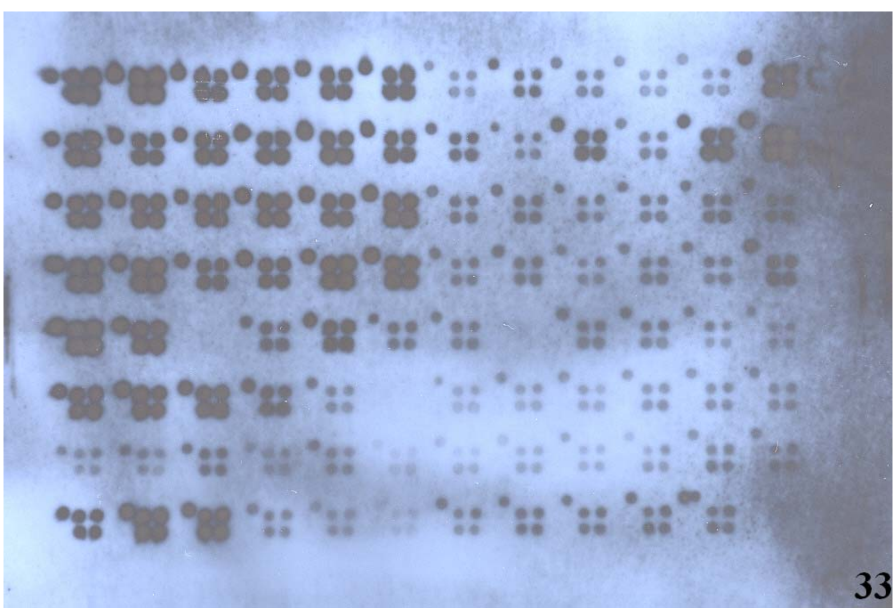

107

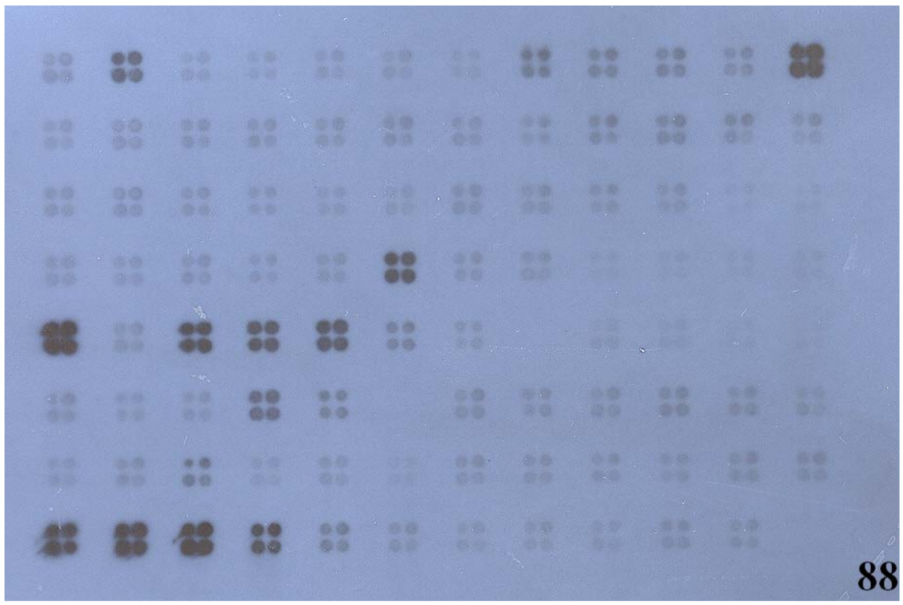

Humus

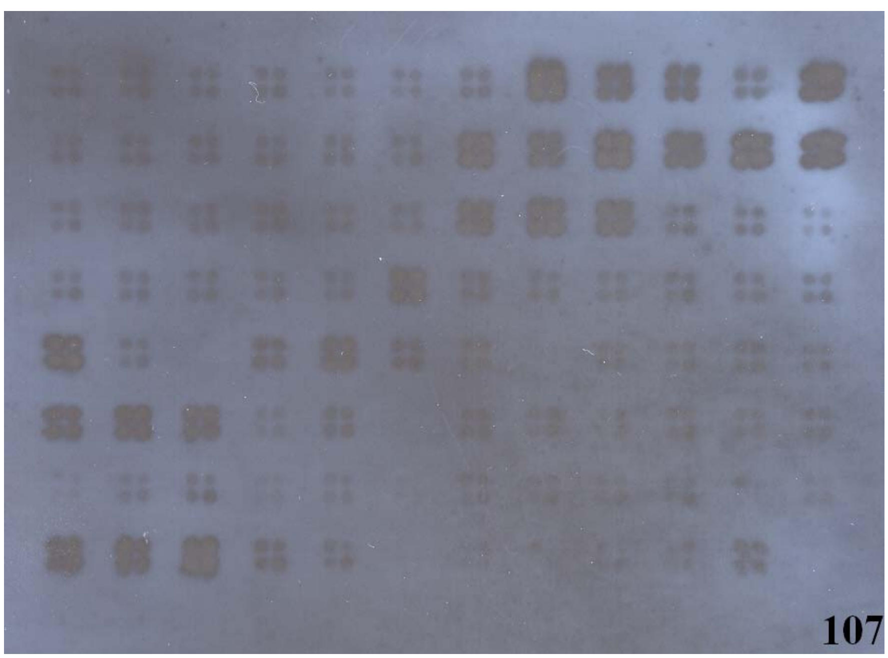

A13

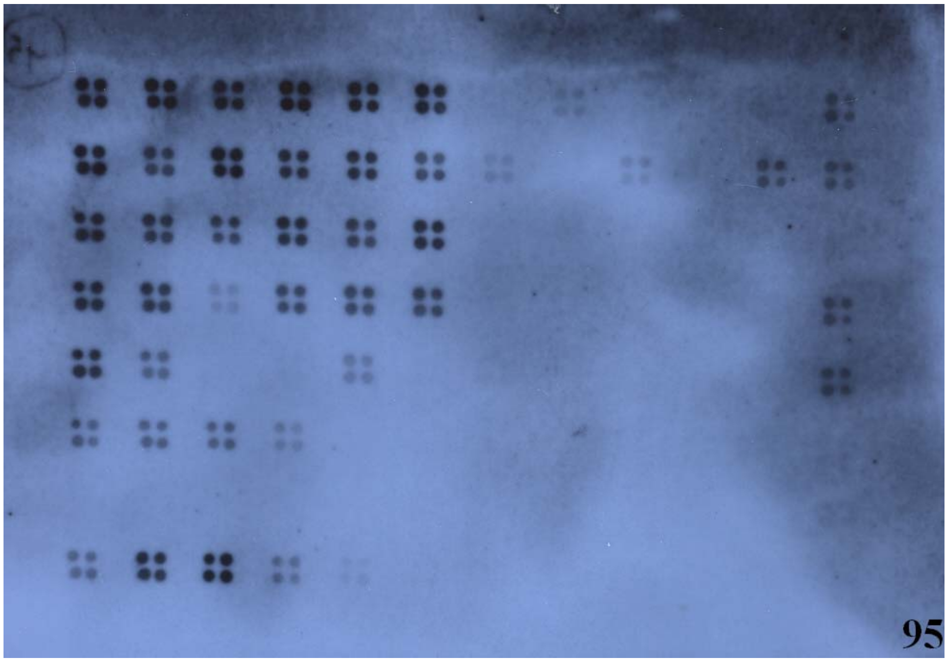

34

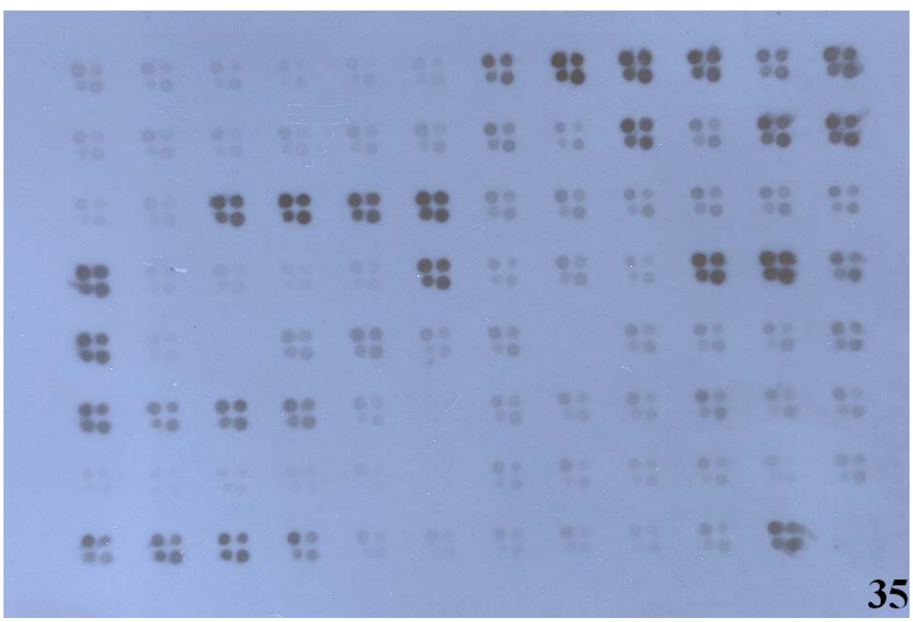

A46

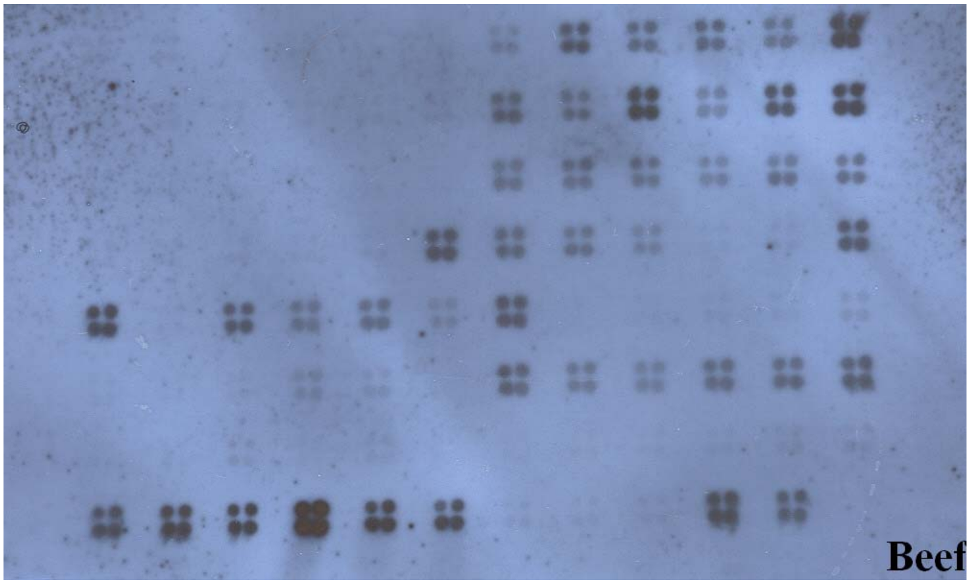

33

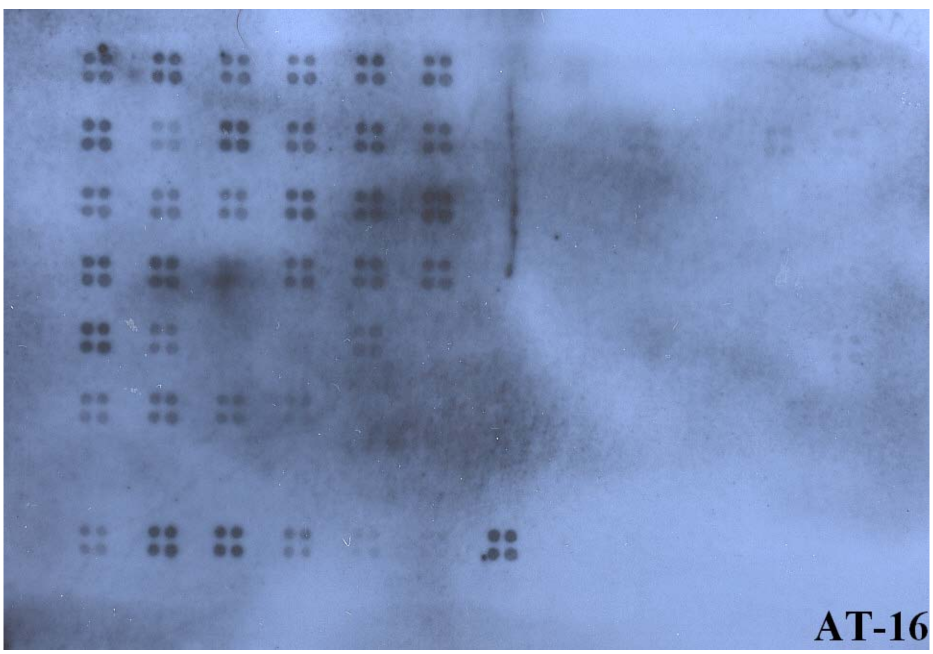

88

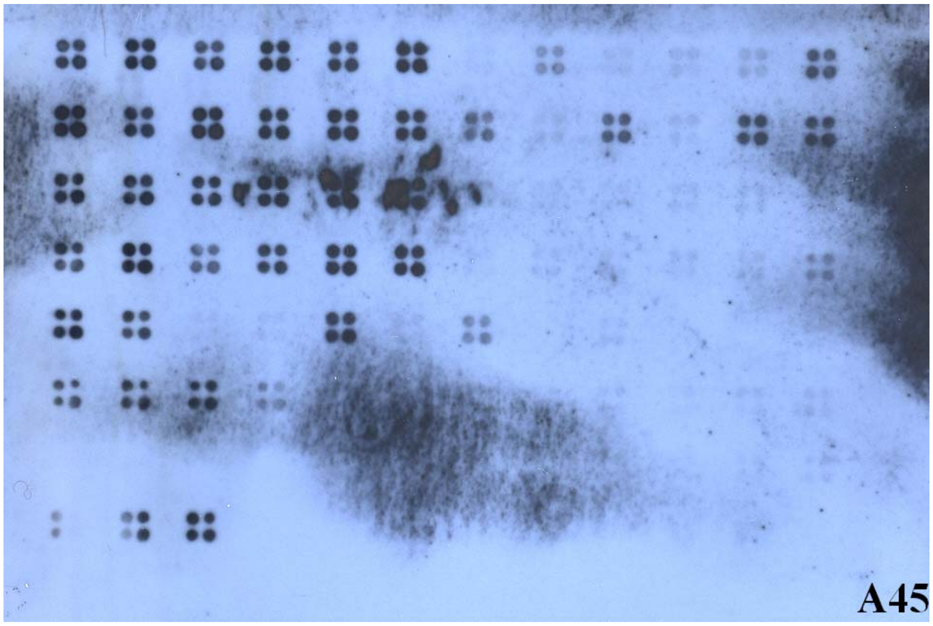

107

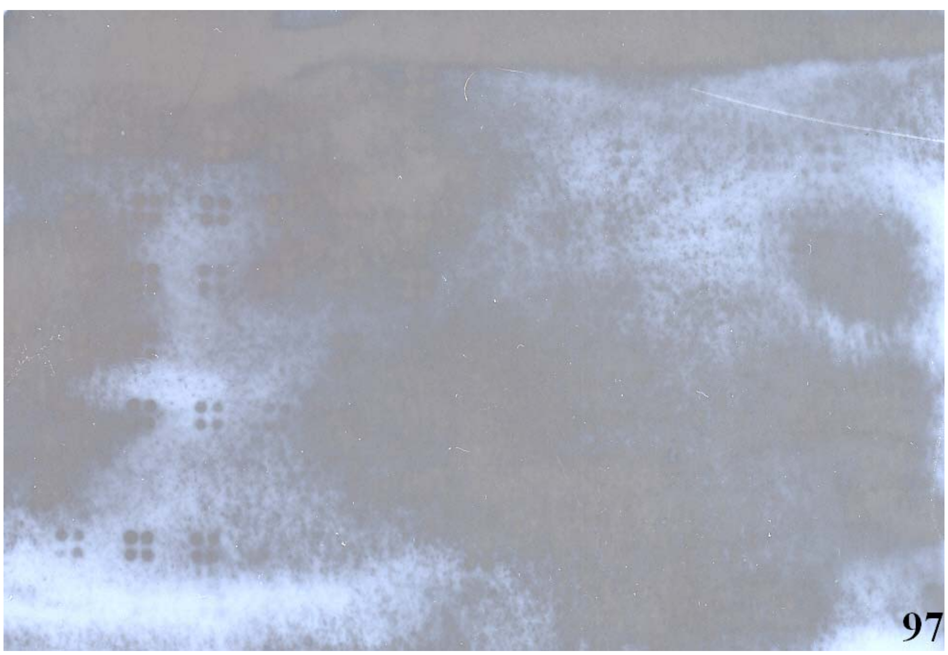

95

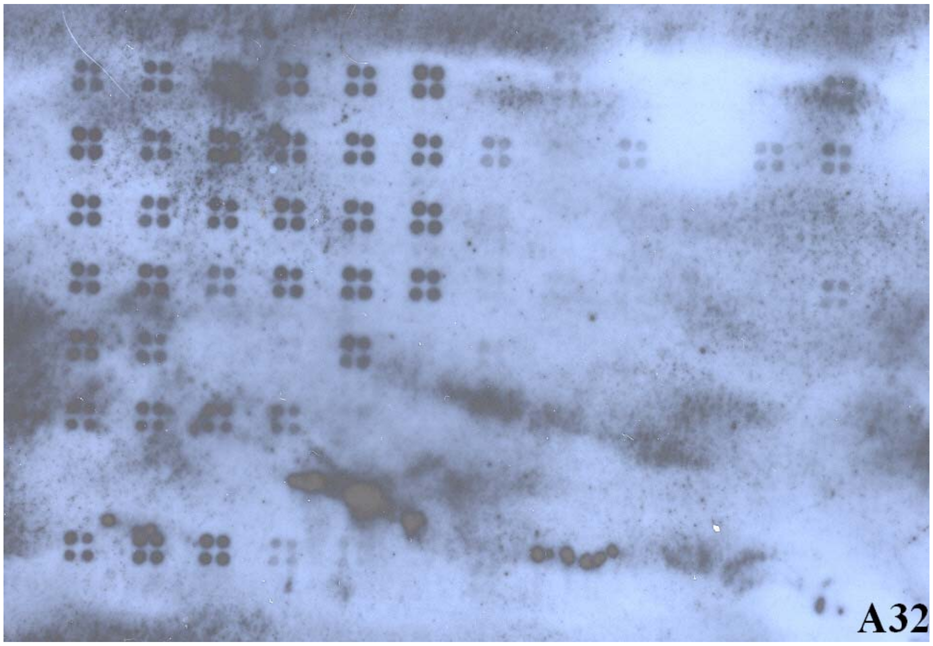

35

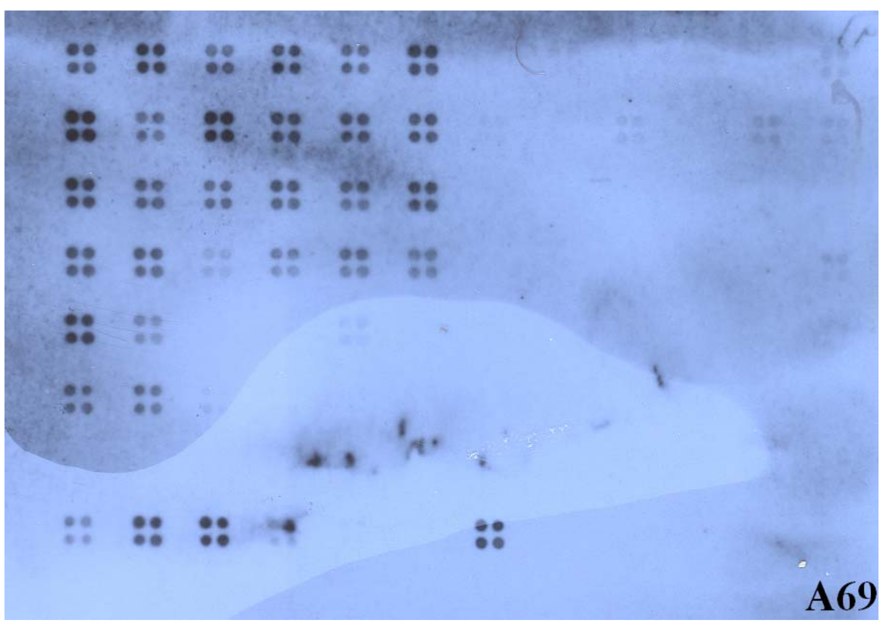

Beef

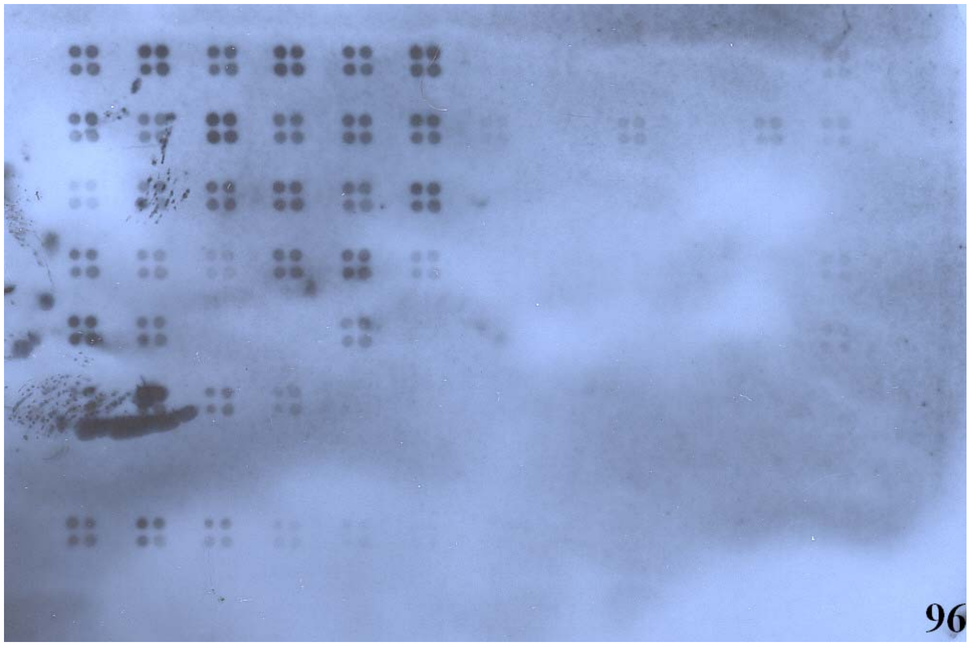

AT-16

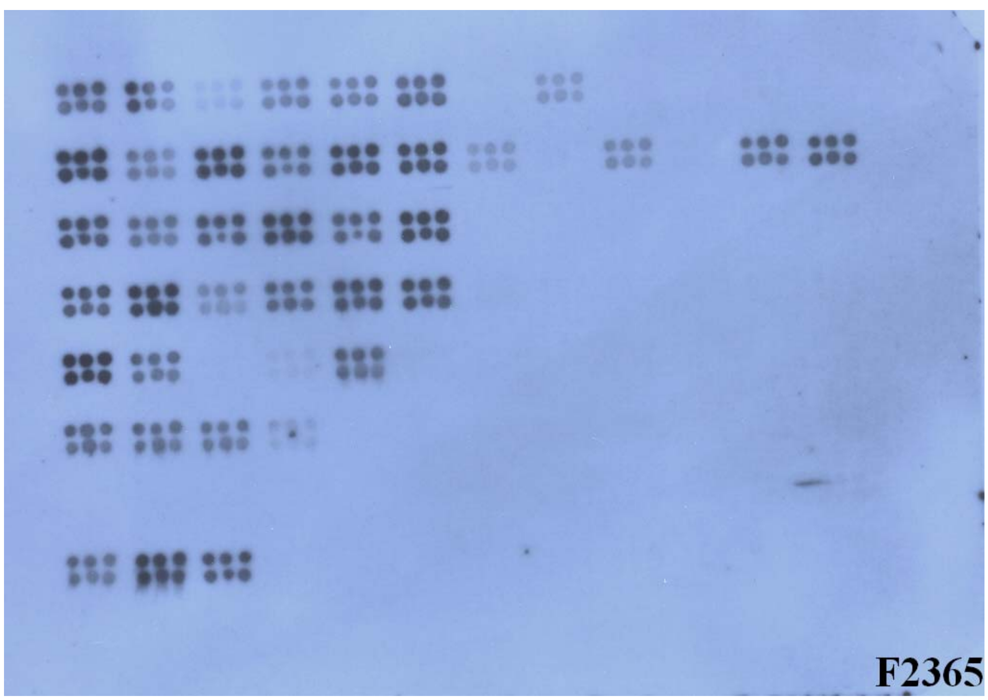

A45

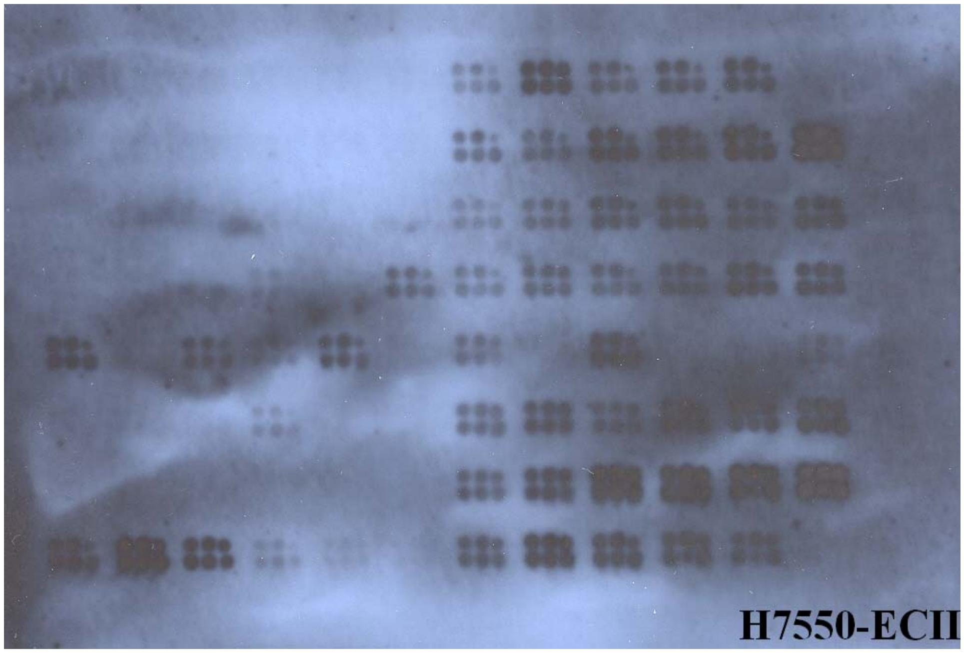

97

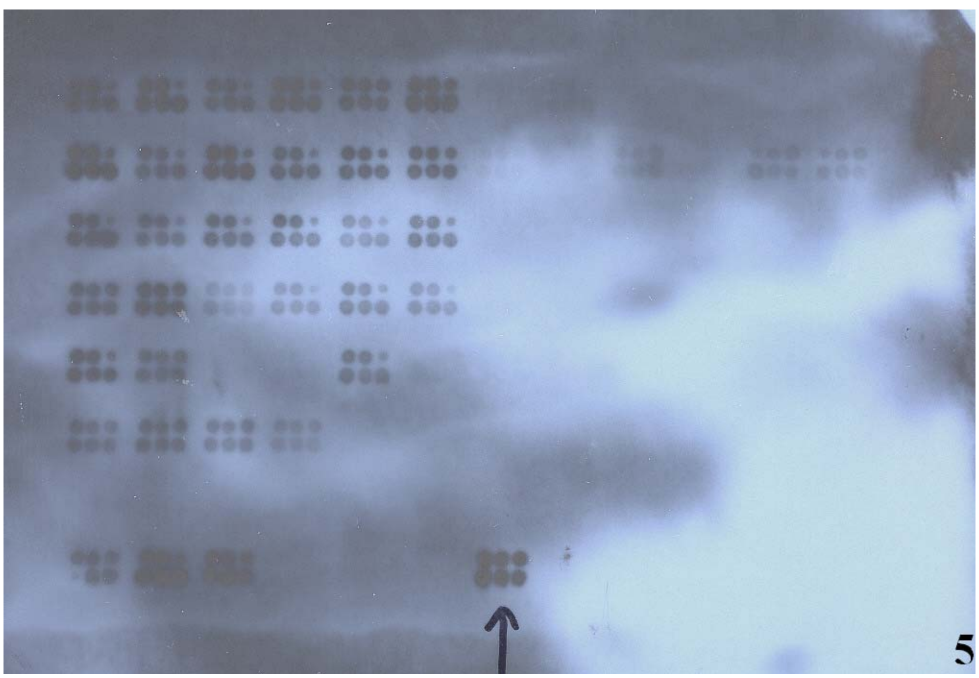

A32


A69


96


F2365


H7550-ECII


5

Appendix B

Sequence alignments of Restriction-Modification system

4bRM vs sero-1 RM pairwise nucleotide alignment

```
# Length: 5108
# Identity:    4973/5108 (97.4%)
# Similarity:  4973/5108 (97.4%)
# Gaps:          11/5108 (0.2%)
# Score: 26593.0
#
##=======================================

Sero1RM            1 GTGAAATTATTATTTGCCACTTTGGAAGTGGAAATTGGT---------TT     41
                     |||||||||||||||||||||||||||||||||||||||         ||
4bRM               1 gtgaaattattatttgccactttggaagtggaaattggtttaggcgaatt     50

Sero1RM           42 TAATAGATTTTTACTTCATAAGACTAAACAAAATTAACTTTTTAGGAGAG     91
                     .|||||||||||||||||||||.|||||||||||||| ||||||||||||
4bRM              51 aaatagatttttacttcataagtctaaacaaaattaa-tttttaggagag     99

Sero1RM           92 GTGAAAatggcctttctgaaaaagtaaggcctttacatcttattacttgc    141
                     |||||||||||||||||||||||.||||.|||||||||||||||||||||
4bRM             100 gtgaaaatggcctttctgaaaaaataagacctttacatcttattacttgc    149

Sero1RM          142 tgccctaagctgtttttttaactcctacacacttggaatattttaccca    191
                     ||.||||||||| ||||||||||||||||.||||||||||||||||||||
4bRM             150 tgtcctaagctg-ttttttaactcctacaaacttggaatattttaccca    198

Sero1RM          192 gaattaaattaaatgatacggatgattagggtgttcggggagcgtttttg    241
                     |||||||||.||||||||||..|||||||.|||||||||||||||||.||
4bRM             199 gaattaaatcaaatgatacgagtgattagagtgttcggggagcgtttatg    248

Sero1RM          242 gtttgggcgacggtgtctgggaggctgtttatactttcgttaccgcgatt    291
                     |.|||||.|||||||||||||..|||||||||.|||||||||||||||||
4bRM             249 gattgggggacggtgtctggggtgctgtttatcctttcgttaccgcgatt    298

Sero1RM          292 tatcagagtgttgttaaaacttcaaggaggttgatgagtttgcttgctac    341
                     |||||||||||||||||||||||||||||||||||||||||.||||||||
4bRM             299 tatcagagtgttgttaaaacttcaaggaggttgatgagtttacttgctac    348

Sero1RM          342 tgcgattggagattgtgagttattcagtgttattacggagcaggcgagct    391
                     ||.||||||.||||||||||||||||||||||||||||||||||.|||||
4bRM             349 tgtgattggtgattgtgagttattcagtgttattacggagcaggtgagct    398

Sero1RM          392 ggacgttttcgaagggggttcacaataatggcttggggaatgtggaattt    441
                     ||||||||||||||.||.|||||||||.||||||||||||||||||||||
4bRM             399 ggacgttttcgaagtggtttcacaatattggcttggggaatgtggaattt    448

Sero1RM          442 ttgatttttctgttgtttttggtcatacttcatcggaaaactggcaataa    491
                     ||||||||||||||||||||||||||||||.|||||.||||||.|.||||
4bRM             449 ttgatttttctgttgtttttggtcatactttatcgggaaactgacgataa    498

Sero1RM          492 gacgctaaataatgtttatcttgtgctaaaatcagcaaaatgcttcatct    541
                     |||||||||.||||||||||||||||||||||||||||||||||||||||
4bRM             499 gacgctaaaaaatgtttatcttgtgctaaaatcagcaaaatgcttcatct    548

Sero1RM          542 agttttcagatgaagcattttgctctttaagtattaatatTTATCGTaaa    591
                     |.||||.||||||||||||||||||||||||||||||||||||.||||||
4bRM             549 attttaagatgaagcattttgctctttaagtattaatatttattgtaaa    598

Sero1RM          592 aggtctttaatttgctcaaatatataatctttgcttaaccaccaagcttg    641
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             599 aggtctttaatttgctcaaatatataatctttgcttaaccaccaagcttg    648

Sero1RM          642 tttcaccatataattgtctgtgtatatatctttttctgattctggcctgt    691
                     ||||||||||||||||||||||||||||||||||||||.|||||||||||
4bRM             649 tttcaccatataattgtctgtgtatatatctttttctgtttctggcctgt    698

Sero1RM          692 taatccactttgagttacatggtaatttactagaattgttcatatatttt    741
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             699 taatccactttgagttacatggtaatttactagaattgttcatatatttt    748

Sero1RM          742 tttttctcttttccattttctataatagttttatagtaaggcgcattata    791
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
4bRM             749 tttttctcttttccattttctataatagttttatagtaaggcgcattata    798

Sero1RM          792 ctgtgattttgaagcaccaggcctcaaatgaagaacgttttccagagatg    841
                     |||.||||||||||||||.|||||||||||||||||.|||||||||||||
4bRM             799 ctgagattttgaagcaccgggcctcaaatgaagaacattttccagagatg    848

Sero1RM          842 gtttaatgaaattgtttaaaattttactttcctgaagaatctttagta    891
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             849 gtttaatgaaattgtttaaaattttactttcctgaagaatctttagta    898

Sero1RM          892 taagttaattcaacgcccttcttcaatttgtctacatctgattcccaaac    941
                     |||||||||||.||.|||||.||||||||||||||||||||||.||||||
4bRM             899 taagttaattcgactcccttttttcaatttgtctacatctgatttccaaac    948

Sero1RM          942 ggctttacatggaccataaatatcagatgccggcatattccagaatttg    991
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             949 ggctttacatggaccataaatatcagatgccggcatattccagaatttg    998

Sero1RM          992 cccctacaaagaatatcttttctgggttagtttcataggtgtttttcca   1041
                     |||||||||||||||||||||||||||||||||||||||||||||||||.
4bRM             999 cccctacaaagaatatcttttctgggttagtttcataggtgtttttccg   1048

Sero1RM         1042 ggatgtttatcatttaaatcattaaatactaaaagtaaaaactttgttct   1091
                     ||.|||||||||||||||||||||||||||||||||||||||||||||||
4bRM            1049 ggctgtttatcatttaaatcattaaatactaaaagtaaaaactttgttct   1098

Sero1RM         1092 atctaagaaatccgcaacacttgactcttcccagttttctgaaactaacg   1141
                     ||||||.||||||||||||||.||.|||||||||||||||||||||||||
4bRM            1099 atctaataaatccgcaacactcgagtcttcccagttttctgaaactaacg   1148

Sero1RM         1142 attcaaaagaagggatatttgaaatttgaaatgttctttaggttttcca   1191
                     ..|||||||||||||||.||||||||||||||||||||||||||||||||
4bRM            1149 cctcaaaagaagggatactttgaaatttgaaatgttctttaggttttcca   1198

Sero1RM         1192 tctcgtaacgttatagttttgagtctaatttcggctttttaatttcttc   1241
                     ||||||||.|||||||||||||||||.|||||||||||||.|||||||||
4bRM            1199 tctcgtaatgttatagttttgagtctgatttcggctttttcaatttcttc   1248

Sero1RM         1242 agaagtaacttcagctttacttgttgggagattaagaattcttctaataa   1291
                     ||||||||||||||||||||||.|||||||||||||||||||||||||||
4bRM            1249 agaagtaacttcagctttactttttgggagattaagaattcttctaataa   1298

Sero1RM         1292 tcatgttgtttaaatttttttgtgattttcgatttgtttttatcttaaaa   1341
                     ||||||||||||||||||||||.||||||..|||||||||||||||.|||
4bRM            1299 tcatgttgtttaaatttttttgcgatttttcatttgtttttatcttgaaa   1348

Sero1RM         1342 ttttcaattaaatcttgttggctccatacttttttgttagttggatagaa   1391
                     |||||||||||||||||||||||||||||||||||||||||.||||||||
4bRM            1349 ttttcaattaaatcttgttggctccatacttttttgttagtcggatagaa   1398

Sero1RM         1392 aggaagaaagcgagagataatgatttcttctagagaattatttttcaatt   1441
                     |||||||||||||||||||||||||||||||||||||||||||||||..|
4bRM            1399 aggaagaaagcgagagataatgatttcttctagagaattatttttcagct   1448

Sero1RM         1442 cattagcatctttaactatactattaatagcagcagaatgatcattacct   1491
                     ||||||||||||||||||||||||||||||||||||||||||.|||||||
4bRM            1449 cattagcatctttaactatactattaatagcagcagaatgattattacct   1498

Sero1RM         1492 ataatttgggtattaattaattcgttaataaactttgttttaaacgagta   1541
                     ||||||||.|||||||||||||||||||||||||||||||||||||||||
4bRM            1499 ataatttgcgtattaattaattcgttaataaactttgttttaaacgagta   1548

Sero1RM         1542 tgcccgtggtttaggcttaatttcgctgcctacttgagtagtgaattctg   1591
                     |||||||||||||||||||||.||||||||||||||.|||||||||||||
4bRM            1549 tgcccgtggtttaggcttaatctcgctgcctacttgggtagtgaattctg   1598

Sero1RM         1592 ccccagtggcagcttttgtacaagcgcctaaataattagtcattccttct   1641
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM            1599 ccccagtggcagcttttgtacaagcgcctaaataattagtcattccttct   1648

Sero1RM         1642 gaaagttcttctgctttactttcattaatcagtttagaaatttttccca   1691
                     ||.|||||||||||||||||||||||||||||||||||||||||||||||
4bRM            1649 gagagttcttctgctttactttcattaatcagtttagaaatttttccca   1698

Sero1RM         1692 atcttgttttataatttcaaaatctttatctgaaatttctaggactgtat   1741
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
4bRM             1699 atcttgttttataatttcaaaatctttatctgaaatttctaggactgtat  1748

Sero1RM          1742 ctttgggaagattaaaatctgttaagtcaaccatagatagaagtgtagca  1791
                      |||||||||||||||||||||||||||||||||||||||||||.||.|||
4bRM             1749 ctttgggaagattaaaatctgttaagtcaaccatagatagaagagtggca  1798

Sero1RM          1792 taagcaattttaaattgtttttgtcttcaattaattctttgcttgagaa  1841
                      ||||||||||||||||||||||.|||||||||||||||||||||||||||
4bRM             1799 taagcaattttaaattgttttttatcttcaattaattctttgcttgagaa  1848

Sero1RM          1842 ttttttcttatcactagtgtctctccgatagtattgaattagttcaatta  1891
                      |||||||||||||||||||||||||||.||||||||.|||||||||||||
4bRM             1849 ttttttcttatcactagtgtctctccggtagtattggattagttcaatta  1898

Sero1RM          1892 ctttatttttatgccaaaaagagctctcttcaaatgattttttatattca  1941
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             1899 ctttatttttatgccaaaaagagctctcttcaaatgattttttatattca  1948

Sero1RM          1942 tttttgtagtggattttatttaaaacgagtctttcaccagcaatcaatct  1991
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             1949 tttttgtagtggattttatttaaaacgagtctttcaccagcaatcaatct  1998

Sero1RM          1992 agtgacttttttgtctttatttcctttactttgtttatctataggagtta  2041
                      ||||||||||||||||||||||||.|||||||||||||||||||||||||
4bRM             1999 agtgacttttttgtctttatttcccttactttgtttatctataggagtta  2048

Sero1RM          2042 ctttaagttcaacaccggcttctttaaagtctggttcagaaatattatca  2091
                      |||||||||||||.||||||||||||||||||||||||||||||||||||
4bRM             2049 ctttaagttcaactccggcttctttaaagtctggttcagaaatattatca  2098

Sero1RM          2092 gcaggatggtcataccaaatttgttcaatgacatttcctggataagattt  2141
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             2099 gcaggatggtcataccaaatttgttcaatgacatttcctggataagattt  2148

Sero1RM          2142 gttattaacattttttaaattattttcatcaatttcacctaaagttttac  2191
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             2149 gttattaacattttttaaattattttcatcaatttcacctaaagttttac  2198

Sero1RM          2192 caatagcttctaaagctctgttgtgaactttctcttttgtatcatagatc  2241
                      ||||||||||||||||||||||||||||||||||||||||||||||.|||
4bRM             2199 caatagcttctaaagctctgttgtgaactttctcttttgtatcataaatc  2248

Sero1RM          2242 ctcttcaattgaattcccccgaatacagtatttctctctaaatatagtat  2291
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             2249 ctcttcaattgaattcccccgaatacagtatttctctctaaatatagtat  2298

Sero1RM          2292 atcaactccttggcagaattaaaactatcggttataattaataaattaat  2341
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             2299 atcaactccttggcagaattaaaactatcggttataattaataaattaat  2348

Sero1RM          2342 atcgaaatcccataaaccataatagaggtgtaaatatgattcgattaaga  2391
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             2349 atcgaaatcccataaaccataatagaggtgtaaatatgattcgattaaga  2398

Sero1RM          2392 attaatgaaatattgaaagagagagaaatgagccaaaaagcattgtgcga  2441
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             2399 attaatgaaatattgaaagagagagaaatgagccaaaaagcattgtgcga  2448

Sero1RM          2442 gctctctggacttaggccgactacagtttctgaaatgtgtagaggagtta  2491
                      |||||||||||||||.||||||||||||||||||||||||||||||||||
4bRM             2449 gctctctggacttagaccgactacagtttctgaaatgtgtagaggagtta  2498

Sero1RM          2492 gaacaactgttaatctacaacatttagaaacattaattgatgcattagaa  2541
                      ||||||||||||||||||||||||||||||||||||||||||||.||||.
4bRM             2499 gaacaactgttaatctacaacatttagaaacattaattgatgcactagag  2548

Sero1RM          2542 atagaggattttaatcaaattttaaagagagataaggtgtgacaatgagt  2591
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             2549 atagaggattttaatcaaattttaaagagagataaggtgtgacaatgagt  2598

Sero1RM          2592 ttaaatatattttcaatgtttgatggtgtaggtggcttcattgtaggact  2641
                      |||||||||||||||||||||||||||||||||||||||||||||||.||
4bRM             2599 ttaaatatattttcaatgtttgatggtgtaggtggcttcattgtaggcct  2648

Sero1RM          2642 caatgatgctaatgaagctatagagaaagaaatgtttagaacgacgtatt  2691
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
4bRM            2649 caatgatgctaatgaagctatagagaaagaaatgtttagaacgacgtatt 2698

Sero1RM         2692 ctaatcagtttgagcctagcaaaaaAGCACAagatgcatatgaagttggt 2741
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM            2699 ctaatcagtttgagcctagcaaaaaagcacaagatgcatatgaagttggt 2748

Sero1RM         2742 gtatatagatttccagaaatgaaccatattccagatgatattatgactgt 2791
                     .|||||||||||||||||||||||||||||||||||||||||||||||||
4bRM            2749 ttatatagatttccagaaatgaaccatattccagatgatattatgactgt 2798

Sero1RM         2792 ttctgataataaatttcaagaaatgcatgatgcaggagtaaatatgatag 2841
                     |||||||||||||||||||||||||||||||||.||||||||||||||||
4bRM            2799 ttctgataataaatttcaagaaatgcatgatgctggagtaaatatgatag 2848

Sero1RM         2842 ttggaggatttccttgtcaagattattcagtggcccgttctttaaaagat 2891
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM            2849 ttggaggatttccttgtcaagattattcagtggcccgttctttaaaagat 2898

Sero1RM         2892 gaacaaggaatcgcagggaaaaaaggagtcttattctgggagattattcg 2941
                     |||||.|||||.||.||.|||||||||||||||||||||||.||||||||
4bRM            2899 gaacagggaattgctggaaaaaaaggagtcttattctgggaaattattcg 2948

Sero1RM         2942 agctattaagaaaattcaacctaaatatttgattttagaaaatgtggatc 2991
                     |||||||||.|||||||||||||||||||||||||||||||||||.||||
4bRM            2949 agctattaataaaattcaacctaaatatttgattttagaaaatgttgatc 2998

Sero1RM         2992 gtttattaaagtctccttctacccagcgtggtcgagacttcgcaattatg 3041
                     ||||||||||||||||.|||||.|||||||||||||||||||||||||||
4bRM            2999 gtttattaaagtctccgtctactcagcgtggtcgagacttcgcaattatg 3048

Sero1RM         3042 ctaggagctttcaataaacttggctattcagtagaatggagagtaattaa 3091
                     ||.|||||||||||||||||.|||||||||||.|||||||||||||||||
4bRM            3049 ctcggagctttcaataaactcggctattcagtggaatggagagtaattaa 3098

Sero1RM         3092 cgcggcagaatatggacgagctcagcgtcgcagaagagttttcttttttg 3141
                     .|||||||||||||||||||||||.|||||||||||.|||||||||||||
4bRM            3099 tgcggcagaatatggacgagctcaacgtcgcagaagggttttcttttttg 3148

Sero1RM         3142 tttatcggaatgatacaaaatGggctaagagaatagatgaaaagtacgaa 3191
                     |||||.|.||||||||||||||||||||||.|||||||||||||||||||
4bRM            3149 tttatagaaatgatacaaaatgggctaagaaaatagatgaaaagtacgaa 3198

Sero1RM         3192 aatgataaactagaagaaggtttattagaagatattgttattgcttctaa 3241
                     |||||.|||||||||||||||||||||||||||.||||||||||||||||
4bRM            3199 aatgaaaaactagaagaaggtttattagaagatgttgttattgcttctaa 3248

Sero1RM         3242 tcgatatgataattatatttttcaggaaggtctatttgcaagacaatttc 3291
                     ||.|||||||.|||||||||||||||||||.|||||||||||||||||||
4bRM            3249 tcaatatgatgattatatttttcaggaaggcctatttgcaagacaatttc 3298

Sero1RM         3292 cgattgaaaataaaatcaaatacgatggcaaacgtccacgtcatgctagt 3341
                     ||||||||||||||||||||||||||||.|||||.|||||||||||||||
4bRM            3299 cgattgaaaataaaatcaaatacgatggaaaacgcccacgtcatgctagt 3348

Sero1RM         3342 cattttctaggtaaagatgcagactatgatgattatattttaaatatatc 3391
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM            3349 cattttctaggtaaagatgcagactatgatgattatattttaaatatatc 3398

Sero1RM         3392 agataactttactgggactgtttggaacactgggattatgcgacatggta 3441
                     |||||||||||||||||||||||||||.||||||||||||||||||||||
4bRM            3399 agataactttactgggactgtttggaatactgggattatgcgacatggta 3448

Sero1RM         3442 gatattttacagcggacacgactcctatagaagagccccctataagccta 3491
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM            3449 gatattttacagcggacacgactcctatagaagagccccctataagccta 3498

Sero1RM         3492 ggtacaattgtagagaaagcgaaacaaagttttataggtcagtatgataa 3541
                     ||||||||||||||||||||||||||||||||||||.|||||||||||||
4bRM            3499 ggtacaattgtagagaaagcgaaacaaagttttataagtcagtatgataa 3548

Sero1RM         3542 tctggatttagaaataaaagaatataattcttatattgagcaatatgtta 3591
                     |||||||||||.||||||||||||||||||||||||||||||||||||||
4bRM            3549 tctggatttaggaataaaagaatataattcttatattgagcaatatgtta 3598

Sero1RM         3592 tttcagatgagattaaaattgaaaaatttaaaactcttagaggttccaag 3641
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
4bRM          3599 tttcagatgagattaaaattgaaaaatttaaaactcttagaggttccaag  3648

Sero1RM       3642 aaaatccctagaaaacgtcctgatggaagtgagtatttttacgcagaagg  3691
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3649 aaaatccctagaaaacgtcctgatggaagtgagtatttttacgcagaagg  3698

Sero1RM       3692 ggcaatgtcgccatatgatagctttgatttacctgctagaactatgttaa  3741
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3699 ggcaatgtcgccatatgatagctttgatttacctgctagaactatgttaa  3748

Sero1RM       3742 ctagtgaagcatctgtaaatcgttcaactcatttattatttgaaaataat  3791
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3749 ctagtgaagcatctgtaaatcgttcaactcatttattatttgaaaataat  3798

Sero1RM       3792 aaataccgacttattacaccaattgaggcagagttattacaagattttcc  3841
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3799 aaataccgacttattacaccaattgaggcagagttattacaagattttcc  3848

Sero1RM       3842 aataaattggaccaaatataaaatgaatgaagaaacaaatgaggtttctg  3891
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3849 aataaattggaccaaatataaaatgaatgaagaaacaaatgaggtttctg  3898

Sero1RM       3892 aggtctctgatagaatgagatatttctttatgggtaacgctttagttact  3941
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3899 aggtctctgatagaatgagatatttctttatgggtaacgctttagttact  3948

Sero1RM       3942 ggaattgtaaaaagaatagggattgaattagctaaaattgagagtcatga  3991
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3949 ggaattgtaaaaagaatagggattgaattagctaaaattgagagtcatga  3998

Sero1RM       3992 atagttaaacggattaaaagtgcgcatttttgtgcactttttttagtatt  4041
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          3999 atagttaaacggattaaaagtgcgcatttttgtgcactttttttagtatt  4048

Sero1RM       4042 gttaaaatggaatcgtatgatgcactttattttataataatacttgatat  4091
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4049 gttaaaatggaatcgtatgatgcactttattttataataatacttgatat  4098

Sero1RM       4092 ttaaaacagtttatctagttagaagatgacgaagtagttagtgaatgtta  4141
                   ||||||.|||||||||||||||||||||||||||||||||||||||||||
4bRM          4099 ttaaaatagtttatctagttagaagatgacgaagtagttagtgaatgtta  4148

Sero1RM       4142 caatctaatagtaaataataatttatgtgaaaatcaatacaaagtatact  4191
                   ||||||||||||||||||||||||||||||||||.|||||||||||||||
4bRM          4149 caatctaatagtaaataataatttatgtgaaaattaatacaaagtatact  4198

Sero1RM       4192 agtataaatctctactatatttatcatagttgcaactaaattggagtggt  4241
                   |||.||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4199 agtgtaaatctctactatatttatcatagttgcaactaaattggagtggt  4248
Sero1RM       4242 tttttgtataatttagagttatttctgctaatgaggacatctttataaa   4291
                   |||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4249 tttttgtataatttagagttatttctgctaatgaggacatctttataaa   4298

Sero1RM       4292 aaccaaatatagatttatctcagagttagaagattttaaagaatttgcac  4341
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4299 aaccaaatatagatttatctcagagttagaagattttaaagaatttgcac  4348

Sero1RM       4342 aaactctaactgatgatgctggaaaagcAAAAGTTAAAAGTGGAAAGGCT  4391
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4349 aaactctaactgatgatgctggaaaagcaaaagttaaaagtggaaaggct  4398

Sero1RM       4392 AGTTCATATACTAGATACCTTATACGACTAATAATTTTTTATTGTGAAAC  4441
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4399 agttcatatactagataccttatacgactaataatttttttattgtgaaac 4448

Sero1RM       4442 TCATAGAGACAATCTAGAGGAATTATCATCATTTGAAGCAGTGAAAAAAC  4491
                   |||.||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4449 tcacagagacaatctagaggaattatcatcatttgaagcagtgaaaaaac  4498

Sero1RM       4492 TTGAAAAAAATAAGATTAGACGAAAGTTTTAAAGCATTCAATCAGGAATCT 4541
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4499 ttgaaaaaataagattagacgaaagttttaaagcattcaatcaggaatct  4548

Sero1RM       4542 AACCGTTTTTATAGCGCGACTTTAAGCTGTTATTTGGCATATGTAACATT  4591
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM          4549 aaccgtttttatagcgcgactttaagctgttatttggcatatgtaacatt  4598
```

```
Sero1RM          4592 TAAAAATTCTGATGCTGAAGAAATAATTGATAACCAAATTATTAGTGGTT    4641
                      |||||||||||||||||||||||||||.||||||||||||||||||..||
4bRM             4599 taaaaattctgatgctgaagaaataatggataaccaaattattagtactt    4648

Sero1RM          4642 TGAATAATTTTGAAAATCAAGCACAAGTGGTAAAGGAAAATAAATCAACT    4691
                      ||||||||||||||||||.|||||||||.|||||||||||||||||||||
4bRM             4649 tgaataattttgaaaatccagcacaagtagtaaaggaaaataaatcaact    4698

Sero1RM          4692 TATCTAGTTAGGGTGCCTAAAAGTAGAGCTAGTAAAAGAAAAAATAATAA    4741
                      |||.||||||||||||||||||||||||||||||||||||||||||||||
4bRM             4699 tatttagttagggtgcctaaaagtagagctagtaaaagaaaaaataataa    4748

Sero1RM          4742 TGTTTATTCGTATCCAAGAAACTATGTAGAGGCCATAGAAGCAAGAAGAA    4791
                      |||||||||.||||||||||||||||||||.|||||||||||||||||||
4bRM             4749 tgtttattcatatccaagaaactatgtagaagccatagaagcaagaagaa    4798

Sero1RM          4792 GATGTGGATGGATATGTGAATTTAATCAAGAGCATAAGACTTTTATTAAC    4841
                      |||||||||||.||||||||||||||||||||||||||||||||||||||
4bRM             4799 gatgtggatgggtatgtgaatttaatcaagagcataagacttttattaac    4848

Sero1RM          4842 ATTAAAGATGGTAAGCCTTATGTGGAGGGACATCATCTTATACCTATGGC    4891
                      |||||.|||||||||||||||||||||.||||||||||||||||||||||
4bRM             4849 attaaggatggtaagccttatgtggaaggacatcatcttatacctatggc    4898

Sero1RM          4892 AGTACAGGATTATTACGATAATActattgatttgcagataatataatct     4941
                      |||||||||||||||.||||||||.|||||||||||||||||||||||||
4bRM             4899 agtacaggattattatgataatacaattgatttgcagataatataatct     4948

Sero1RM          4942 gcctttgtccaacatgccatagaaaaattcattatgctgttcaaagtgag    4991
                      |.||||||||||||||.|||||||||||||||||||||||||||||||||
4bRM             4949 gtctttgtccaacatgtcatagaaaaattcattatgctgttcaaagtgag    4998

Sero1RM          4992 aaaaaagaaatgttaataaaaatttttaatgaaagggaaaaattttattt    5041
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             4999 aaaaaagaaatgttaataaaaatttttaatgaaagggaaaaattttattt    5048

Sero1RM          5042 acaacatggtatagagattgatagaaaaacacttttaagtttttatggta    5091
                      ||||||||||||||||||||||||||||||||||||||||||||||||||
4bRM             5049 acaacatggtatagagattgatagaaaaacacttttaagtttttatggta    5098

Sero1RM          5092 ttttttaa    5099
                      ||||||||
4bRM             5099 ttttttaa    5106
```

Sau3AI Endonuclease

```
# Length: 555
# Identity:     542/555 (97.7%)
# Similarity:   551/555 (99.3%)
# Gaps:           0/555 ( 0.0%)
# Score: 2837.0
#
#
#=======================================

Sero-4bENase       1 MKRIYDTKEKVHNRALEAIGKTLGEIDENNLKNVNNKSYPGNVIEQIWYD      50
                     :|||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1-ENase       1 LKRIYDTKEKVHNRALEAIGKTLGEIDENNLKNVNNKSYPGNVIEQIWYD      50

Sero-4bENase      51 HPADNISEPDFKEAGVELKVTPIDKQSKGNKDKKVTRLIAGERLVLNKIH     100
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1-ENase      51 HPADNISEPDFKEAGVELKVTPIDKQSKGNKDKKVTRLIAGERLVLNKIH     100

Sero-4bENase     101 YKNEYKKSFEESSFWHKNKVIELIQYYRRDTSDKKKFSSKELIEDKKQFK     150
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1-ENase     101 YKNEYKKSFEESSFWHKNKVIELIQYYRRDTSDKKKFSSKELIEDKKQFK     150

Sero-4bENase     151 IAYATLLSMVDLTDFNLPKDTVLEISDKDFEIIKQDWEKISKLINESKAE     200
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1-ENase     151 IAYATLLSMVDLTDFNLPKDTVLEISDKDFEIIKQDWEKISKLINESKAE     200

Sero-4bENase     201 ELSEGMTNYLGACTKAATGAEFTTQVGSEIKPKPRAYSFKTKFINELINT     250
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1-ENase     201 ELSEGMTNYLGACTKAATGAEFTTQVGSEIKPKPRAYSFKTKFINELINT     250

Sero-4bENase     251 QIIGNNHSAAINSIVKDANELKNNSLEEIIISRFLPFYPTNKKVWSQQDL     300
                     |||||:||||||||||||||||||||||||||||||||||||||||||||
Sero-1-ENase     251 QIIGNDHSAAINSIVKDANELKNNSLEEIIISRFLPFYPTNKKVWSQQDL     300

Sero-4bENase     301 IENFKIKTNEKSQKNLNNMIIRRILNLPKSKAEVTSEEIEKAEIRLKTIT     350
                     |||||||||.|||||||||||||||||||.||||||||||:|||||||||
Sero-1-ENase     301 IENFKIKTNRKSQKNLNNMIIRRILNLPTSKAEVTSEEIKKAEIRLKTIT     350

Sero-4bENase     351 LRDGKPKEHFKFQSIPSFEALVSENWEDSSVADLLDRTKFLLLVFNDLND     400
                     |||||||||||||:||||||:|||||||:|||||.|||||||||||||||
Sero-1-ENase     351 LRDGKPKEHFKFQNIPSFESLVSENWEESSVADFLDRTKFLLLVFNDLND     400

Sero-4bENase     401 KQPGKNTYETNPEKIFFVGAKFWNMPASDIYGPCKAVWKSDVDKLKKGVE     450
                     |.||||||||||||||||||||||||||||||||||||:|||||||||||
Sero-1-ENase     401 KHPGKNTYETNPEKIFFVGAKFWNMPASDIYGPCKAVWESDVDKLKKGVE     450

Sero-4bENase     451 LTYTKDSSGKVKILNNFIKPSLENVLHLRPGASKSQYNAPYYKTIIENGK     500
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1-ENase     451 LTYTKDSSGKVKILNNFIKPSLENVLHLRPGASKSQYNAPYYKTIIENGK     500

Sero-4bENase     501 EKKKYMNNSSKLPCNSKWINRPETEKDIYTDNYMVKQAWWLSKDYIFEQI     550
                     ||||||||||||||||||||||||:|||||||||||||||||||||||||
Sero-1-ENase     501 EKKKYMNNSSKLPCNSKWINRPESEKDIYTDNYMVKQAWWLSKDYIFEQI     550

Sero-4bENase     551 KDLLQ    555
                     ||||:
Sero-1-ENase     551 KDLLR    555


#---------------------------------------
```

Regulatory protein:

```
Length: 68
# Identity:      68/68 (100.0%)
# Similarity:    68/68 (100.0%)
# Gaps:           0/68 ( 0.0%)
# Score: 337.0
#
#
#=======================================

4b-C               1 MIRLRINEILKEREMSQKALCELSGLRPTTVSEMCRGVRTTVNLQHLETL     50
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1C            1 MIRLRINEILKEREMSQKALCELSGLRPTTVSEMCRGVRTTVNLQHLETL     50

4b-C              51 IDALEIEDFNQILKRDKV     68
                     ||||||||||||||||||
Sero-1C           51 IDALEIEDFNQILKRDKV     68
```

Methylase:

```
# Length: 469
# Identity:     460/469 (98.1%)
# Similarity:   466/469 (99.4%)
# Gaps:           0/469 ( 0.0%)
# Score: 2416.0#
#=======================================

4b-Met             1 MSLNIFSMFDGVGGFIVGLNDANEAIEKEMFRTTYSNQFEPSKKAQDAYE     50
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1Met          1 MSLNIFSMFDGVGGFIVGLNDANEAIEKEMFRTTYSNQFEPSKKAQDAYE     50

4b-Met            51 VGLYRFPEMNHIPDDIMTVSDNKFQEMHDAGVNMIVGGFPCQDYSVARSL    100
                     ||:|||||||||||||||||||||||||||||||||||||||||||||||
Sero-1Met         51 VGVYRFPEMNHIPDDIMTVSDNKFQEMHDAGVNMIVGGFPCQDYSVARSL    100

4b-Met           101 KDEQGIAGKKGVLFWEIIRAINKIQPKYLILENVDRLLKSPSTQRGRDFA    150
                     |||||||||||||||||||||.||||||||||||||||||||||||||||
Sero-1Met        101 KDEQGIAGKKGVLFWEIIRAIKKIQPKYLILENVDRLLKSPSTQRGRDFA    150

4b-Met           151 IMLGAFNKLGYSVEWRVINAAEYGRAQRRRRVFFFVYRNDTKWAKKIDEK    200
                     |||||||||||||||||||||||||||||||||||||||||||||:||||
Sero-1Met        151 IMLGAFNKLGYSVEWRVINAAEYGRAQRRRRVFFFVYRNDTKWAKRIDEK    200

4b-Met           201 YENEKLEEGLLEDVVIASNQYDDYIFQEGLFARQFPIENKIKYDGKRPRH    250
                     |||:|||||||||:|||||:||:|||||||||||||||||||||||||||
Sero-1Met        201 YENDKLEEGLLEDIVIASNRYDNYIFQEGLFARQFPIENKIKYDGKRPRH    250

4b-Met           251 ASHFLGKDADYDDYILNISDNFTGTVWNTGIMRHGRYFTADTTPIEEPPI    300
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1Met        251 ASHFLGKDADYDDYILNISDNFTGTVWNTGIMRHGRYFTADTTPIEEPPI    300

4b-Met           301 SLGTIVEKAKQSFISQYDNLDLGIKEYNSYIEQYVISDEIKIEKFKTLRG    350
                     ||||||||||||||.|||||||.|||||||||||||||||||||||||||
Sero-1Met        301 SLGTIVEKAKQSFIGQYDNLDLEIKEYNSYIEQYVISDEIKIEKFKTLRG    350

4b-Met           351 SKKIPRKRPDGSEYFYAEGAMSPYDSFDLPARTMLTSEASVNRSTHLLFE    400
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1Met        351 SKKIPRKRPDGSEYFYAEGAMSPYDSFDLPARTMLTSEASVNRSTHLLFE    400

4b-Met           401 NNKYRLITPIEAELLQDFPINWTKYKMNEETNEVSEVSDRMRYFFMGNAL    450
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1Met        401 NNKYRLITPIEAELLQDFPINWTKYKMNEETNEVSEVSDRMRYFFMGNAL    450

4b-Met           451 VTGIVKRIGIELAKIESHE     469
                     |||||||||||||||||||
Sero-1Met        451 VTGIVKRIGIELAKIESHE     469

#---------------------------------------
```

Conserved Domain Protein

```
# Length: 284
# Identity:     279/284 (98.2%)
# Similarity:   282/284 (99.3%)
# Gaps:           0/284 ( 0.0%)
# Score: 1468.0
#
#
#=======================================

4bConsDomPro       1 MYNFRVISANEDIFIKTKYRFISELEDFKEFAQTLTDDAGKAKVKSGKAS     50
                     :|||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1ConsDom      1 LYNFRVISANEDIFIKTKYRFISELEDFKEFAQTLTDDAGKAKVKSGKAS     50

4bConsDomPro      51 SYTRYLIRLIIFYCETHRDNLEELSSFEAVKKLEKIRLDESFKAFNQESN    100
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1ConsDom     51 SYTRYLIRLIIFYCETHRDNLEELSSFEAVKKLEKIRLDESFKAFNQESN    100

4bConsDomPro     101 RFYSATLSCYLAYVTFKNSDAEEIMDNQIISTLNNFENPAQVVKENKSTY    150
                     ||||||||||||||||||||||||:||||||.||||||.|||||||||||
Sero-1ConsDom    101 RFYSATLSCYLAYVTFKNSDAEEIIDNQIISGLNNFENQAQVVKENKSTY    150

4bConsDomPro     151 LVRVPKSRASKRKNNNVYSYPRNYVEAIEARRRCGWVCEFNQEHKTFINI    200
                     ||||||||||||||||||||||||||||||||||||:|||||||||||||
Sero-1ConsDom    151 LVRVPKSRASKRKNNNVYSYPRNYVEAIEARRRCGWICEFNQEHKTFINI    200

4bConsDomPro     201 KDGKPYVEGHHLIPMAVQDYYDNTIDFADNIICLCPTCHRKIHYAVQSEK    250
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Sero-1ConsDom    201 KDGKPYVEGHHLIPMAVQDYYDNTIDFADNIICLCPTCHRKIHYAVQSEK    250

4bConsDomPro     251 KEMLIKIFNEREKFYLQHGIEIDRKTLLSFYGIF    284
                     ||||||||||||||||||||||||||||||||||
Sero-1ConsDom    251 KEMLIKIFNEREKFYLQHGIEIDRKTLLSFYGIF    284
```