UNIVERSIDADE FEDERAL DE PELOTAS
Programa de Pós-Graduação em Ciência e Tecnologia Agroindustrial

Tese

COMPARAÇÕES FILOGENÔMICAS ENTRE CEPAS DE *Listeria monocytogenes* ISOLADAS DE DIFERENTES FONTES E REGIÕES GEOGRÁFICAS

Élen Silveira Nalério

Pelotas, 2009

ÉLEN SILVEIRA NALÉRIO

COMPARAÇÕES FILOGENÔMICAS ENTRE CEPAS DE *Listeria monocytogenes* ISOLADAS DE DIFERENTES FONTES E REGIÕES GEOGRÁFICAS

Tese apresentada ao Programa de Pós-Graduação em Ciência e Tecnologia Agroindustrial da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Doutor em Ciências (Microbiologia de Alimentos).

Orientador: Prof. Dr. Wladimir Padilha da Silva

Co-Orientador (es): Dr. Brendan William Wren
Dr. Richard Stabler

Pelotas, 2009

Banca examinadora:

Prof. Dr. Wladimir Padilha da Silva

Prof. Dr. Fábio Pereira Leivas Leite

Prof[a] Dr[a] Márcia Monks Jantzen

Prof[a] Dr[a] Amanda de Souza da Motta

À minha família
Mãe, Pai, Natália e Leonardo

AGRADECIMENTOS

A Deus, por estar sempre junto a mim, me guiando.

A minha família, mãe, pai, Natália e Léo, por serem a base de tudo e, incansavelmente, estarem ao meu lado, incentivando e torcendo pelos meus ideais.

A grande família Nalério, onde todos, mesmo que indiretamente, sempre estiveram presentes na minha vida me apoiando, principalmente, minha "irmã" Ivna, que desde a infância sonhou comigo o futuro.

As minhas queridas amigas Márcia e Karla que sempre me apoiaram e fizeram presentes de alguma forma.

Ao Paulo Henrique, que apesar de ter entrado na minha vida recentemente, tem me acolhido de forma muito especial e, com seu carinho, tem demonstrado o quanto acredita em mim.

Ao meu orientador e amigo, Wladimir Padilha da Silva, que desde a graduação me mostrou uma conduta profissional exemplar e, acima disso, me proporcionou crescer.

Aos amigos do Laboratório de Microbiologia de Alimentos Prof. Celso, Marcelo, Carol, Rodrigo, Andréia, Milena, Ana Eucares, Márcia Mata, Kátia, Denise, Júlia, Greici, Valmor, Simone e Janaina.

Aos professores e funcionários do Departamento de Ciência e Tecnologia Agroindustrial.

Aos amigos da London School of Hygiene and Tropical Medicine, especialmente Ozan Gudogdu, Suaad Al-Jaberi, Pornpan Pumirat e Pippa Strong pela grande amizade.

Ao Brendan Wren e Richard Stabler pela disponibilização de estrutura, material e de conhecimentos científicos, além disso pela confiança depositada em mim.

À Universidade Federal de Pelotas e ao Programa de Pós-Graduação em Ciência e Tecnologia por proporcionarem a realização do curso.

À CAPES (Coordenação de Apoio a Pesquisa em Nível Superior) pela concessão da bolsa de estudos no Brasil e no exterior.

À todas as pessoas que direta ou indiretamente, contribuíram para este estudo.

RESUMO

NALÉRIO, ÉLEN SILVEIRA**. COMPARAÇÕES FILOGENÔMICAS ENTRE CEPAS DE *Listeria monocytogenes* ISOLADAS DE DIFERENTES FONTES E REGIÕES GEOGRÁFICAS.** 2009. 65f. Tese (Doutorado). Programa de Pós-Graduação em Ciência e Tecnologia Agroindustrial. Universidade Federal de Pelotas, Pelotas.

Listeria monocytogenes é o agente causador da listeriose, uma infecção severa que pode cursar com sintomas que variam desde gastroenterites, meningites e até mesmo a morte. De fato, o desenvolvimento da doença pode ser relacionado a determinados sorotipos/linhagens das cepas de *Listeria*. Análises moleculares dos diferentes sorotipos/linhagens de *L. monocytogenes,* demonstraram que esta espécie é amplamente diversa, a qual pode ser agrupada em três linhagens. O estudo completo de genomas, baseado na técnica de *microarray*, pode ser empregado para estudar a relação filogenética entre cepas de *Listeria* tanto em nível de espécie, quanto em nível de sorotipo. Não obstante, a técnica de *microarray* visa evidenciar as diferenças no potencial patogênico e/ou adaptativo das cepas. O objetivo deste estudo foi a comparação filogenética entre cepas de *L. monocytogenes* isoladas de diferentes fontes. Foram analisadas 99 cepas de *L. monocytogenes* de diferentes origens geográficas (Brasil, Dinamarca, Áustria, Irlanda, Estados Unidos da América e fontes desconhecidas), incluindo cepas clínicas (humanas e animais), de alimentos e de indústrias alimentícias. O DNA das cepas teste foi hibridizado em DNA *microarrays* de *L. monocytogenes* baseado em seqüências do genoma de *L. monocytogenes* EGD-e. Os protocolos para marcação e hibridização do DNA seguiram as recomendações de Dorrell et al., (2001). A aquisição de dados, o processamento e as comparações filogenômicas foram realizadas conforme previamente descrito por Stabler et al. (2006). Comparações filogenômicas agruparam as cepas de *L. monocytogenes* em dois *clades* centrais, os quais são representativos das duas principais linhagens desta espécie. Além disso, cada um destes *clades* foram sub-divididos em mais dois sub-clades. A formação dos *clades* foi independente da origem geográfica das cepas, com exceção do *clade* contendo cepas persistentes (cepas que persistem no ambiente de processamento de alimentos), onde nenhuma cepa Brasileira esteve presente. Foram identificados 18 genes específicos para as cepas da linhagem I (sorotipos 1/2a e 1/2c). Esses genes são relacionados ao metabolismo de carboidratos, sistema regulatório *two component*, complexo de transporte ABC e aos genes *bvrB* e *bvrC*. A grande maioria das cepas persistentes se agrupou no mesmo *clade* pertencente à linhagem I. Foi obtido um conjunto de genes únicos pertencentes exclusivamente às cepas persistentes de *L. monocytogenes*, os quais sugerem serem os responsáveis pelo perfil adaptativo destas cepas. Os genes são envolvidos em resistência ao estresse e são relacionados ao transporte e metabolismo de carboidratos, processamento de informação ambiental, mecanismos de transdução de sinais, proteína de superfície celular, transporte e metabolismo de aminoácidos, transporte e metabolismo de nucleotídeos, tradução, biogênese de parede celular, replicação, recombinação e reparo, transporte de pequenas moléculas similar ao transportador ABC, metabolismo de lipídios e de função desconhecida. Dos 14 genes de virulência listados a

maioria deles esteve presente em todas as cepas de *L. monocytogenes* estudadas, com exceção dos genes *inlE* e *inlG*. Estes dados sugerem que, apesar das distintas origens de isolamento, a variabilidade genética das cepas de *L. monocytogenes* é direcionada para adaptação ambiental, ao invés da diferenciação visando virulência.

Palavras-chave: *Listeria monocytogenes*, comparações filogenômicas, genes de virulência, genes de persistência, microarray.

ABSTRACT

NALÉRIO, ÉLEN SILVEIRA. **PHYLOGENOMIC COMPARISONS BETWEEN *Listeria monocytogenes* STRAINS ISOLATED FROM DIFFERENT SOURCES AND GEOGRAPHIC REGIONS.** 2009. 65f. Tese (Doutorado). Programa de Pós-Graduação em Ciência e Tecnologia Agroindustrial. Universidade Federal de Pelotas, Pelotas.

Listeria monocytogenes is the causative agent of listeriosis which may cause a range of diseases from gastroenteritis, meningitis and death. In fact, disease outcome can be related to strain serotype/lineage thus molecular analyses has demonstrated that *L. monocytogenes* is a highly diverse species which can be grouped into three lineages. Whole-genome microarray can be employed to study phylogenetic relationships among *Listeria* strains either species or serotype level, in addition to demonstrate differences on their virulence potential and/or environmental adaptation. The aim of this study was the whole genome comparison of *L. monocytogenes* strains from different origins. Ninety-nine *L. monocytogenes* strains from different geographical origins (Brazil, Denmark, Austria, Ireland, USA and unknown), including clinical strains (humans and animals), food and food industries strains were analysed. DNA from all strains were competitively hybridized on to a *L. monocytogenes* DNA microarray based on the whole-genome sequence *L. monocytogenes* EGD-e. DNA labeling and hybridization protocol were followed according to Dorrell et al., (2001). Data acquisition, processing and comparative phylogenomics were performed as previously described by Stabler et al. (2006). Comparative phylogenomics clustered the *L. monocytogenes* strains into two central clades which is representative of the two main lineages of this species. In addition each of these clades were divided into two further subclades. Clade formation was independent of the geographical origin of strains with the exception of the clade containing persistent strains (strains that persist in food-processing environment), where none of the Brazilian strains were present. It was found 18 specific genes for lineage I strains (1/2a and 1/2c serotypes). These genes are related to carbohydrate metabolism, two component regulatory system, ABC transporter complex and *bvrB* and *bvrC* genes. Significantly all persistent strains clustered together in the same lineage I clade. We achieved a set of unique genes belonging exclusively to *L. monocytogenes* persistent strains pointing to be responsible for its adaptation profile. The genes are involved in stress resistance and are related to carbohydrate transport and metabolism, environmental information processing, signal transduction mechanisms, cell surface protein, amino acid transport and metabolism, nucleotide transport and metabolism, translation, cell wall biogenesis, replication, recombination and repair, transport of small molecules similar to ABC transporter, metabolism of lipids and unknown function. Interestingly from 14 virulence listed genes most of them were present in all studied *L. monocytogenes* strains with exception of *inlE* and *inlG* genes. These findings indicate that genetic variability of *L. monocytogenes* strains point to niche adaptation instead virulence differentiation despite of different origins. Persistent strains clustered suggesting genetic origin to survival in this environment.

Keywords: *Listeria monocytogenes,* phylogenomic comparisons, virulence genes, persistent genes.

Sumário

1 INTRODUÇÃO

As bactérias pertencentes ao gênero *Listeria* possuem a característica de serem difundidas na natureza, uma vez que estão presentes no solo, água, alimentos e na microbiota intestinal do homem e dos animais (SEELIGER, 1961; FARBER AND PETERKIN, 1991; ROCOURT & COSSART, 1997). *Listeria monocytogenes* é um microrganismo de grande interesse em saúde pública e também para as indústrias alimentícias (GUDMUNDSDÓTTIR et al., 2005), embora tenha sido reconhecido como um patógeno humano antes dos anos 70 (GOMBAS et al., 2003), somente nas três últimas décadas foi associado aos alimentos e classificado como patógeno de origem alimentar. É o agente causador da listeriose, uma infecção severa, a qual apresenta taxa de mortalidade entre 20-30% (DUSSURGET, PIZZARO-CERDA & COSSART, 2004; WALLS & BUCHANAM, 2005).

Alimentos contaminados são as maiores fontes de infecção de *Listeria*, e o trato gastrointestinal é o sitio primário de entrada da bactéria no organismo do hospedeiro (VASQUES-BOLAND et al., 2001). Qualquer cepa de *L. monocytogenes* pode ser considerada potencialmente patogênica para os humanos. Entretanto, várias observações sugerem que *L. monocytogenes* apresenta virulência heterogênea, em função de seus sorotipos, os quais estariam associados, de alguma forma, com o potencial de patogenicidade do microrganismo (JACQUET et al., 2002). Dos 13 sorotipos conhecidos, somente três (1/2a, 1/2b e 4b) têm sido relacionados a 90% dos surtos e casos esporádicos de listeriose (CABRITA et al., 2004), sendo as cepas do sorotipo 4b mais comumente isoladas em surtos, e cepas dos sorotipos 1/2a e 1/2b em casos esporádicos (FARBER & PERTERKIN, 1991; LIANOU et al., 2006).

Os avanços nas tecnologias de biologia molecular permitiram o desenvolvimento de técnicas eficazes para a análise em apenas, um único experimento, de um grande número de genes bacterianos (DORRELL, HINCHLIFFE & WREN, 2005). Assim, a técnica de *microarray* fornece uma revolução para estudos de diversidade genética de patógenos como *L. monocytogenes* (CALL, BORUCKI & LOGE, 2003; RUDI, KATLA & NATERSTAD, 2003; GARAIZAR, REMENTERIA & PORWOLLIK, 2006).

Os *microarray* são pequenos chips de DNA aderidos em lâminas onde milhares de seqüências de DNA ou cDNA são arranjados, representando genes conhecidos ou seqüências inteiras de um gene com todas suas possibilidades de mutação (BENEDETTI et al., 2000; DORRELL, HINCHLIFFE & WREN, 2005). Desta forma, em uma única lâmina é possível a deposição de milhares de seqüências, que são imobilizadas de forma organizada e em posições conhecidas (DORRELL, HINCHLIFFE & WREN, 2005; GARAIZAR, REMENTERIA & PORWOLLIK, 2006).

A aplicação de DNA *microarray* em subtipagem epidemiológica microbiana é ainda esporádica, este fato se deve, em parte, a pouca disponibilidade de genomas completos para a execução dos arranjos e também pelo elevado custo do método (GARAIZAR, REMENTERIA & PORWOLLIK, 2006). Estes micro-arranjos são baseados no genoma de uma cepa bacteriana completamente seqüenciada e em uma coleção de sondas de DNA fixadas em uma superfície sólida. Assim, em um único experimento de hibridização, o status de ausência\presença de todos os genes de um patógeno seqüenciado e de uma requerida cepa pode ser examinada (GARAIZAR, REMENTERIA & PORWOLLIK, 2006).

A disponibilidade do genoma completo de *L. monocytogenes* (GLASER et al., 2001) permitiu o desenvolvimento de DNA *microarray*. Com isso, muitas estratégias baseadas nesta técnica podem ser desenvolvidas para diferenciar *L. monocytogenes* entre seus sorotipos e linhagens filogenéticas (NITHINGALE et al., 2005; HAIN, STEINWEG & CHAKRABORTY, 2006). Além disso, seguindo a expansão de dados oriundos das análises de biologia molecular, uma série de aplicações práticas podem ser estabelecidas para esta técnica de *DNA microarray* (PANDA et al., 2003), principalmente, aquelas propondo contribuir para a elucidação dos mecanismos que determinam as diferenças

entre cepas epidemiologicamente relacionadas e também para compreender a adaptação, persistência e/ou virulência de *L. monocytogenes* (DOUMITH et al., 2006).

TÍTULO 1: Whole genome comparison of *Listeria monocytogenes* isolated from several countries and sources

Whole genome comparison of *Listeria monocytogenes* isolated from several countries and sources

Nalério, E. N.*; Silva, W. P.*; Stabler, R. A.; Wren, B. W.****

*Department of Agro-industrial Science and Technology, Laboratory of Food Microbiology, University Federal of Pelotas, Pelotas, Brazil

**Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom

1. ABSTRACT

L. monocytogenes is the causative agent of listeriosis which may cause a range of diseases from gastroenteritis to death. The disease outcome can be related to strain serotype/lineage thus molecular analyses has demonstrated that *L. monocytogenes* is a highly diverse species which can be grouped into three lineages. Whole-genome microarray can be employed to study phylogenetic relationships among *Listeria* strains either species or serotype level, in addition to demonstrate differences on their virulence potential and/or environmental adaptation. The aim of this study was to compare the whole genome of *L. monocytogenes* strains from different origins. Thus 99 *L. monocytogenes* strains from different geographical origins (Brazil, Denmark, Austria, Ireland, USA and unknown), including clinical strains (humans and animals), food and food industries strains were analysed, plus one *L. innocua* strain. The Bayesian algorithm allowed the clustering of *L. monocytogenes* strains into two central clades which is representative of the two main lineages of this species. The clustering were independent of the geographical origin of strains, the exception was the clade gathering the vast majority of probable persistent strains, where none of the Brazilian strains were present. It was found 18 specific genes for lineage I strains (1/2a and 1/2c serotypes). These genes are related to carbohydrate metabolism, two component regulatory system, ABC transporter complex and *bvrB* and *bvrC* genes. Interestingly from the 14 virulence listed genes most of them were present in every studied *L. monocytogenes* strain with exception of *inlE* and *inlG* genes. These results demonstrate that virulence genes involved in *L. monocytogenes* pathogenicity are conserved among the different strains.

Key-words: *Listeria monocytogenes,* microarray, genome comparison

2. INTRODUCTION

Listeria monocytogenes is the causative agent of listeriosis, which may cause a range of diseases from gastroenteritis to death (Schlech, 1998). Although listeriosis has low incidence in comparison with other food-borne diseases, it is a great public health concern due its mortality rate which is around 20-30% (Mead et al., 1999). Importantly, the epidemiology of listeriosis in Europe is changing and the incidence is increasing (Goulet et al., 2008). In the last five years the incidence has increased 59% in the European countries (EFSA, 2007). These increases have been reported in Belgium, Denmark, Germany, Finland, France and Switzerland (Goulet et al., 2008). In England and Wales morbidity rate has increased up to 44% in the patients for whom outcome data was available. The reason for this upsurge is still uncertain but it might be due in part to improvement in the surveillance network systems (Gillespie et al., 2006).

Interestingly, listeriosis has not been frequently reported in developing countries despite the fact that the bacterium has been widely isolated in food and food-processing plants (Lima et al., 2005; Nalério et al., 2009; Hellstorm et al., 2010; Myia et al., 2010). First impressions may suggest a divergence on virulence according to geographic regions; however this situation is probably due to inappropriate health surveillance networks. In Brazil, for instance, like many developing countries, notification of listeriosis is not compulsory therefore few has been reported as a consequence few official statistic data, which makes comparison to epidemiologic data from European countries, United States of America and Canada difficult. However in a study carried out in Brazil with 3,112 *Listeria* strains isolated from 1971 to 1997 reported that 74.9% of those strains were isolated from food, 9.5% from environment, 7.9% from human cases and 7.8% from animal cases. *L. monocytogenes* was isolated in 24.9% of all samples and in 96% of strains isolated from humans moreover most of those strains were related to pathological processes (Hofer et al., 2000).

Although nearly all *L. monocytogenes* strains have been considered pathogenic, it appears that are not all strains capable of developing disease. In

fact, disease outcome can be related to strain serotype/lineage. There are 13 recognised *L. monocytogenes* serotypes but only three (4b, 1/2a and1/2b) have been associated with approximately 95% of cases and outbreaks of listeriosis (Schuchat et al., 1991; Kathariou, 2002). Molecular analyses as like AFLP (Amplified Fragment Length Polymorphism) (Aarts et al., 1999), MEE (Multilocus Enzyme Electrophoresis) (Bibb et al., 1989) and PFGE (Pulsed-field Gel electrophoresis) (Brosh et al., 1994) demonstrated that *L. monocytogenes* is a highly diverse species and which can be grouped into three lineages. Lineage I (consisting of serotypes 1/2a, 3a, 1/2c and 3c), lineage II (1/2b, 3b, 4b, 4d, 4e and 7) and lineage III (4a and 4c) (Doumith et al., 2004a; Nightingale et al., 2005). This classification was additionally subdivided in 6 phylogenetic groups, lineage I.1 (serotypes 1/2a and 3a), I.2 (1/2c and 3c), II.1 (4b, 4d and 4e), II.2 (1/2b, 3b and 7), III.1 (4a) and III.2 (4c) (Doumith et al, 2004b). Furthermore it has been suggested that lineage I represent an environmental adapted lineage whilst lineage II strains represent a human host adapted lineage (Wiedmann, 2002).

Powerful and highly discriminatory methods are extremely necessary for understanding the divergences on epidemiology and evolutionary relationships about *L. monocytogenes.* Consequently microarray is efficient on comparing whole-genomes, allowing a deep insight into genetic content of large strain collections (Dorrell et al., 2005).Thus the construction of a whole-genome microarray (Hinds et al., 2002) to study this associations among *Listeria* strains either species or serotype level may fulfill this lacking of knowledge. Comparative phylogenomics employing Bayesian mathematical analyses is able to determine phylogenetics relationships, permitting robust approach into genetic relatedness of bacterial pathogens (Dorrel et al., 2005; Champion et al., 2005). Additionally this technique may be used as an alternative to characterise pathogenic isolates (Call et al., 2003a) by subtypification and also providing an understanding in the molecular phylogeny and evolution of *L. monocytogenes* strains giving information about transmission, pathogenesis and niche adaptation (Call et al., 2003b; Nightingale et al., 2005).

3. MATERIALS AND METHODS

Strains collection

The 100 strains used in this study were from diverse countries and sources, to date: 47 strains from Brazil, 10 from Denmark, 08 from Austria, 6 from Ireland, 1 from Unites States and 26 from unknown origins (Table 1). These strains were isolated from food (58), equipment and environment of industries (16), clinical cases both humans and animals (14), asymptomatic animal carriers (2), environment (6) and unidentified sources of isolation (5). On this collection is also included one *L. innocua* strain.

Genomic DNA (gDNA) extraction

The genomic DNA extraction was carried out with Wizard (Promega) purification kit and its quality and quantification was performed by 1% agarose gel electrophoresis and GeneQuant spectrophotometer (Amersham Pharmacia), respectively. The gDNA samples were kept at 4℃ to minimize damage which can be caused by freeze thawing.

Microarray Design

From 2944 annotated sequences of the whole-sequenced strain EGD-e (Glaser et al., 2001), 2855 were included in the design process. The *L. monocytogenes* arrays were developed and printed by BµG@S (Bacterial Microarray Group at St. Georges's Hospital Medical School), Department of Cellular and Molecular Medicine, St George's Hospital Medical School, London (Hinds et al., 2002). Briefly, ten pairs of gene-specific primers were designed to each sequence by applying PRIMER3 (Rozen & Skaletsky, 2000). Primers were designed to have a length of 20-25 base pairs (bp), melting temperature between 50℃ and 80℃ and amplicon ranging in size from 100 to 800 bp in length with an optimum size of 600 bp. The PCR primers were selected based on the BLAST similarity of the predicted PCR product to other genes on the microarray. Products with no similarity or least similarity were selected to ensure the least possible cross-hybridization on the microarray. The PCR products are spotted at a high density on poly-L-lysine-coated glass microscope

slide, using a MicroGridII (BioRobotics) arraying robot. Array design is available from ArrayExpress database (www.ebi.ac.uk/arrayexpress accession number A-BUGS-19).

DNA labelling and hybridization

For each microarray slide, 5µg of both test genomic DNA labelled with cyanine dye Cy3 Fluorolink™ (Amersham Bioscience) and *L. monocytogenes* EGD-e DNA labelled with cyanine dye Cy5 Fluorolink™ (Amersham Bioscience) was used in the microarray hybridization. The *L. monocytogenes* EGD-e is a standard strain (control) used as general reference for all hybridizations as previously described by Dorrell et al. (2001). Microarray slides were prehybridized into a Coplin jar within a solution which consist in 3.5 X SSC (1X SSC - 0.15M NaCl plus 0.015M Sodium Citrate), 0.1% SDS (Sodium Dodecyl Sulfate) and 10mg/mL BSA (Bovine Serum Albumin) at 65ºC for 20 minutes. After this the slides were rinsed thoroughly in distillate water followed for rinsing in isopropanol 1 minute each, centrifuged and kept way from light until the moment of hybridization.

After DNA labelling, test and control DNA, Cy3-labelled DNA and Cy5-labelled DNA, respectively, were mixed together and then purified with MinElute Purification Kit (Qiagen). Just before the hybridization the solution was mixed with 20X SSC and 2% SDS followed for denaturation at 95ºC for 2 minutes, applied on microarray slide and covered with 22 x 22mm LifterSlips (Eyrie Scientific), hybridized overnight sealed in a humidified hybridization chamber (Telechem International) and immersed in a bath at 65ºC for 16-20 hours. Afterwards, the slides were washed first in a pre-heated at 65ºC Wash A solution (1x SSC, 0.05%SDS and distillated water) for 2 minutes and then twice in the Wash B solution (0.06X SSC and distillated water), centrifuged and scanned.

Data acquisition, processing and comparative phylogenomics

The scanning was carried out in an Affymatrix 418 Array Scanner (MWG Biotech) and signal data were obtained through ImaGene 5.2 (BioDiscovery). All generated data were further analysed for whole-genome comparisons of

tests strains against the *L. monocytogenes* EGD-e strain by GeneSpring GX 7.3.1. (Agilent Technologies UK Limited).

The signal data transformation set measurements values below 0.01 were set to 0.01. The experiment was normalized to 50th percentile of all measurements in that sample which was calculated using all genes not marked absent. For each gene represented by a signal spot the light intensity was measured dividing by control channel value in each sample. If the control channel was below 0.01 then 0.01 was used instead. If the control channel and the signal channel were both below 0.01 then no data was reported.

The presence, divergence or absence of each gene of the array was determined by GACK analysis for cut-off assignment (GACK software) by transforming the normalized data from GeneSpring in log of ratio of raw and control data (Howard et al., 2006), since this software calculate an estimated probability of presence (EPP) value for each gene (Kim et al., 2002). The relationship of the strains was determined by Mr. Bayes v.3.0 software with a Bayesian method-based algorithm (Ronquist & Huelsenbeck, 2003), however before it the GACK output trinary format was transformed into binary data which means divergent or marginal genes were converted to present. The Bayesian model employ four-chain Markov chain Monte Carlo and 16-category gamma distribution with 1 million iterations to model the presence or absence rate heterogeneity per gene throughout the genome of strain as previously described by Howard et al. (2006) and Stabler et al. (2006). Treeview was the software used to view the tree generated.

4. RESULTS AND DISCUSSION

Overall microarray analysis

A broad range of 99 *L. monocytogenes* strains from diverse countries and sources were used in this research (Table 1). To date 47 strains from Brazil, 11 from Denmark, 09 from Austria, 06 from Ireland, 01 from United States of America and 25 unknown origins and also a *L. innocua* strain, totaling 100 strains. In this set was included 58 food strains, 16 equipment and

environment of industries, 16 clinical strains (12 from human cases and 2 from animal cases), 2 strains from asymptomatic animal carriers, 6 strains from environment and 05 unknown ones (Table 1). All strains were competitively hybridized with a *L. monocytogenes* DNA microarray based on the whole-genome sequence *L. monocytogenes* EGD-e (serotype 1/2a), which has 2855 target genes.

The Bayesian algorithm allowed the clustering of the *L. monocytogenes* strains into 2 central clades (Figure 1) which represent the two major lineages of this species, lineage I and II as previously described (Piffareti et al., 1989), proofing the huge power of this method for strain differentiation. Additionally our study provided concordant results as earlier reported in other phylogenetic studies of *L. monocytogenes* genome comparison (Call et al., 2003a; Zhang et al., 2003; Doumith et al, 2004b; Nightingale et al., 2005).

Furthermore the lineage II strains were sub-divided in further two main clades, the first (clade A) bearing in its huge part serotype 1/2b strains and the other clade gathering serogroup 4 (4b and 4e serotypes) strains. This scenario has been an indicative of clustering according to *L. monocytogenes* lineages sub-divisions, lineage II.1 (4b, 4d and 4e) and lineage II.2 (1/2b, 3b and 7) (Doumith et al., 2004b) as shown in the figure 1. Surprisingly the clade assembling lineage II.1 strains, representing serotypes 4b and 4e split the strains in additional two sub-clades, for easier understanding we called them clades B and C, respectively. Zhang and colleagues (2003) mentioned in their research that 1/2b, 3b and 4b strains were split in more two subgroups and it was congruent with serotype distribution. Meanwhile, in our study we report, based in the comparative Bayesian approach, that lineage II strains were divided in further three sub-groups which enhances the high power of strain differentiation of microarrays allied to robust algorithm.

Based in this information and by previous findings (Nightingale et al., 2005) we can infer which lineage II strains behave as sub-populations according to its serotype since the strains serotypes 1/2b, 4b and 4e were allocated almost exclusively in separated sub-clades (A, B and C). The highly degree of clonality demonstrated by lineage II strains may be indicative of a recent population bottleneck (Meinersmann et al., 2004).

Nevertheless this pattern was not observed in lineage I strains which did not demonstrate a clear serotype outline of clustering. Hence lineage I of *L. monocytogenes* represent a “generalist” lineage which is better adapted to survive and multiply in the environment while still maintaining the ability to cause human disease (Nighingale et al., 2004). Call et al. (2003b) in their assessments also found that 1/2a serotype strains were distributed among three clusters with one of them being a mix of both serotypes 1/2a and 1/2c. According to Bibb et al. (1990) serotype 1/2a strains have nearly twice as much genetic variation as other *L. monocytogenes* serotypes.

Additionally we can deduce that lineage I and II strains represent cohesive and separate subpopulations and this hypothesis is supported by the observations of Nightingale and colleagues (2005) which agreed that horizontal genes transfer of core genes predominantly occur between strains that belong to the same lineage. Recombination should be rare event between strains belonging to different *L. monocytogenes* genetic lineages (Salcedo et al., 2003), indicating that *L. monocytogenes* lineages represents species or subspecies-like entities, with distinct ecological preferences and then a deep separation involving lineage I and II strains appears to act as a barrier to the exchange of genetic information between these two lineages (Nightingale et al., 2005).

This hypothesis can be supported making an analogy with the findings of Doumith et al. (2004b). Which comparing genomes of a lineage I (serotype 1/2a) and lineage II (serotype 4b) strains of *L. monocytogenes* found that 8% of lineage II strain was missing in lineage I strain even though both being of the same species. Considering and comparing the genetic diversity between two different species of *Listeria*, *L. monocytogenes* EGD-e and *L. innocua,* which are about are 10.5%, we can suppose *L. monocytogenes* lineages could be a sub-species of this bacterium. As a consequence of this divergence are the evolutionary histories and population structure of these *L. monocytogenes* lineages. While lineage II strains appear to be highly clonal with limited horizontal gene transfer, lineage I strains demonstrate greater genetic diversity and evidence of a larger number of horizontal gene transfer events than lineage II strains (Nightingale et al., 2005).

Comparative phylogenomics of core genes

Aiming to know the chief divergence in the core genes of lineage I and II strains we generated a Venn diagram. This diagram was created through Genespring software analyses using only genes considered present (EPP of 100%) in the GACK analysis. Consequently it was possible to notice that lineage I and II strains share 45.5% (comprising 1293 core genes) of all 2855 genes present in the genome of sequenced *L. monocytogenes* EGD-e strain. The lineage I strains have 217 unique core genes in its genome while lineage II strains got 199 unique core genes.

The Genespring clustering analyses allowed the identification of specific regions present in *L. monocytogenes* lineages I and II strains. Thus it was found 18 specific genes for lineage I strains (Table 2 and Figure 2) which comprising five regions of *L. monocytogenes* genome, to date: lmo0734-lmo0739, lmo1060-lmo1063, lmo1968-lmo1971, lmo1973-lmo1974 and lmo2786-lmo2787. The functions of these genes are related to carbohydrate metabolism, two component regulatory system, ABC transporter complex and *bvrB* and *bvrC* genes. Some of these regions are considered as lineage-specific marker genes (Doumith et al., 2004b; Zhang et al., 2003) and were diagnosed like specific as well as in the lineage I strains in this study. Nelson et al. (2004) comparing whole-genome of serotype 4b (lineage II) and 1/2a strains (lineage I) found 83 genes restricted to the serotype 1/2a strains and 51 genes restricted to serotype 4b strains.

Noteworthy the genes *brvB* and *brvC* are part of a system that couples β-glucoside transport to catabolite repression of the PrfA-dependent virulence genes (Brehm et al., 1999). These genes are associated to down regulation of virulence genes in the presence of β-glucosides, which are found chiefly in plant tissues and seems to be an important feature in *L. monocytogenes* (Zhang et al., 2003) and according to Call et al. (2003b) this region is missing in lineage II strains, such information is in agreement to the data found in our study.

The lineage-specific marker genes for *L. monocytogenes* lineage II strains were only the regions lmo0466-lmo0467, lmo0469-lmo0471, lmo0151 and lmo1118-lmo1119. These regions code for hypothetical proteins and the only gene that has known function is lmo1119 which is a methylase. Maybe the complexity of identify specific markers for lineage II strains can be a failure of

this work. It may possibly be explained because the array employed in this study is based on the genome of *L. monocytogenes* EGD-e serotype 1/2a; turning the identification of lineage II strains specific genes complicated.

There are two simple explanations for the origin of the conserved lineage-specific and serotype-specific divergences in genome content of *L. monocytogenes.* First, the genes within the region of divergence (RD) may have been present in the most recent common ancestor of the two *L. monocytogenes* lineages and subsequently lost during divergence of lineage II. Alternatively the region could be acquired by lineage I strains, it is also possible that some can be acquired and some lost. These findings indicate that divergences between lineages I and II strains genome occurred through multiple gene acquisition and deletions events (Zhang et al., 2003). Also Nelson et al. (2004) suggested that the gene loss from a lineage ancestral to *L. monocytogenes* and *L. innocua* preceded the genomic diversification of *L. monocytogenes* into genomic lineages I and II. Consequently, such gene loss may have contributed to genomic improvements of *L. monocytogenes*, perhaps conferring high fitness to the pathogen.

Apart from the important divergence between the two main *L. monocytogenes* lineages, analyses identified remarkable genomic conservation within these major lineages and subgroups but variations between the different subgroups. These trails seem to mirror the evolution within the genus *Listeria* (Doumith et al., 2004b).

Comparative diversity among geographical origins, sources and serotypes of *L. monocytogenes* strains

In the clade A bearing in its biggest part serotype 1/2b strains were present strains isolated from food, equipment and environment of food industries; the only exception was the strain 1727 which is clinical strain isolated of an outbreak. It can be noted that almost every one of strains were originated from Brazil, maybe indicating epidemiologically linked strains as they were grouped as adjacent neighbours (Call et al., 2003b) in the phylogenetic tree (Figure 1).

However in the clade B gathering in its vast totality serotype 4b strains was present food and clinical strains from Brazil, Ireland and unknown origins. It can be evidence demonstrating that 4b strains besides its apparent virulence also kept features of adaptation in the environment. Food, industry equipment and asymptomatic animal strains were present into the clade C (majority serotype 4e strains), being almost exclusively grouped by strains from food origins.

Interestingly every persistent strain clustered into clade F was isolated from food and food-processing plants (industry environment). Highlighting the environment adaptation characteristic rather than the ability of cause disease of lineage I *L. monocytogenes* strains. Surprisingly this is the only clade where none Brazilian strains were present. However there was no evidence that geographical origin of strains was a factor determinant by which could contribute for this clustering since the persistent strains were isolated from European countries as Denmark, Austria and Ireland.

In this study, some analysed strains were studied previously by Jensen et al. (2008); where lineage I strains such as N53-1 and La111 were the less efficient causing mortality in fruit flies; 7418 and Scott A being the most efficient strains developing mortality. It is worth note that N53-1 and La111 are considered persistent strains by these authors and were also clustered in our persistent clade (clade F), while 7418 (1/2b) and Scott A (4b) both from the main clade of lineage II strains, related to disease outcome. Consequently, with these findings we may infer the real ability of strains from lineage II in causing disease, while the persistent strains are less capable to outcome disease but environmentally adapted.

Among the *L. monocytogenes* strains we also analysed a *L. innocua* strain and observing the phylogenetic tree is possible to note that *L. innocua* is clustered among serogroup 4 strains (clade C). This finding is in line with previous report (Zhang et al., 2003) that serotype 4b and *L. innocua* genome sequence are similar in some genomic regions. According to Doumith et al. (2004b) *L. innocua* evolved by successive gene loss from an ancestor of *L. monocytogenes* serogroup 4. The serogroup 4 of *L. monocytogenes* is closer to serogroup 6 of *L. innocua* than to *L. monocytogenes* serogroup 1/2 strains (Fiedler, 1988). There are two genes of *L. innocua* that are associated with

teichoic biosynthesis which were uniquely shared with *L. monocytogenes* strains of serogroup 4 and *L. welshimeri* (Doumith et al., 2004b).

Virulence genes

Virulence traits are usually considered the highest point in a pathogen. Consequently when is studied a bacterium as *L. monocytogenes*, considering its huge genetic diversity it is easy to believe that virulence genes should be quite variable as well. Nevertheless in our study from the 14 listed virulence genes (*inlA, inlB, inlC, inlI, inlG, inlH, inlE, inlJ (*precursor*),uhpT, prfA, plcA, plcB, hly* and *actA*) 85.7% were present in every studied *L. monocytogenes* strain (Figure 3). The exceptions were *inlE* and *inlG* genes, this last was absent mainly in lineage II strains. Interestingly there was no correlation with origin of isolation, serotype and/or geographical distribution within the presence of virulence genes. Despite of the conserved presence of virulence genes in *L. monocytogenes* strains, it is possible to note that gene expression regulation is the intriguing factor that trigger off the virulence profile and then the divergent profile of pathogenicity of *L. monocytogenes* strains.

Our results are in line with Doumith and colleagues (2004b) findings that assessing the genome content of 93 *L. monocytogenes* strains through microarray hybridizations found the virulence cluster (*prfA, plcA, hly, mpl, actA* and *plcB*), also *inlA*, *inlB, uhpT* and *bsh* genes present in all studied strains.

The internalin family proteins are surface proteins which are involved in the virulence process. In our study the internalin A was present in all strains independent from origin however only the presence is not enough to predict that strain is virulent. Based on current findings even whether persistent strains bear internalin A (*inl A*) gene, it does not mean they are virulent. Because single point mutations found in internalin genes from persistent strains could result in incorrect protein folding leading to a lower affinity to the E-cadherin (Jensen et al., 2008). The internalin G (*inlG*) was absent in the vast majority of lineage II strains, data that corroborate the findings of Doumith et al. (2004b) which reported that this gene in fact is specifically absent in lineage II strains.

According to Nelson et al. (2004) most of the *Listeria* virulence genes are conserved either in virulent or in less virulent *Listeria* strains and the presence

of these genes alone is not enough to explain the differences in virulence of any particular strain. Virulence is not a constant property and it has been speculated that it can be modulated, for instance, by components or conditions in processing of foods (reviewed by Jensen et al., 2008). For instance, in a recent research was reported that persistent RAPD type 9 strains had a lower virulence potential when compared to the clinical strains, with regards to the invasion of Caco-2 cells, killing fruits flies and nematodes worms, as well as faecal shedding and infection of guinea pig tissues (Jensen et al., 2008). Indicating which persistent strains although bearing virulence genes should keep their apparatus of virulence silent probably due some gene regulation.

Firstly our intriguing question regarding this study was whether strains isolated from different geographic regions and different sources would have similar gene content. Thus overall comparative phylogenomics of *L. monocytogenes* strains from Europe and Brazil have identified the two main lineages of this pathogen. Confirming that these lineages represent strains with similar genetic content independent from geographic area. Remarkably we identified a set of genes possible involved in the adaptation of strains in the environment of food industries however no significant divergences were noted among virulence genes.

5. REFERENCES

Aarts, H. J., Hakemulder, L. E., AVan Hoef., M. Genomic typing of *Listeria monocytogenes* strains by automated laser fluorescence analysis of amplified fragment length polymorphism fingerprint patterns. 1999. **International Journal of Food Microbiology.** 49:95–102.

Bibb, W. F., Schwartz, B., Gellin, B. G., Plikaytis, B. D., Weaver, R. E. Analysis of *Listeria monocytogenes* by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. 1989. **International Journal of Food Microbiology**. 8:233–239.

Bibb, W. F., Gellin, B. G., Weaver, R., Schwartz, B., Plikaytis, B. D., Reeves, M. W.,Pinner, R. W., Broome, C. V. 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. 1990. **Applied and Environmental Microbiology.** 56:2133-2141

Brehm, K., Ripio, M. T., Kreft, J., Vasquez-Boland, J.-A. The *bvr* locus of *Listeria monocytogenes* mediates virulence gene expression by β-glucosides. 1999. **Infection and Immunity.** 181: 5024-5032.

Brosch, R., Chen, J., Luchansky, J. B. Pulsed-Field Fingerprinting of listeriae: identication of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. 1994. **Applied and Environmental Microbiology.** 60: 2584-2592.

Call, D. R., Borucki, M. K., Loge, F. J. Detection of bacterial pathogens in environmental samples using DNA microarrays. 2003a. **Journal of Microbiological Methods.** 53:235-243.

Call, D. R., Borucki, M. K., Besser, T. E. Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of *Listeria monocytogenes.* 2003b. **Journal of Clinical Microbiology.** 41:632-639.

Champion, O. L., Gaunt, M. W., Gundogdu, O., Elmi, A., Witney, A. A., Hinds, J., Dorrel, N., Wren, B. W. Comparative phylogenomics of the food-borne pathogen *Campylobacter jejuni* reveals genetic markers predictive of infection source. 2005. **Proceedings of the National Academy of Sciences**. 102:16043-16048.

Dorrell, N., Mangan, J. A., Laing, K. G., Hinds, J., Linton, D., Al-Ghusein, H., Barrell, B. G., Parkhill, J., Stoker, N. G., Karlyshev, A. V., Butcher, P. D., Wren, B. W. Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. 2001. **Genome Research.** 11:1706–1715.

Dorrell, N., Hinchliffe, S. J., Wren, B. W. Comparative phylogenomics of the pathogenic bacteria by microarray analysis. 2005. **Current Opinion in Microbiology.** 8:620-626.

Doumith, M., Buchrieser, C., Glaser, P., Jacquet, C. and Martin, P. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. 2004a. **Journal of Clinical Microbiology.** 42:3819-3822.

Doumith, M., Cazalet, C., Simoes, N., Frangeul, L., Jacquet, C., Kunst, F., Martin, P., Cossart, P., Glaser, P., Buchrieser, C. New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. 2004b. **Infection and Immunity.** 72:1072-1083.

Dumas, E., Meunier, B., Berdagué, J.-L., Chambon, C. Comparative analysis of extracellular and intracellular proteomes of *Listeria monocytogenes* strains reveals a correlation between protein expression and serovar. 2008. **Applied and Environmental Microbiology.** 74:7399-7409.

European Food Safety Authority (EFSA). 2007. The community summary report on trends and sources of zoonoses, zoonotic agents, antimicrobial

resistance and foodborne outbreaks in the European Union in 2006. Euro Food Safety Authority. J. 130:1-310.

Fiedler, F. Biochemistry of the cell surface of *Listeria* strains: a locating general view. 1988. **Infection.** 16:92-97.

Gillespie, I. A., McLauchlin, J., Grant, K. A., Little, C. L., Mithani, V., Penman, C., Lane, C.; Regan, M. 2006. **Emerging Infectious Diseases**. 12:1361-1366.

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

Goulet, V., Hedberg, C., Le Monnier, A., Valk, H. Increasing incidence of listeriosis in France and other European countries. 2008. **Emerging Infectious Diseases**. 14: 734-740.

Hinds, J., Laing, K. G., Mangan, J. A., Butcher, P. D. Microarrays for microbes: the BμG@S approach. 2002. **Comparative and Functional Genomics**. 3:333-337.

Hofer, E., Ribeiro, R., Feitosa, D. P. Species and serovars of the genus *Listeria* isolated from different sources in Brazil from 1971 to 1997. 2000. **Memórias do Instituto Oswaldo Cruz.** 95:615-620.

Howard, S. L., Gaunt, M. W., Hinds, J. Witney, A. A., Stabler, R., Wren, B. W. Application of comparative phylogenomics to study the evolution of *Yersinia enterocolitica* and to identify genetic differences relating to pathogenicity. 2006. **Journal of Bacteriology**. 188:3645-3653.

Jensen, A., Thomsen, L. E., Jorgensen, R. L., Larsen, M. H., Roldgaard, B. B., Christensen, B. B., Vogel, B. F., Gram, L., Ingmer, H. Processing plant persistent strains of *Listeria monocytogenes* appear to have a lower virulence potential than clinical strains in selected virulence models. 2008. **International Journal of Food Microbiology.** 123:254-261.

Kathariou, S. *Listeria monocytogenes* virulence and pathogenicity, a food safety perspective. 2002. **Journal of Food Protection.**, 5: 1811-1829.

Kim, C. C., Joyce, E. A., Chan, K., Falkow, S. Improved analytical methods for microarray-based genome-composition analysis. 2002. **Genome Biology.** 3:RESEARCH0065.

Lima, A. S., Laer, A. E., Trindade, P. S., Silva, W. P. Disseminação de *Listeria monocytogenes* no processamento de lingüiça mista frescal avaliada por sorologia e RAPD. 2005. **Alimentos e Nutrição.** 16:245-251.

Mead, P. S., L. Slutsker, V. Dietz. L. F. McCaig, J. S. Bresee, C. Shapiro, P. M. Griffin; Tauxe, R. V. 1999. Food-related illness and death in the United States. **Emerging Infectious Diseases**. 5:607-625.

Meinersmann, R. J., Phillips, R. W., Wiedmann, M., Berrang, M. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes reveals clonal and recombination histories of three lineages. 2004. **Applied and Environment Microbiology.** 70:2193-2203.

Miya, S., Takahashi, H., Ishikawa, T., Fujii, T., Kimura. Brisk of *Listeria monocytogenes* contamination of raw ready-to-eat seafood products available at retail outlets in Japan. 2010. **Applied and Environmental Microbiology.** 76:3383-3386.

Nalério, E. N., Araújo, M. R., Mendonça, K. S., Bassani, M. T., Silva, W. P. *Listeria monocytogenes*: monitoramento desse perigo biológico na cadeia produtiva de frangos do sul do Rio Grande do Sul. 2009. **Ciência e Tecnologia de Alimentos.** 29.

Nelson, K. E., Fouts, D. E., Mongodin, E. F., Ravel, J., DeBoy, R. T., Kolonay, J. F., Rasko, D. A., Angiuoli, S. V., Gill, S.R., Paulsen, I. T., Peterson, J., White, O., Nelson, W. C., Nierman, W., Beanan, M. J., Brinkac, L. M., Daugherty, S. C., Dodson, R. J., Durkin, A. S., Madupu, R., Haft, D. H., Selengut, J., Aken, S. V., Khouri, H., Fedorova, N., Forberger, H., Tran, B., Kathariou, S., Wonderling, L. D., Uhlich, G. A., Bayles, D. O., Luchansky, J. B., Fraser, C. M. 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. 2004. **Nucleic Acids Research.** 32:2386-2395.

Nightingale, K. K., Schukken, Y. H. Nightingale C. R., Fortes, E. D., Ho, A. J. Her, Z. Grohn, Y. T., McDonough, P. L. Wiedmann. M. 2004. Ecology and transmission of *Listeria monocytogenes* in ruminants and the farm environment. **Applied and Environment Microbiology.** 70:4458–4467

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

Piffaretti, J. C., Kressebuch, H., Aeschbacher, M., Bille, J., Bannerman, J. M., Selander, R. K., Rocourt, J. 1989. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. **Proceedings of the National Academy of Sciences**. 86:3818-3822.

Ronquist, F.; Huelsenbeck, J. P. MrBayes3: Bayesian phylogenetic inference under mixed models. 2003. **Bioinformatics.** 19:1572-1574.

Rozen, S. & Skaletsky, H**.** 2000. **Methods in Molecular Biology**.132, 365-386.

Salcedo, C., Arreaza L., Alcala, B., de la Fuente L., Vazquez J. A.. 2003. Development of a multilocus sequence typing method for analysis of *Listeria monocytogenes* clones. **Journal of Clinical Microbiology.** 41:757–762.

Schlech, W. F. 1998. Foodborne listeriosis. **Clinical Infectious Disease.** 31:770-775.

Shuchat, A., Swaminathan, B. and Broome, C. 1991. Epidemiology of human listeriosis. **Clinical Microbiology Reviews.** 4:169-183.

Stabler, R. A., Gerding, D. N., Songer, J. G., Drudy, D., Brazier, J.S., Trinh, H. T., Witney, A. A., Hinds, J., Wren, B. W. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. **Journal of Bacteriology**. 2006. 188:7297-7305.

Hellström, S**.,** Laukkanen, R**.,** Siekkinen, K. M**.,** Ranta, J**.,** Maijala, R**.,** Korkeala, H**.** *Listeria monocytogenes* contamination in pork can originate from farms. **Food Protection.** 2010. 73:641-648.

Wiedmann, M. 2002. Molecular subtyping methods for *Listeria monocytogenes*. **Journal of AOAC International.** 85:524-531.

Zhang, C., Zhang, M., Ju, J. Nietfeltdt, J., Wise, J., Terry, P. M., Olson, M., Kachmann, S. D., Wiedmann, M., Samadpour, M., Benson, A. K. Genome diversification in phylogenetic lineages I and II of *Listeria monocytogenes*: identification of segments unique to lineage II populations. 2003. **Journal of Bacteriology.** 185:5573-5584.

Table 1 – Identification and information about strains

Strain	Origin	Year of isolation	Group	Serotype/ Serogroup	Lineage	City -Country
19	Cheese	2007	Food	-	-	Cork - Ireland
30	Dry faeces – cow	-	Animal asymptomatic	4b	II-1	Cork - Ireland
92	Cheese	-	Food	-	-	Cork - Ireland
411	Unknown	-	Unknown	-	-	Cork - Ireland
535	Vacherin Cheese	1996	Food	-	-	Unknown
758	Producer B – smear	1997	Environment – Persistent	-	-	Vienna - Austria
1454	Clinical – animal (cow – mastitis)	1997	Animal	1/2a	I-1	Vienna - Austria
1458	Producer A – smear	1997	Environment – non-persistent	-	-	Vienna - Austria
1727	Outbreak	1998	Human	-	-	Atlanta, CDC - USA
2315	Producer B – Cheese	1999	Food- Persistent	-	-	Vienna - Austria
2540	Producer A - Cheese rasp	1999	Food – Persistent	-	-	Vienna - Austria
2566	Producer B – smear	1999	Environment - non-persistent	-	-	Vienna - Austria
4239	Human case (52 year-old male)	1998	Human	1/2a	I-1	Unknown
4359	Producer A - water	2004	Food – Persistent	-	-	Vienna - Austria
4423	Producer B – smear	2004	Environment - Persistent	-	-	Vienna - Austria
4446	Clinical – Human case (63 year-old female)	1998	Human	4b	II-1	Unknown
4542	Clinical – Human case (84 year-old male)	1998	Human	4b	II-1	Unknown
4810	Clinical – Human case (84 year-old male)	1998	Human	4	II	Unknown
4898	Producer A - water	2007	Food – Persistent	-	-	Vienna - Austria
4905	Human case (70 year-old female)	1998	Human	1/2	-	Unknown
6179	Cheese	2000	Food – Persistent	1/2a	I-1	Cork - Ireland
6895	Ham	1998	Food	1/2	II	Unknown
7291	Pasta with chicken	1998	Food	4b	II-1	Unknown
7418	Spreadable sausage	1998	Food – Sporadic	1/2b	II-2	Unknown
2V903b	Smokehouse V - cold-smoked salmon - final product - plant 1	1999	Food	4	I	Denmark
3R23-1	Smokehouse 1 – steel	2003	Industry Environment	-	II	Denmark
Ag28	Seawater fish farm 1 - stone with algae	2004	Environment - Sporadic	-	-	Unknown
Ap143	Slaughterhouse 2 - mud from bleeding area	2004	Industry Environment	-	-	Unknown
Br21	Fresh water fish farm 2 - water from detesting tank	2004	Environment - Persistent	1/2a	I-1	Unknown
Br22	Fresh water fish farm 2 - sediment from detesting tank	2004	Environment	-	-	Unknown
BRA 1	Rubber plucker fingers of defeathering machine	2005	Industry Environment	4b	II-1	Pelotas - Brazil

Table1 - Continuation

BRA 102	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas - Brazil
BRA 104	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas - Brazil
BRA 107	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas - Brazil
BRA 108	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas - Brazil
BRA 10s	Pork sausage	-	Food	1b	-	Pelotas - Brazil
BRA 111	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas - Brazil
BRA 11o	Sheep carcass	-	Food	4b	II-1	Pelotas - Brazil
BRA 12o	Sheep carcass	-	Food	1/2c	I-2	Pelotas - Brazil
BRA 13s	Pork sausage	-	Food	4b	II-1	Pelotas - Brazil
BRA 18o	Sheep carcass	-	Food	1/2b	II-2	Pelotas - Brazil
BRA 1s	Pork sausage	-	Food	4b	II-1	Pelotas - Brazil
BRA 20	Evisceration floor	2006	Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 21o	Sheep carcass	-	Food	4b	II-1	Pelotas - Brazil
BRA 22	Washing water prechiller	2006	Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 22o	Sheep carcass	-	Food	1/2c	I-2	Pelotas - Brazil
BRA 22q	Cottage cheese	-	Food – retail	1/2a	I-1	Pelotas - Brazil
BRA 23q	Cottage cheese	-	Food – retail	4b	II-1	Pelotas - Brazil
BRA 24q	Cottage cheese	-	Food – retail	1/2b	II-2	Pelotas - Brazil
BRA 25s	Handling (pork sausage)	-	-	1c	-	Pelotas - Brazil
BRA 27	Washing water chiller	2006	Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 2o	Sheep carcass	-	Food	4b	II-1	Pelotas - Brazil
BRA 32	Conveyor cut belt	2006	Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 42	Chicken Final product	2006	Food - Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 45	Washing water chicken from evisceration	2006	Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 46	Evisceration floor	2006	Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 53	Conveyor cut belt	2006	Industry Environment	1/2b	II-2	Pelotas - Brazil
BRA 58	Chilled chicken	2006	Food - retail	4e	II-1	Pelotas - Brazil
BRA 63	Chilled chicken	2006	Food -retail	4e	II-1	Pelotas - Brazil
BRA 6o	Sheep carcass	-	Food	3c	I-2	Pelotas - Brazil
BRA 71	Chilled chicken	2006	Food - Retail	4e	II-1	Pelotas - Brazil
BRA 74	Chilled chicken	2006	Food - Retail	4e	II-1	Pelotas - Brazil
BRA 76	Chicken Final product	2006	Food – Industry Environment	4e	II-1	Pelotas - Brazil
BRA 77	Chilled chicken	2006	Food – Retail	4e	II-1	Pelotas - Brazil
BRA 82	Chilled chicken	2006	Food – Retail	4e	II-1	Pelotas - Brazil
BRA 86	Chicken cloacae swab (faeces)	2006	Animal asymptomatic	4e	II-1	Pelotas - Brazil
BRA 88	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas - Brazil

Table 1 - Continuation

BRA 8q	Cottage cheese	-	Food – Retail	1/2c	I-2	Pelotas - Brazil
BRA 8s	Pork sausage	-	Food – Retail	1c	-	Pelotas - Brazil
BRA 91	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas - Brazil
BRA 94	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas - Brazil
BRA 98	Chilled chicken	2006	Food – Retail	1/2c	I-2	Pelotas - Brazil
BRA 99	Chilled chicken	2006	Food – Retail	-	-	Pelotas - Brazil
BRA 9o	Sheep carcass	-	Food	1/2a	I-1	Pelotas - Brazil
BRA St07	Clinical – human	-	Human	1/2b	II-2	Brazil
BRA St08	Clinical – human	-	Human	1/2a	I-1	Brazil
BRA St247	Clinical – human	-	Human	4b	II-1	Brazil
EDG	Animal case – rabbit	1926	Animal - Standard strain	1/2a	I-1	Unknown
H025	Fresh water fish farm 1 - surface of rainbow trout	2004	Environment	-	-	Unknown
L. innocua	-	-	-	6a	-	Unknown
La111	Smokehouse G - cold-smoked salmon	1996	Food – Industry environment	1/2a	I-1	Denmark
La22	Smokehouse V - cold-smoked salmon - final product	1995	Food - Industry environment	1/2a	I-1	Denmark
11994	Unknown	-	Unknown	-	-	Unknown
Lm 7973	Unknown	-	Unknown	-	-	Unknown
Lo28	Clinical - Pregnant woman	-	Human Faeces	1/2c	I-2	Unknown
M103-1	Slaughterhouse A - rubber conveyor belt	2003	Industry Environment	1/2a	I-1	Denmark
N22-2	Smokehouse 2 - floor in slicing area	2002	Sporadic - food industry environment	-	-	Unknown
N53-1	Smokehouse 2 - scraper mat	2002	Industry Environment	1/2a	I-1	Denmark
No37-1	Smokehouse 3 - waste container (pet food)	2001	Industry Environment	-	II	Denmark
No40-1	Smokehouse 3 - floor swap	2003	Industry Environment	1/2a	I-1	Denmark
R416a	Smokehouse 1 - slicing machine - processing plant 2	1996	Industry Environment	-	-	Denmark
R479a	Smokehouse 1 - cold-smoked salmon - final product - processing plant 2	1996	Food – Persistent	1/2a	I-1	Denmark
S2583	Lettuce processing equipment	1998	Industry Environment	1/2	II	Unknown
Scott A	Human	-	Human – Standard strain	4b	II-1	Unknown
SVS2719	Animal case – goat	1998	Animal	1/2	II	Unknown
SVS2827	Ensilage	1998	Food	1/2	II	Unknown
V518a	Smokehouse V - cold-smoked salmon- final product slicing mach. 2 - proc plant 1	1998	Food - Industry Environment	4b	II-1	Denmark
Vo14	Fresh water fish farm 3 - grass and plants around the race-way-system	2004	Environment	-	-	Unknown
Vo26	Fresh water fish farm 3 - sediment from incoming stream	2004	Environmentl	-	-	Unknown
29	Unknown	-	Unknown	-	-	Cork - Ireland

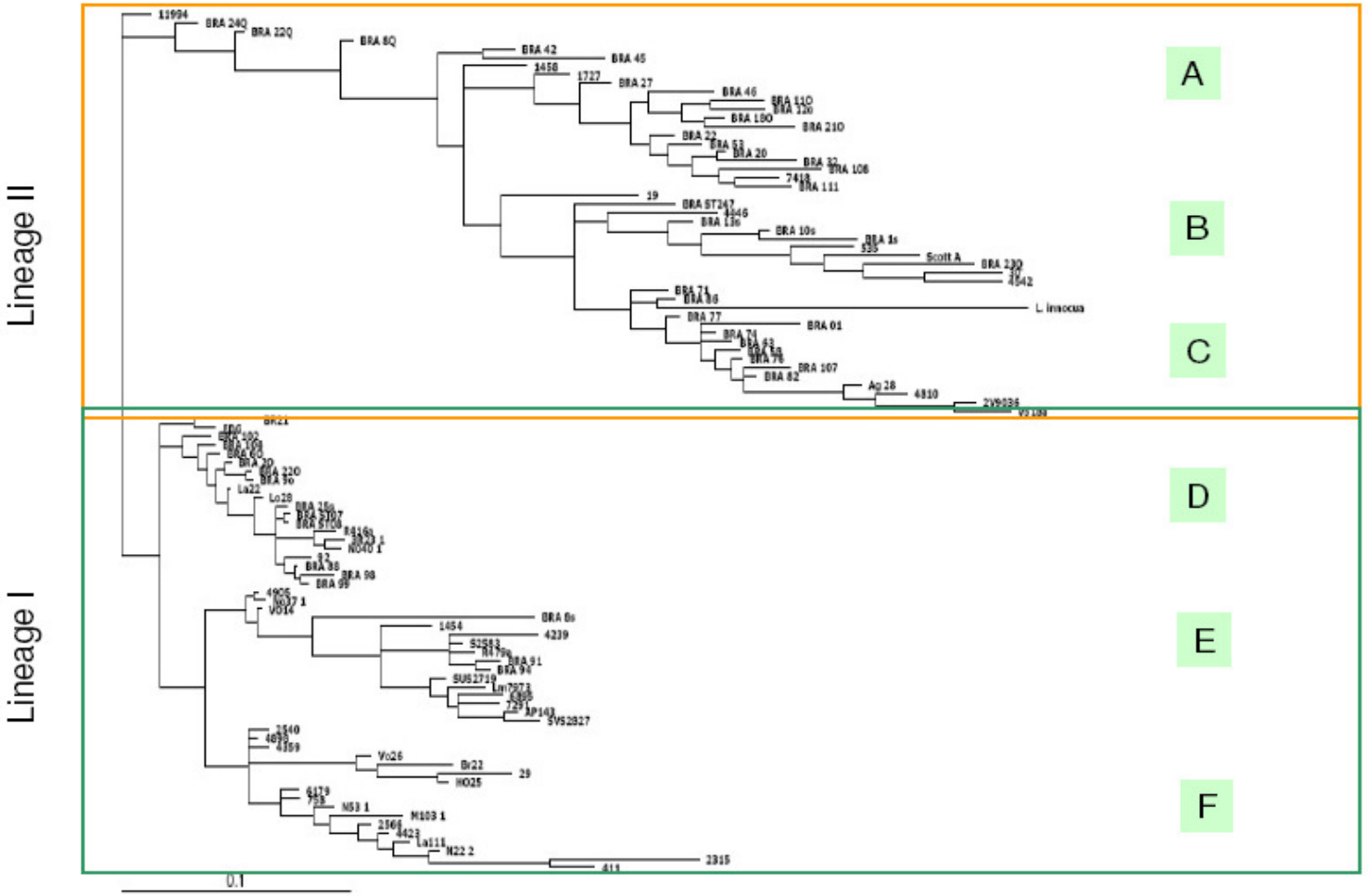

Figure 1 – Phylogenetic tree showing the genetic relationship among *L. monocytogenes* strains. A – clade of majority 1/2b strains; B – clade of majority 4b strains; C – clade of majority 4e strains; D, E and F – clade of majority 1/2a.

Table 2 - Specific-gene markers of *L. monocytogenes* lineage I strains

Gene ID	***Gene function***
Lmo0734	Similar to transcriptional regulator (LacI family)
Lmo0735	Similar to ribulose 5- phosphate 3-epimerase
Lmo0736	Similar to ribose 5-phosphate isomerise
Lmo0737	Conserved hypothetical protein
Lmo0738	Similar to phosphotransferase system (PTS) beta-glucoside-specific enzyme II ABC component
Lmo0739	Similar to 6-phospho-beta-glucosidase
Lmo1060	Similar to transcription response regulator
Lmo1061	Similar to two-component sensor histidine kinase
Lmo1062	Similar to ABC transporters (permease protein)
Lmo1063	Similar to ABC transporter (ATP binding protein)
Lmo1968	Similar to creatinine amidohydrolases
Lmo1969	similar to 2-keto-3-deoxygluconate-6-phosfate aldolase
Lmo1970	similar to putative to phosphotriesterase related protein
Lmo1971	Similar to pentitol PTS system enzyme II C component
Lmo1973	Similar to PTS system enzyme II A component
Lmo1974	Similar to transcriptions regulators (GntR family)
Lmo2786	*bvrC*
Lmo2787	*bvrB* – beta-glucoside-specific phosphotransferase enzyme II ABC component

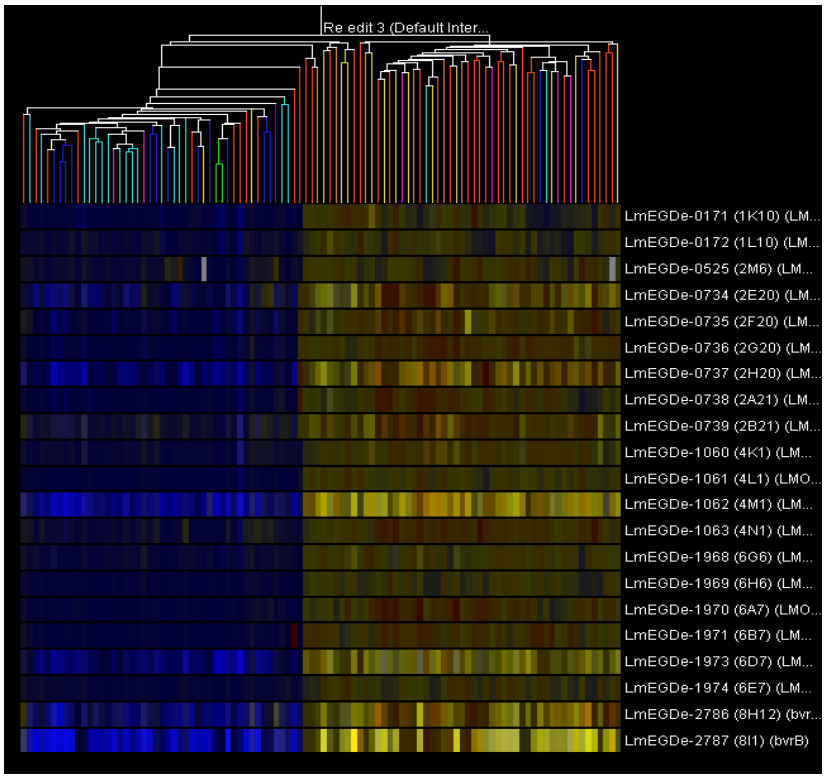

Figure 2 – Genespring clustering of lineage I and II strains. Yellow bars represent present genes and blue bars represent absent genes in the strains.

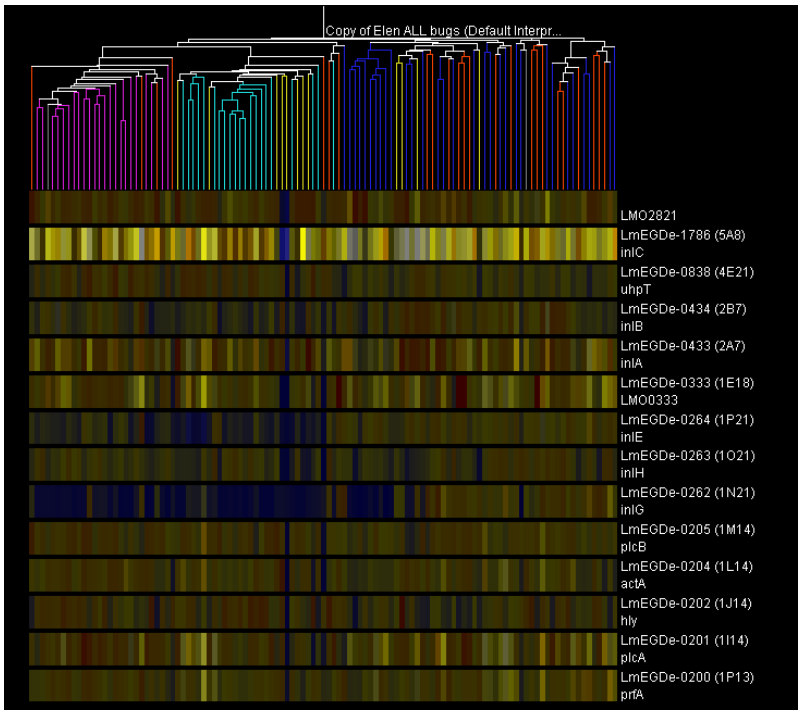

Figure 3 – GeneSpring clustering of virulence genes of *L. monocytogenes*, yellow bars indicating present genes and blue bars absent genes.

TÍTULO 2: Comparative phylogenomics of *Listeria monocytogenes* reveals specific genes related to adaptation profile

Comparative phylogenomics of *Listeria monocytogenes* reveals specific genes related to adaptation profile

Nalério, E. S.*; Silva, W. P.*; Stabler, R. A.; Wren, B. W.****
*Department of Agro Industrial Science and Technology, Laboratory of Food Microbiology, University Federal of Pelotas, Pelotas, Brazil
**Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom

1. ABSTRACT

Listeria monocytogenes is a food-borne pathogen ubiquitous in the environment that often adapts and overcomes a range of hostile conditions. The peaces that make *L. monocytogenes* machinery of adaptation works is still a reason for several researches intending to bring the light on this peculiar feature of the pathogen. Therefore microarray assay is capable of accomplish this task comparing whole-genomes, allowing a deep insight into genetic content of large strain collections, trying to decipher questions about adaptation and virulence profile. We performed DNA microarray analyses of 99 *L. monocytogenes* strains from diverse countries and sources plus one *L. innocua* strain. The Bayesian algorithm allowed the clustering of *L. monocytogenes* strains into two central clades which is representative of the two main lineages of this species. The clustering were independent of the geographical origin of strains, the exception was the clade gathering the vast majority of persistent strains, where none of the Brazilian strains were present. With bioinformatics approach we achieved in the first time a set of unique genes (lmo0105, lmo0137, lmo0333, lmo0381, lmo0461, lmo1286, lmo1467, lmo1642, lmo1843, lmo1851, lmo1969, lmo1970, lmo1991, lmo2155, lmo2192, lmo2442, lmo2696) belonging exclusively to *L. monocytogenes* persistent strains pointing to be responsible for its adaptation profile. This information may supply the lacking of understanding about how *L. monocytogenes* gets to trickle environment harsh conditions aiming its survival.
Keywords: *Listeria monocytogenes*, DNA microarray, adaptation profile

2. INTRODUCTION

Listeria monocytogenes is a bacterium ubiquitous in the environment (Rocourt, 1999) and is the causative agent of listeriosis a severe infection acquired generally due to contaminated food consumption (Rocourt, 1996). This food-borne pathogen frequently adapts and overcomes various forms of hostile conditions during its transmissions for foods and related environments (Tasara & Stephan, 2007). Although *L. monocytogenes* is not able to form spores and has no capsule, it can survive in adverse conditions, found both in nature and in the food chain, due to its specific metabolic features (Vasquez-Boland et al., 2001).

L. monocytogenes strains persistent in food-processing environments are more prone to be food products contaminants than non-persistent strains (Jensen et al., 2008). In addition study has demonstrated that this microorganism is able to persist in the food-processing environment for several months and even up to 10 years (Ojeniyi et al., 2000). It has been reported that food contamination by *L. monocytogenes* increases up to high levels during food processing (Autio et al., 1999; Wulff et al., 2006)

Considerable diversity is present among different strains of *L. monocytogenes* then clonal expansion of highly virulent subpopulations of the pathogen may exist (Doumith et al., 2004), which make *Listeria* strains demonstrates differences on their virulence potential (Brosch et al., 1994; Ragon et al., 2008) and environmental adaptation skills (Nightingale et al., 2004).

There are several methods widely employed to provide information trying to solve the questions still uncertain about this pathogen (Hain et al., 2006). In addition powerful and highly discriminatory methods are extremely necessary for understanding the divergences on epidemiology and also evolutionary relationships about *L. monocytogenes*. Therefore microarray assay is capable of accomplish this task comparing whole-genomes, allowing a deep insight into genetic content of large strain collections (Dorrell et al., 2005).

Furthermore the microarray technique may be used as an alternative to characterise pathogenic strains (Call et al., 2003) by sub-typification and also providing an understanding in the molecular phylogeny and evolution of *L.*

monocytogenes strains giving information about host transmission, pathogenesis and niche adaptation (Call et al., 2003; Nightingale et al., 2005). Thus the construction of a *L. monocytogenes* whole-genome microarray (Hinds et al., 2002) is achievable due to the availability of the genome sequence of *L. monocytogenes* EGD-e (Glaser et al., 2001) enabling study the relationships among *Listeria* strains either species or serotype level and more than that trying to decipher its questions about adaptation and virulence profile.

Aiming to compare the genomic content of *L. monocytogenes* strains from distinct geographic regions and sources, we performed DNA microarray analyses allied to Bayesian algorithm. Additionally through bioinformatics studies we intended identify the possibly genes associated with the persistent feature of these strains.

3. MATERIALS AND METHODS

Strains Collection

The 100 strains (99 *L. monocytogenes* and one *L. innocua*) used in this study were obtained from diverse countries and sources, to date 47 strains from Brazil, 10 from Denmark, 08 from Austria, 6 from Ireland, 1 from Unites States and 26 from unknown origins (Table 1). These strains were isolated from food (58), equipment and environment of industries (16), clinical cases both humans and animals (14), asymptomatic animal carriers (2), environment (6) and unidentified sources (5).

Genomic DNA (gDNA) extraction

The genomic DNA extraction was carried out with Wizard (Promega) purification kit and its quality and quantification was performed by 1% agarose gel electrophoresis and GeneQuant spectrophotometer (Amersham Pharmacia), respectively. The gDNA samples were kept at 4°C to minimize damage which can be caused by freeze thawing.

Microarray Design

From 2944 annotated sequences of the whole-sequenced strain EGD-e (Glaser et al., 2001), 2855 were included in the design process. The *L. monocytogenes* arrays were developed and printed by BµG@S (Bacterial Microarray Group at St. Georges's Hospital Medical School), Department of Cellular and Molecular Medicine, St George's Hospital Medical School, London (Hinds et al., 2002). Briefly, ten pairs of gene-specific primers were designed to each sequence by applying PRIMER3 (Rozen & Skaletsky, 2000). Primers were designed to have a length of 20-25 base pairs (bp), melting temperature between 50℃ and 80℃ and amplicon ranging in size from 100 to 800 bp in length with an optimum size of 600 bp. The PCR primers were selected based on the BLAST similarity of the predicted PCR product to other genes on the microarray. Products with no similarity or least similarity were selected to ensure the least possible cross-hybridization on the microarray. The PCR products are spotted at a high density on poly-L-lysine-coated glass microscope slide, using a MicroGridII (BioRobotics) arraying robot. Array design is available from ArrayExpress database (www.ebi.ac.uk/arrayexpress accession number A-BUGS-19).

DNA labelling and hybridization

For each microarray slide, 5µg of both test genomic DNA labelled with cyanine dye Cy3 Fluorolink™ (Amersham Bioscience) and *L. monocytogenes* EGD-e DNA labelled with cyanine dye Cy5 Fluorolink™ (Amersham Bioscience) was used in the microarray hybridization. The *L. monocytogenes* EGD-e is a standard strain (control) used as general reference for all hybridizations as previously described by Dorrell et al. (2001). Microarray slides were prehybridized into a Coplin jar within a solution which consist in 3.5 X SSC (1X SSC - 0.15M NaCl plus 0.015M Sodium Citrate), 0.1% SDS (Sodium Dodecyl Sulfate) and 10mg/mL BSA (Bovine Serum Albumin) at 65ºC for 20 minutes. After this the slides were rinsed thoroughly in distillated water followed for rinsing in isopropanol 1 minute each, centrifuged and kept way from light until the moment of hybridization.

After DNA labelling, test and control DNA, Cy3-labelled DNA and Cy5-labelled DNA, respectively, were mixed together and then purified with MinElute Purification Kit (Qiagen). Just before the hybridization the solution was mixed with 20X SSC and 2% SDS followed for denaturation at 95ºC for 2 minutes, applied on microarray slide and covered with 22 x 22mm LifterSlips (Eyrie Scientific), hybridized overnight sealed in a humidified hybridization chamber (Telechem International) and immersed in a bath at 65ºC for 16-20 hours. Afterwards, the slides were washed first in a pre-heated at 65ºC Wash A solution (1x SSC, 0.05%SDS and distillated water) for 2 minutes and then twice in the Wash B solution (0.06X SSC and distillated water), centrifuged and scanned.

Data acquisition, processing and comparative phylogenomics

The scanning was carried out in an Affymatrix 418 Array Scanner (MWG Biotech) and signal data were obtained through ImaGene 5.2 (BioDiscovery). All generated data were further analysed for whole-genome comparisons of tests strains against the *L. monocytogenes* EGD-e strain by GeneSpring GX 7.3.1. (Agilent Technologies UK Limited).

The signal data transformation set measurements values below 0.01 were set to 0.01. The experiment was normalized to 50th percentile of all measurements in that sample which was calculated using all genes not marked absent. For each gene represented by a signal spot the light intensity was measured dividing by control channel value in each sample. If the control channel was below 0.01 then 0.01 was used instead. If the control channel and the signal channel were both below 0.01 then no data was reported.

The presence, divergence or absence of each gene of the array was determined by GACK analysis for cut-off assignment (GACK software) by transforming the normalized data from GeneSpring in log of ratio of raw and control data (Howard et al., 2006), since this software calculate an estimated probability of presence (EPP) value for each gene (Kim et al., 2002). The relationship of the strains was determined by Mr. Bayes v.3.0 software with a Bayesian method-based algorithm (Ronquist & Huelsenbeck, 2003), however before it the GACK output trinary format was transformed into binary data which

means divergent or marginal genes were converted to present. The Bayesian model employ four-chain Markov chain Monte Carlo and 16-category gamma distribution with 1 million iterations to model the presence or absence rate heterogeneity per gene throughout the genome of strain as previously described by Howard et al. (2006) and Stabler et al. (2006). Treeview was the software used to view the tree generated.

4. RESULTS AND DISCUSSION

A wide collection of 99 *L. monocytogenes* strains from diverse countries and sources were used in this study (Table 1) and also a *L. innocua* strain, totalling 100 strains. All strains were competitively hybridized with a *L. monocytogenes* DNA microarray based on the whole-genome sequence *L. monocytogenes* EGD-e (serotype 1/2a), which has 2855 target genes.

The Bayesian algorithm allowed the clustering of the *L. monocytogenes* strains into 2 central clades (Figure 1) which represent the two major lineages of this species, lineage I and II as previously described (Piffareti et al., 1989), proofing the huge power of this method for strain differentiation. Additionally our study provided concordant results as earlier reported in other phylogenetic studies of *L. monocytogenes* genome comparison (Call el al., 2003; Zhang et al., 2003; Doumith et al, 2004b; Nightingale et al., 2005). Briefly, we have found an outlined split among lineage II strains enhancing the idea of other authors (Nightingale et al., 2005) by which strains of this lineage behave such as sub-populations of *L. monocytogenes*. The highly degree of clonality correlated by lineage II-serotype strains may be indicative of a recent population bottleneck (Meinersmann et al., 2004). Paradoxally a model was not verified for lineage I strains which not show a clear serotype delineated of clustering. Hence *L. monocytogenes* lineage I represent a “generalist” lineage which is better adapted to survive and multiply in the environment while still maintaining the ability to cause human disease (Nightingale et al., 2004).

Apart from the important divergence between lineage I and II strains, on further observations we have noted a remarkable clustering of persistent strains into one of the lineage I clades (clade F). This is the only clade where the

clustering of strains are reasonable, since the other did not separated strains regarding, either serotype nor origin of strain, corroborating the theory of *L. monocytogenes* lineage I strains have a greater history of horizontal gene transfer (Nightingale et al., 2005) and so large variability among strains.

Surprisingly this is the only clade where none Brazilian strains were present. Nevertheless there was no evidence that geographical origin of strains was a determinant factor by which could contribute for this clustering as the persistent strains were isolated from distinct European countries as Denmark, Austria and Ireland. We could speculate that Brazilian strains would be young sample of *L. monocytogenes* strains less adapted to some niches in the environment. In addition none clinical strain was included in this cluster corroborating the idea of specific adaptation to environment of these strains or specific manner of gene regulation of environmental strains.

Besides the fact of the huge variety of strains from different geographic regions and sources used in this study, there was not strict correlation among their clustering and strains origins; apart from the serotyping profile. However, as previously mentioned, one of the most interesting finding in this work was the clustering of the vast majority of persistent strains in the same clade (some of them cited by Jensen et al., 2008). All strains present in this clade were isolated from food and food-processing plants (industry environment) highlighting the characteristic of adaptation and survival of these strains in the environment rather than their ability of cause disease.

The strains La22, No37-1, N53-1, La111, M103-1, No40-1 and 3R23-1 used in this study were isolated from Danish fish slaughter- and smokehouses and characterized as persistent by Jensen et al., (2008) mainly due its isolation up to 10 years in the food-processing plants (Vogel et al., 2001). Interesting the strains N53-1, La111 and M103-1 grouped in clade F (persistent strains) have the RAPD (Random Amplified Polymorphic DNA) type 9 which is a RAPD profile that has been isolated from Denmark, Faroe Island and in products from Greenland over years (Vogel et al., 2001; Jensen et al., 2008). Besides that RAPD 9 strains have the characteristic of being predominant into the processing plants assessed (Jensen et al., 2008). These observations are in line and corroborate with the idea that the clade F really clustered strains based in its persistent feature. The reasons for persistence of *L. monocytogenes*

strains could be: (I) specific RAPD type can be better colonizer (more able to adhere surfaces, for example); (II) this RAPD type is more resistant to disinfectants or (III) it is able to outcompete other RAPD types on the processing equipment and during enrichment on laboratory works. Also RAPD type 9 belongs to *L. monocytogenes* lineage I and so it is an indicative of survival advantage; besides that lineage I strains could be more-efficient in the uptake and is more stress tolerant in competition situations and/or during nutrient depletion and then better suited to survive (Jensen et al., 2008).

Persistent or adaptive genes identification

As reported before certainly our more striking observation in this study was the clustering of persistent strains in one same clade. With this approach we achieved in the first time the possibility of identify the probable genes involved in the adaptation and persistence of *L. monocytogenes* strains in the environment of industry. It can be a clue for the knowledge about this relevant feature of *L. monocytogenes* and so the starting point to develop new methods to avoid and maybe to eradicate it from food industries.

In order to address questions regarding to adaptation and persistence profile of strains, through Genespring analyses we generated a Venn diagram (Figure 2) of strains clustered into lineage I thief clade (that comprises three sub-clades, Figure 1) by which it was found 1510 common genes shared by the three clades belonging to lineage I strains.

Interestingly the persistent strains clade shared only 17 genes (Figure 2 and Table 2), all of them being exclusively present in this group of strains. The genes are mainly involved in stress resistance and are related to carbohydrate metabolism (lmo0105), carbohydrate transport and metabolism (lmo2696), environmental information processing (membrane transport) (lmo0137), signal transduction mechanisms (lmo1467 and lmo1642), cell surface protein (adhesion) (lmo0333), amino acid transport and metabolism (lmo1991 and lmo2192), nucleotide transport and metabolism (lmo2155), translation (lmo1843), cell wall (membrane biogenesis) (lmo1851), replication, recombination and repair (lmo1286 - parE), transport of small molecules similar to ABC transporter (lmo1969), metabolism of lipids (lmo1970) and unknown

function (lmo2442). The genes lmo0381 and lmo0461 are genes coding to hypothetical proteins with unknown function for *L. monocytogenes* strains, however they are paralogs genes with well-defined functions in other bacteria as *Bacillus* spp. (BLAST alignments).

The most part of genes reported in our study are first time reported in a research as *L. monocytogenes* strains persistence-indicative, the exception were the genes lmo1969 and lmo1970 that have been reported by Zhang et al. (2003) and by Doumith et al. (2004) as exemplars of specific lineage I strains markers.

The most remarkable genes grouped in this set of 17 genes are: lmo0105, lmo0381, lmo0461 and lmo1642. The first region, lmo0105 coding sequence for chitinase B (*chiB*) a glycosyl hydrolase which catalyzes the degradation of chitin, an insoluble linear -1,4-linked polymer of *N*-acetylglucosamine (Henrissat & Bairoch, 1993), this compound is well-known as involved in biofilm formation. The enzyme plays important physiological and ecological roles and is present in a wide range of organisms, including those that do not contain chitin (Henrissat & Bairoch, 1993). This enzyme has been reported in several bacterial researches (Barboza-Corona et al., 2003; Carranza et al., 2009; Hirano et al, 2009; Stewart et al., 2009) as a manner to keep advantage in the environment. *ChiB* consists of two domains being one of those a small carboxy-terminal putative chitin-binding domain which shows homology to fibronectin type III. Due to the bonding action of fibronectin its presence in these enzymes may be related to the substrate attachment. This protein was initially found in mammals and forms extracellular, multifunctional matrices which are important in cell bonding (Barboza-Corona et al., 2003).

L. monocytogenes has been isolated from shrimp (Destro et al., 1996; Gudmundsdóttir et al., 2006) which has chitin in its carapace. Besides a recent study verified that *L. monocytogenes* has a greater attachment to shrimp carapace than *Salmonella* (Wan et al., 2009) corroborating our findings that this enzyme someway can help in the attaching of the bacterium and also helping it in its environmental adaptation. It is worth of note which some of persistent strains used in this study came from fish processing plants and then could be one of the reasons of the presence of this gene in its genome content.

The lmo0381 gene codes a hypothetical protein with unknown function for *L. monocytogenes.* Nevertheless similarity analyses (BLAST) found correlation among this coding region with other related to paralogs genes of *Bacillus* spp. The same situation was noted with the gene lmo0461 by which was found the paralog gene polyketide-type polyunsaturated fatty acid synthase PfaA.

The paralog gene of lmo0381 is an M-protease which has an outstanding feature of conferring high alkaline resistance (Shirai et al., 1997). The enzyme belongs to the extremely alkaliphilic group related to alkaline serine proteases and shows optimal activity at pH 12.3. There is an extensive use of alkaline proteases in such industrial applications as laundry detergent additives (Markland & Smith, 1971). Certainly we can hypothesize that this paralog gene in *L. monocytogenes* persistent strains could give benefits for its survival in processing plants industries during the cleaning with detergents; based on the alkaline protease production and then adaptation to a non-physiological pH (Shirai et al., 1997).

The gene lmo0461 codes a hypothetical protein in *L. monocytogenes* with unknown function so alignments (BLAST) of this region, showed similarity with a polyketide-type polyunsaturated fatty acid synthase PfaA. The omega-3 polyunsaturated fatty acids (PUFAs) were once thought to be absent in bacterial membranes (Erwin & Bloch, 1964), but numerous bacterial species of marine origin have now been shown to produce very-long-chain PUFAs. The PUFA-producing strains have led to speculation that PUFA synthesis is an important adaptation for countering the effects of elevated hydrostatic pressure and low temperature. Based in this information is possible to outline a parallel with *L. monocytogenes* cell membrane by which confers to this bacterium singular resistance in the environments mainly in those cold found in food-processing plants. Thus it is possible to release a hypothesis that *L. monocytogenes* persistent strains possessing similar gene could be better adapted to food industries environment, since is very well-known that in this sort of ambient low temperatures and also high pressure are constantly found, depending on the processing of the food.

In Gram-positive bacteria, such as *Bacillus subtilis* and the human pathogens *L. monocytogenes* and *Staphylococcus aureus*, one such signaling

cascade leads to the activation of the general stress sigma factor and enhance transcription of its large regulon, providing a global response to the imposed stress (Price, 2002). The stressosome is the signaling hub that integrates multiple physical stress signals (Avila-Perez et al., 2006) and orchestrates a single signaling outcome: the activation of sigma B. The gene lmo1642 of *L. monocytogenes* code for a putative sigma factor regulator involved in signal transduction mechanisms. Probably lmo1642 acts similarly as the stressosome of *B. subtilis* when it experience changes in the environment such fluctuations in the temperature, pH or ionic strength, by which can trigger the expression of two distinct sets of genes. One set constitutes a response specific to the imposed stress; the second set corresponds to a more general response to stress. This sigma factor provides multiple stress resistances to the cell. Thus, the stressosome appears to have evolved to provide a common solution to the problem of signal integration of bacteria (reviewed by Marles-Wright et al., 2008). We could speculate that this mechanism can be a parallel path of regulation developed by *L. monocytogenes* persistent strains aiming to work out questions regarding to stress regulation.

The lmo 1467 gene is similar to phosphatase starvation induced protein PhoH and it is also involved in signaling transduction mechanism. It may reply to a variety of different stimuli such as heat shock, oxidative products nutrient limitation, acidic condition, and damage caused by toxic chemicals and physical agents (Hanawa et al., 2006).

L. monocytogenes EGD-e strains have a substantial number of ABC transporters and sugar phosphotransferase system (PTS) comprising 4% of the genome (Glaser et al., 2001). It has been reported an ABC transport complex in the region lmo1060-lmo1061 to be specific for lineage I strains (Doumith et al., 200b) and according to Zhang et al. (2003) the gene lmo1062 which also codes for an ABC permease is present only in 1/2a strains. However in our study we found another region (lmo0137) also with ABC transporter function but supposed to be specific for persistent strains. This finding may point to that different strains can get different paths for nutrient utilization and consequently to adaptation.

According to our analyses, *L. monocytogenes* persistent strains apparently shaped a sort of machinery developed specifically aiming to create

advantages under harsh conditions of physical and chemical stress faced in the environment. The character of persistence of strains is strictly linked to their enhanced adherence and biofilm formation ability (Borucki et al., 2003) which is also a way of survival. Despite of the progress which has been made to unmask *L. monocytogenes* biofilm formation yet none gene is really associated to it. Recently it was proposed that BapL (*Listeria monocytogenes* Biofilm-Associated Protein) could contribute to *L. monocytogenes* adherence to abiotic surfaces. It is a protein similar to a *Staphylococcus aureus* protein associated to its binding to abiotic surfaces. However the BapL (gene lmo0435) is not an essential requirement for all strains adherence; since even BapL-negative strains demonstrate the skill of adhere to abiotic surface (Jordan et al., 2008). In our work we could not find the gene in question even in the lineage II strains; meanwhile it was found other genes related to attachment such as referred above.

The persistent types appear to be unique to each processing unit however it could be hypothesized that specific traits are required for a specific *L. monocytogenes* strain to persist as like as special resistance to cleaning and disinfecting agents or a special attachment capability. With this study, we release the hypothesis that this set of genes can be the clue for understanding the persistence of *L. monocytogenes* strains. Taken together these data can help in the reduction of this pathogen in food products since cross-contamination in food processing plants is the main reason of food contamination. Likely there are some more genes involved in this complex process of adaptation however we identify just those uniquely present in the persistent clade, probably the ones by which surely characterize persistent strains.

Additionally there was evidence that strains found to persist in food production environments but from different countries were also very similar, possible showing a transcontinental spread of persister isolates or co-evolution of persistent genotypes. This study provides an understanding in the molecular phylogeny and evolution of *L. monocytogenes* strains giving information about transmission and niche adaptation.

5. REFERENCES

Ávila-Pérez, M., Hellingwerf, K. J., Kort, R. Blue light activates the sigma^{B}-dependent stress response of *Bacillus subtilis* via YtvA. 2006. **Journal of Bacteriology.** 188: 6411-6414.

Autio, T., Hielm, S., Miettinem, M., Sjoberg, A. M., Aarnisalo, K., Bjorkroth, J., Mattila-Sandholm, T., Korleala, H. Source of *Listeria monocytogenes* contamination in a cold-smoked rainbow trout processing plant detected by Pulsed-field gel electrophoresis typing. 1999. **Applied Environmental Microbiology.** 65:150-155.

Barboza-Corona, J. E., Nieto-Mazzocco, E., Velázquez-Robledo, R., Salcedo-Hernandez, R., Bautista, M., Jiménez, B., Ibarra, J. E. Cloning, sequencing, and expression of the chitinase gene *chiA74* from *Bacillus thuringiensis.* 2003. **Applied and Environment Microbiology.** 63:1023-1029.

Borucki, M. K., Peppin, J. D., White, D., Loge, F., Call, D. R. Variation in biofilm formation among strains of *Listeria monocytogenes*. 2003. **Applied and Environment Microbiology.** 69:7336-7342.

Brosch, R., Chen, J.; Luchansky, J. B. Pulsed-field fingerprinting of listeriae: identification of genomic divisions for *Listeria monocytogenes* and their correlation with serovar. 1994. **Applied and Environment Microbiology***.* 60**:** 2584-92.

Call, D. R., Borucki, M. K., Besser, T. E. Mixed-genome microarrays reveal multiple serotype and lineage-specific differences among strains of *Listeria monocytogenes.* 2003. **Journal of Clinical Microbiology.** 41:632-639.

Carranza, P., Hartmann, I., Lehner, A., Stephan, R., Gehrig, P., Grossmann, J., Barkow-Oesterreicher, S., Roschitzki, B., Eberl, L., Riedel, K. Proteomic profiling of *Cronobacter turicensis* 3032, a food-borne opportunistic pathogen. 2009. **Proteomics.** 9:3564 – 3579.

Destro, M. T., Leitão, M., Farber, J. M. Use of molecular typing methods to trace the dissemination of *Listeria monocytogenes* in a shrimp processing plant. 1996. **Applied and Environment Microbiology.** 62:1852-1853.

Dorrell, N., Mangan, J. A., Laing, K. G., Hinds, J., Linton, D., Al-Ghusein, H., Barrell, B. G., Parkhill, J., Stoker, N. G., Karlyshev, A. V., Butcher, P. D., Wren, B. W. Whole genome comparison of *Campylobacter jejuni* human isolates using a low-cost microarray reveals extensive genetic diversity. 2001. **Genome Research.** 11:1706–1715.

Dorrell, N., Hinchliffe, S. J., Wren, B. W. Comparative phylogenomics of the pathogenic bacteria by microarray analysis. 2005. **Current Opinion in Microbiology.** 8:620-626.

Doumith, M., Cazalet, C., Simoes, N., Frangeul, L., Jacquet, C., Kunst, F., Martin, P., Cossart, P., Glaser, P. and Buchrieser, C. New aspects regarding evolution and virulence of *Listeria monocytogenes* revealed by comparative genomics and DNA arrays. 2004. **Infection and Immunity.** 72:1072-1083.

Erwin, J. & Bloch, K. 1964. Biosynthesis of unsaturated fatty acids in microorganisms. **Science.** 143: 1006-1012.

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

Gudmundsdóttir, S., Roche, S.M., Kristinsson, K. G., Kristjánsson, M. Virulence of Listeria monocytogenes isolates from humans and smoked salmon, peeled shrimp, and their processing environments. 2006. **Journal of Food Protection.** 69:2157-2160.

Hain, T., Steinweg, C., Chakraborty, T. Comparative and functional genomics of Listeria spp. 2006. **Journal of Biotechnology.** 126(1):37-51.

Henrissat, B., A. Bairoch**.** New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. 1993. **Biochemistry Journal.** 293:781–788.

Hanawa, T., Kai, M., Kamiya, S., Yamamoto, T. Cloning, sequencing, and transcriptional analysis of the dnaK heat shock operon of *Listeria monocytogenes.* 2000. **Cell Stress & Chaperones.** 5:21-29.

Hinds, J., Laing, K. G., Mangan, J. A., Butcher, P. D. Microarrays for microbes: the BµG@S approach. 2002. **Comparative and Functional Genomics**. 3:333-337.

Hirano, T., Kadokura, K., Ikegami, T., Shigeta, Y., Kumaki, Y., Hakamata, W., Oku, T., Nishio, T. Heterodisaccharide 4-O-(N-acetyl-β-D-glucosaminyl)-D-glucosamine is a specific inducer of chitinolytic enzyme production in Vibrios harboring chitin oligosaccharide deacetylase genes. 2009. **Glycobiology**. 19:1046-1053.

Howard, S. L., Gaunt, M. W., Hinds, J. Witney, A. A., Stabler, R., Wren, B. W. Application of comparative phylogenomics to study the evolution of *Yersinia enterocolitica* and to identify genetic differences relating to pathogenicity. 2006. **Journal of Bacteriology**. 188:3645-3653.

Jensen, A., Thomsen, L. E., Jorgensen, R. L., Larsen, M. H., Roldgaard, B. B., Christensen, B. B., Vogel, B. F., Gram, L., Ingmer, H. Processing plant persistent strains of *Listeria monocytogenes* appear to have a lower virulence potential than clinical strains in selected virulence models. 2008. **International Journal of Food Microbiology.** 123:254-261.

Jordan, S. J., Perni, S., Glenn, S., Fernandes, I., Barbosa, M., Sol, M., Tenreiro, R. P., Chambel, L., Barata, B., Zilhão, I., Aldsworth, T. G., Adriao, A., Faleiro, M. L., Shama, G., Andrew, P. W. *Listeria monocytogenes* Biofilm-associated Protein (BapL) may contribute to surface attachment of *L. monocytogenes* but si absent from many field isolates. 2008. **Applied and Environment Microbiology.** 74:5451-5456.

Kim, C. C., Joyce, E. A., Chan, K., Falkow, S. Improved analytical methods for microarray-based genome-composition analysis. 2002. **Genome Biology.** 3:RESEARCH0065.

Markland,F. S.; Smith,E.L. 1971. In Boyer,P.D. (ed.), **The Enzymes**. Academic Press, New York, vol.3, pp. 561–608.

Marles-Wright, J., Grant, T., Delumeau, O., van Duinen, G., Firbank, S. J., Lewis, P. J., Murray, J. W., Newman, J. A., Quin, M. B., Race, P. R., Rohou, A., Tichelaar, W., van Heel, M., Lewis, R. J. Molecular Architecture of the "Stressosome," a Signal Integration and Transduction Hub. 2008. **Science.** 322:92-96.

Meinersmann, R. J., Phillips, R. W., Wiedmann, M., Berrang, M. Multilocus sequence typing of *Listeria monocytogenes* by use of hypervariable genes reveals clonal and recombination histories of three lineages. 2004. **Applied and Environment Microbiology.** 70:2193-2203.

Nightingale, K. K., Schukken, Y. H. Nightingale C. R., Fortes, E. D., Ho, A. J. Her, Z. Grohn, Y. T., McDonough, P. L. Wiedmann. M. 2004. Ecology and transmission of *Listeria monocytogenes* in ruminants and the farm environment. **Applied and Environmental Microbiology.** 70:4458–4467.

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

Ojeniyi, B., Christensen, J., Bisgaard. Comparative investigation of *Listeria monocytogenes* isolated from a turkey processing plant, turkey products and from human cases of listeriosis in Denmark. 2000. **Epidemiol. Infect.** 125:303-308.

Piffaretti, J. C., Kressebuch, H., Aeschbacher, M., Bille, J., Bannerman, J. M., Selander, R. K., Rocourt, J. Genetic characterization of clones of the bacterium *Listeria monocytogenes* causing epidemic disease. 1989. **Proceedings of the National Academy of Sciences**. 86:3818-3822.

Price, C. W. ***Bacillus subtilis*** **and its closest relatives** (Sonenshein, A. L., Hoch, J. A., and Losick, R., eds) 2002. American Society for Microbiology, Washington, D. C. pp. 369–384

Ragon, M., Wirth, T., Hollandt, F., Lavenir, R., Lecuit, M., Le Monnier, A., Brisse, S. A new perspective on *Listeria monocytogenes* evolution. 2008. **Plos Pathogens.** 4:1-14.

Rocourt, J. 1996. Risk factors for listeriosis. **Food Control.** 7:195-202

Rocourt, J. The genus *Listeria* and *Listeria monocytogenes*: phylogenetic position, taxonomy, and identification, p. 1–20. *In* E. T. Ryser and E. H. Marth (ed.), 1999. **Listeria, listeriosis, and food safety**. 2nd ed. Marcel Dekker Inc., New York, N.Y.

Ronquist, F.; Huelsenbeck, J. P. MrBayes3: Bayesian phylogenetic inference under mixed models. 2003. **Bioinformatics.** 19:1572-1574.

Rozen, S.; Skaletsky, H**.** 2000. **Methods in Molecular Biology.** 132:365-386.

Shirai, T., Suzuki, A., Yamane, T., Ashida, T., Kobayashi, T., Hitomi, J., Ito, S. 1997. High-resolution crystal structure of M-protease: phylogeny aided analysis of the high-alkaline adaptation mechanism. **Protein Engineering.** 10:627-634.

Stabler, R. A., Gerding, D. N., Songer, J. G., Drudy, D., Brazier, J.S., Trinh, H. T., Witney, A. A., Hinds, J., Wren, B. W. Comparative phylogenomics of *Clostridium difficile* reveals clade specificity and microevolution of hypervirulent strains. **Journal of Bacteriology**. 2006. 188:7297-7305.

Stewart, C. R., Rossier, O., Cianciotto, N. P. Surface translocation by *Legionella pneumophila*: a form of sliding motility that is dependent upon Type II protein secretion. **Journal of Bacteriology**. 191:1537–1546.

Tasara, T., Stephan, R. Evaluation of housekeeping genes in *L. monocytogenes* as potential internal control references for normalizing mRNA expression levels in stress adaptation models using real-time PCR. 2007. **FEMS Microbiological Letters.** 269:265-272.

Vasquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., Gonzalez-Zorn, B., Wehland, J. and Kreft, J. *Listeria* Pathogenesis and Molecular Virulence Determinants. 2001. **Clinical Microbiology Reviews**. p. 584–640.

Vogel, B. F., Jorgensen, L. V., Ojeniyi, B., Huss, H. H., Gram, L. Diversity of *Listeria monocytogenes* isolates from cold-smoked salmon produced in different smokehouses as assessed by Random Amplified Polymorphic DNA analyses. 2001. **International Journal of Food Microbiology.** 65:83-92.

Wan, M. N., Goulter, R. M., Poole, S. E., Deeth, H. C., Dykes, G. A. Relationship between the physicochemical properties of no-chitinolytic *Listeria* and *Salmonella* and their attachment to shrimp carapace. 2009. **Journal Food Protection**. 72:1181-1189.

Wulff, G., Gram, L., Ahrens, P., Vogel, B. F. One group of genetically similar *Listeria monocytogenes* strains frequently dominates and persists in several fish slaughter- and smokehouses. 2006. **Applied and Environmental Microbiology.** 72: 4313-4322.

Zhang, C., Zhang, M., Ju, J. Nietfeltdt, J., Wise, J., Terry, P. M., Olson, M., Kachmann, S. D., Wiedmann, M., Samadpour, M., Benson, A. K. Genome diversification in phylogenetic lineages I and II of *Listeria monocytogenes*: identification of segments unique to lineage II populations. 2003. **Journal of Bacteriology.** 185:5573-5584.

Table 1 – Identification and information about *Listeria* strains

Strain	Origin	Year of isolation	Group	Serotype/ Serogroup	Lineage	City -Country
19	Cheese	2007	Food	-	-	Cork - Ireland
30	Dry faeces – cow	-	Animal asymptomatic	4b	II-1	Cork - Ireland
92	Cheese	-	Food	-	-	Cork - Ireland
411	Unknown	-	Unknown	-	-	Cork - Ireland
535	Vacherin Cheese	1996	Food	-	-	Unknown
758	Producer B – smear	1997	Environment – Persistent	-	-	Vienna - Austria
1454	Clinical – animal (cow – mastitis)	1997	Animal	1/2a	I-1	Vienna - Austria
1458	Producer A – smear	1997	Environment – non-persistent	-	-	Vienna - Austria
1727	Outbreak	1998	Human	-	-	Atlanta, CDC - USA
2315	Producer B – Cheese	1999	Food- Persistent	-	-	Vienna - Austria
2540	Producer A - Cheese rasp	1999	Food – Persistent	-	-	Vienna - Austria
2566	Producer B – smear	1999	Environment - non-persistent	-	-	Vienna - Austria
4239	Human case (52 year-old male)	1998	Human	1/2a	I-1	Unknown
4359	Producer A - water	2004	Food – Persistent	-	-	Vienna - Austria
4423	Producer B – smear	2004	Environment - Persistent	-	-	Vienna - Austria
4446	Clinical – Human case (63 year-old female)	1998	Human	4b	II-1	Unknown
4542	Clinical – Human case (84 year-old male)	1998	Human	4b	II-1	Unknown
4810	Clinical – Human case (84 year-old male)	1998	Human	4	II	Unknown
4898	Producer A - water	2007	Food – Persistent	-	-	Vienna - Austria
4905	Human case (70 year-old female)	1998	Human	1/2	-	Unknown
6179	Cheese	2000	Food – Persistent	1/2a	I-1	Cork - Ireland
6895	Ham	1998	Food	1/2	II	Unknown
7291	Pasta with chicken	1998	Food	4b	II-1	Unknown
7418	Spreadable sausage	1998	Food – Sporadic	1/2b	II-2	Unknown
2V903b	Smokehouse V - cold-smoked salmon - final product - plant 1	1999	Food	4	I	Denmark
3R23-1	Smokehouse 1 – steel	2003	Industry Environment	-	II	Denmark
Ag28	Seawater fish farm 1 - stone with algae	2004	Environment - Sporadic	-	-	Unknown
Ap143	Slaughterhouse 2 - mud from bleeding area	2004	Industry Environment	-	-	Unknown
Br21	Fresh water fish farm 2 - water from detesting tank	2004	Environment - Persistent	1/2a	I-1	Unknown
Br22	Fresh water fish farm 2 - sediment from detesting tank	2004	Environment	-	-	Unknown

Table 1 – continuation

BRA 1	Rubber plucker fingers of defeathering machine	2005	Industry Environment	4b	II-1	Pelotas – Brazil
BRA 102	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas – Brazil

BRA 104	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas – Brazil
BRA 107	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas – Brazil
BRA 108	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas – Brazil
BRA 10s	Pork sausage	-	Food	1b	-	Pelotas – Brazil
BRA 111	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas – Brazil
BRA 11o	Sheep carcass	-	Food	4b	II-1	Pelotas – Brazil
BRA 12o	Sheep carcass	-	Food	1/2c	I-2	Pelotas – Brazil
BRA 13s	Pork sausage	-	Food	4b	II-1	Pelotas – Brazil
BRA 18o	Sheep carcass	-	Food	1/2b	II-2	Pelotas – Brazil
BRA 1s	Pork sausage	-	Food	4b	II-1	Pelotas – Brazil
BRA 20	Evisceration floor	2006	Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 21o	Sheep carcass	-	Food	4b	II-1	Pelotas – Brazil
BRA 22	Washing water prechiller	2006	Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 22o	Sheep carcass	-	Food	1/2c	I-2	Pelotas – Brazil
BRA 22q	Cottage cheese	-	Food – retail	1/2a	I-1	Pelotas – Brazil
BRA 23q	Cottage cheese	-	Food – retail	4b	II-1	Pelotas – Brazil
BRA 24q	Cottage cheese	-	Food – retail	1/2b	II-2	Pelotas – Brazil
BRA 25s	Handling (pork sausage)	-	-	1c	-	Pelotas – Brazil
BRA 27	Washing water chiller	2006	Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 2o	Sheep carcass	-	Food	4b	II-1	Pelotas – Brazil
BRA 32	Conveyor cut belt	2006	Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 42	Chicken Final product	2006	Food – Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 45	Washing water chicken from evisceration	2006	Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 46	Evisceration floor	2006	Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 53	Conveyor cut belt	2006	Industry Environment	1/2b	II-2	Pelotas – Brazil
BRA 58	Chilled chicken	2006	Food – retail	4e	II-1	Pelotas – Brazil
BRA 63	Chilled chicken	2006	Food –retail	4e	II-1	Pelotas – Brazil
BRA 6o	Sheep carcass	-	Food	3c	I-2	Pelotas – Brazil
BRA 71	Chilled chicken	2006	Food – Retail	4e	II-1	Pelotas – Brazil
BRA 74	Chilled chicken	2006	Food – Retail	4e	II-1	Pelotas – Brazil
BRA 76	Chicken Final product	2006	Food – Industry Environment	4e	II-1	Pelotas – Brazil
BRA 77	Chilled chicken	2006	Food – Retail	4e	II-1	Pelotas – Brazil
BRA 82	Chilled chicken	2006	Food – Retail	4e	II-1	Pelotas – Brazil
BRA 86	Chicken cloacae swab (faeces)	2006	Animal asymptomatic	4e	II-1	Pelotas – Brazil

Table 1- continuation

BRA 88	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas – Brazil
BRA 8q	Cottage cheese	-	Food – Retail	1/2c	I-2	Pelotas – Brazil
BRA 8s	Pork sausage	-	Food – Retail	1c	-	Pelotas – Brazil
BRA 91	Chilled chicken	2006	Food – Retail	1/2a	I-1	Pelotas – Brazil
BRA 94	Chilled chicken	2006	Food – Retail	1/2b	II-2	Pelotas – Brazil
BRA 98	Chilled chicken	2006	Food – Retail	1/2c	I-2	Pelotas – Brazil
BRA 99	Chilled chicken	2006	Food – Retail	-	-	Pelotas – Brazil
BRA 9o	Sheep carcass	-	Food	1/2a	I-1	Pelotas - Brazil
BRA St07	Clinical – human	-	Human	1/2b	II-2	Brazil
BRA St08	Clinical – human	-	Human	1/2a	I-1	Brazil
BRA St247	Clinical – human	-	Human	4b	II-1	Brazil
EDG	Animal case – rabbit	1926	Animal - Standard strain	1/2a	I-1	Unknown
H025	Fresh water fish farm 1 - surface of rainbow trout	2004	Environment	-	-	Unknown
L. innocua	-	-	-	6a	-	Unknown
La111	Smokehouse G - cold-smoked salmon	1996	Food – Industry environment	1/2a	I-1	Denmark
La22	Smokehouse V - cold-smoked salmon - final product	1995	Food - Industry environment	1/2a	I-1	Denmark
11994	Unknown	-	Unknown	-	-	Unknown
Lm 7973	Unknown	-	Unknown	-	-	Unknown
Lo28	Clinical - Pregnant woman	-	Human Faeces	1/2c	I-2	Unknown
M103-1	Slaughterhouse A - rubber conveyor belt	2003	Industry Environment	1/2a	I-1	Denmark
N22-2	Smokehouse 2 - floor in slicing area	2002	Sporadic - food industry environment	-	-	Unknown
N53-1	Smokehouse 2 - scraper mat	2002	Industry Environment	1/2a	I-1	Denmark
No37-1	Smokehouse 3 - waste container (pet food)	2001	Industry Environment	-	II	Denmark
No40-1	Smokehouse 3 - floor swap	2003	Industry Environment	1/2a	I-1	Denmark
R416a	Smokehouse 1 - slicing machine - processing plant 2	1996	Industry Environment	-	-	Denmark
R479a	Smokehouse 1 - cold-smoked salmon - final product - processing plant 2	1996	Food – Persistent	1/2a	I-1	Denmark
S2583	Lettuce processing equipment	1998	Industry Environment	1/2	II	Unknown
Scott A	Human	-	Human – Standard strain	4b	II-1	Unknown
SVS2719	Animal case – goat	1998	Animal	1/2	II	Unknown
SVS2827	Ensilage	1998	Food	1/2	II	Unknown
V518a	Smokehouse V - cold-smoked salmon- final product slicing mach. 2 - proc plant 1	1998	Food - Industry Environment	4b	II-1	Denmark
Vo14	Fresh water fish farm 3 - grass and plants around the race-way-system	2004	Environment	-	-	Unknown
Vo26	Fresh water fish farm 3 - sediment from incoming stream	2004	Environmentl	-	-	Unknown
29	Unknown	-	Unknown	-	-	Cork – Ireland

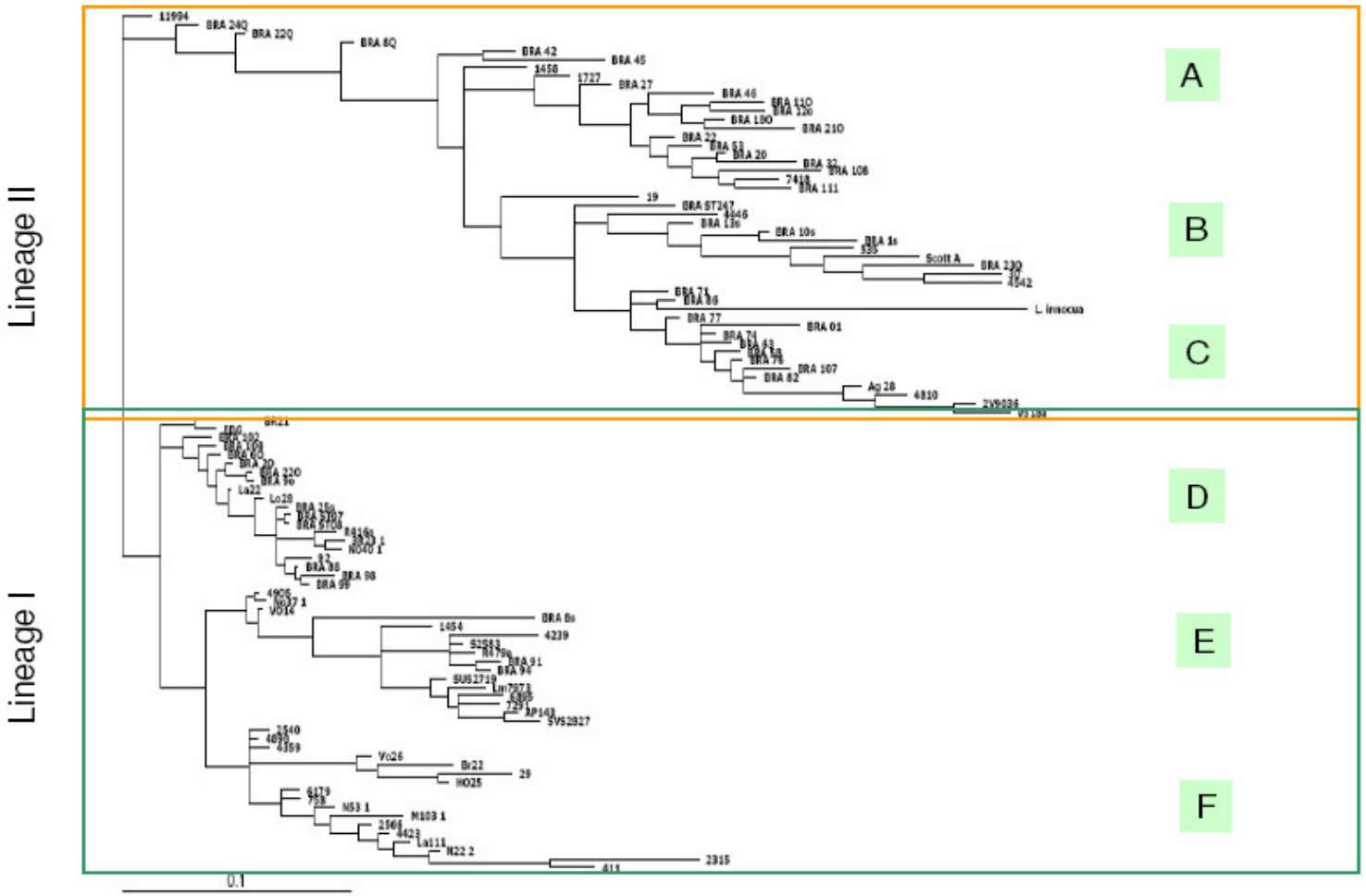

Figure 1 – Phylogenetic tree showing the genetic relationship among *L. monocytogenes* strains. A – clade of majority 1/2b strains; B – clade of majority 4b strains; C – clade of majority 4e strains; D, E and F – clade of majority 1/2a.

Table 2 – Unique genes present in *L. monocytogenes* persistent strains

Gene identification	***Unique genes present in clade F***	***Product of the gene***
LmEGDe-0105	righly similar to chitinase B	hypothetical protein
LmEGDe-0137	similar to oligopeptide ABC transporter, permease protein	hypothetical protein
LmEGDe-0333	similar to internalin proteins, putative peptideoglycan bound protein (LPXTG motif)	hypothetical protein
LmEGDe-0381	lmo 0381	hypothetical protein
LmEGDe-0461	lmo 0461	hypothetical protein
LmEGDe-1286	parE - highly similar to DNA gyrase-like protein (subunit B), decatenates newly replicated chromosomal DNA and relaxes positive and negative DNA supercoiling	DNA topoisomerase IV subunit B
LmEGDe-1467	similar to phosphatase starvation induced protein PhoH	hypothetical protein
LmEGDe-1642	similar to putative sigma factor regulator	hypothetical protein
LmEGDe-1843	lmo1843	hypothetical protein
LmEGDe-1851	similar to carboxy-terminal processing proteinase	hypothetical protein
LmEGDe-1969	similar to 2-keto-3-deoxygluconate-6-phosfate aldolase	hypothetical protein
LmEGDe-1970	similar to putative to phosphotriesterase related protein	hypothetical protein
LmEGDe-1991	ilvA - similar to threonine dehydratase; catalyses the formation 2-oxobutanoate from L-threonine; biosynthetic	threonine dehydratase
LmEGDe-2155	similar to ribonucleoside-diphosphate reductase, subunit alpha; catalyses the rate-limiting step in dNTP synthesis	ribonucleotide-diphosphate reductase subnit alpha
LmEGDe-2192	similar to oligopeptide ABC transporter (ATP-binding protein)	hypothetical protein
LmEGDe-2442	lmo 2442	hypothetical protein
LmEGDe-2696	similar to hypothetical dihydroxyacetone kinase	hypothetical protein

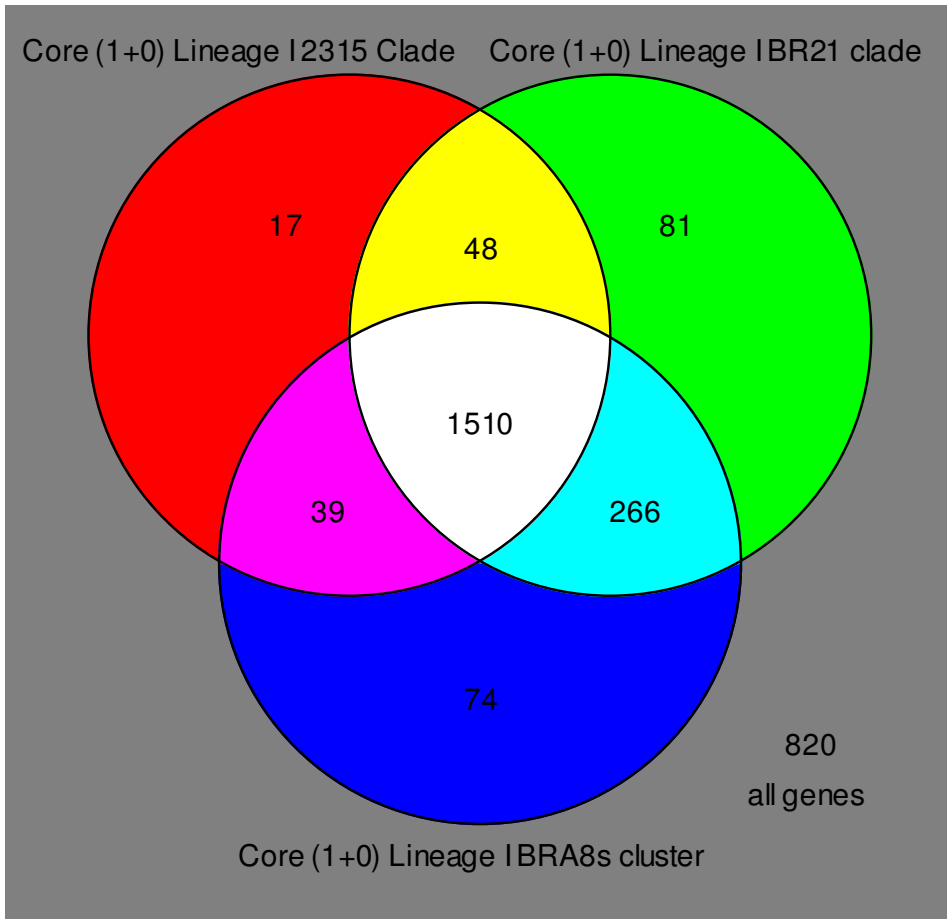

Figure 2 – GeneSpring Venn diagram of the *L. monocytogenes* lineage I clades. In red clade F showing its 17 unique genes; in green clade D showing its 81 unique genes; in yellow the 48 common genes between clade F and D; in blue clade E showing its 74 unique genes; in light blue the 266 common genes between clade D and E; in pink the 39 common genes between clade E and F and in white the 1510 common genes among clades D, E and F.

4 CONCLUSÕES

- Comparações filogenômicas entre cepas Européias, Brasileiras e Americana de *L. monocytogenes* confirmaram que o conteúdo genético das cepas, independentes da região geográfica, é muito similar;

- Não houve correlação entre a presença dos genes de virulência com a origem de isolamento, sorotipo e/ou região geográfica das cepas de *L. monocytogenes*;

- Não foi presenciada divergência significativa entre os genes de virulência nas cepas estudadas;

- A maioria dos genes de virulência de *L. monocytogenes* são conservados, tanto em cepas virulentas, quanto naquelas consideradas menos virulentas, tais como aquelas isoladas do ambiente;

- Há evidências de que, cepas persistentes em indústrias alimentícias, mesmo que pertencentes a diferentes países são muito similares. Possivelmente apresentando dispersão transcontinental deste tipo de microrganismo ou até mesmo que exista a evolução de genótipos com características de persistência;

- As cepas persistentes avaliadas, provavelmente, adaptaram um conjunto específico de genes, os quais podem fornecer determinadas vantagens sob as condições adversas encontradas por estes patógenos. Este conjunto de genes pode ser um indício para todo o entendimento sobre persistência das cepas de *L. monocytogenes* em plantas processadoras de alimentos;

- Estes dados quando utilizados em conjunto podem auxiliar na redução de *L. monocytogenes* nos alimentos.

5 REFERÊNCIAS

BENEDETTI, V M G., BIGLIA, N., SISMONDI, P., BORTOLI, M. DNA chips: the future of biomarkers. **International Journal of Biological Markers.** v.15, n. 1, 2000.

CABRITA, P., CORREIA, S., FERREIRA-DIAS, S., BRITO, L. Genetic characterization of *Listeria monocytogenes* food isolates and pathogenic potential within serovars 1/2a and 1/2b. **System. Applied Microbiol.** n. 27, p. 454-461, 2004.

CALL, D. R., BORUCKI, M. K., LOGE, F. J. Detection of bacterial pathogens in environmental samples using DNA microarrays. **Journal of Microbiological Methods**. v. 53, p. 235– 243, 2003.

DORRELL, N., HINCHLIFFE, S. J., WREN, B. W. Comparative phylogenomics of pathogenic bacteria by microarray analysis. **Current Opinion in Microbiology**. v. 8, p.620–626, 2005.

DOUMITH, M., JACQUET, C., GOULET, V., OGGONI, C., LOOCK, F. V., BUCHRIESER, C., MARTIN, P. Use of DNA arrays for the analysis of outbreak-related strains of *Listeria monocytogenes.* **International Journal of Medical Microbiology**. v. 296, p. 559–562, 2006.

DUSSURGET, O., PIZARRO-CERDA, J., COSSART, P. Molecular determinants of *Listeria monocytogenes* virulence. **Annual Review of Microbiology.** v. 58, p. 587-610, 2004.

FARBER, J. M. & PETERKIN, P. I. 1991. Listeria monocytogenes, a food-borne pathogen. *Microbiol Rev,* 55**,** 476-511.

GARAIZAR, J., REMENTERIA, A., PORWOLLIK, S. DNAmicroarray technology: a newtool for the epidemiological typing of bacterial pathogens? **FEMS Immunol Med Microbiol.** v.47, p. 178–189, 2006.

GLASER, P., FRANGEUL, L., BUCHRIESER, C., RUSNIOK, C., AMEND, A., BAQUERO, F., BERCHE, P., BLOECKER, H., BRANDT, P., CHAKRABORTY, T., CHARBIT, A., CHETOUANI, F., COUVÉ, E., DE DARUVAR, A., DEHOUX, P., DOMANN, E., DOMÍNGUEZ-BERNAL, G., DUCHAUD, E., DURANT, L., DUSSURGET, O., ENTIAN, K.-D., FSIHI, H., GARCIA-DEL PORTILLO, F., GARRIDO, P., GAUTIER, L., GOEBEL, W., GÓMEZ-LÓPEZ, N., HAIN, T., HAUF, J., JACKSON, D., JONES, L.-M., KAERST, U., KREFT, J., KUHN, M., KUNST, F., KURAPKAT, G., MADUEÑO, E., MAITOURNAM, A., MATA VICENTE, J., NG, E., NEDJARI, H., NORDSIEK, G., NOVELLA, S., DE PABLOS, B., PÉREZ-DIAZ, J.-C., PURCELL, R., REMMEL, B., ROSE, M., SCHLUETER, T., SIMOES, N., TIERREZ, A., VÁZQUEZ-BOLAND, J.-A., VOSS, H., WEHLAND, J., COSSART, P. Comparative Genomics of *Listeria* Species**. Science.** v. 294, n. 849, 2001.

GOMBAS, D. E., CHEN, Y., CLAVERO, R. S., SCOTT, V. N. Survey of *Listeria monocytogenes* in ready-to-eat foods. **Journal of Food Protection**, v. 66, n. 4, p. 559 – 569, 2003.

GUDBJORNSDOTTIR, S., GUDBJORNSDOTTIR, B., LAUZON, H. L., EINARSSON, H., KRISTINSSON, K. G., KRISTJANSSON, M. Tracing *Listeria monocytogenes* isolates from cold-smoked salmon and its processing environment in Iceland using pulsed-field gel electrophoresis. **International Journal of Food Microbiology.** v. 101, p. 41-45, 2005.

JACQUET, C.; GOUIN, E.; JEANNEL, D.; COSSART, P.; ROCOURT, J. Expression of ActA, Ami, InlB, and Listeriolysin O in *Listeria monocytogenes* of Human and Food Origin. **Applied and Environmental Microbiology**, v. 68, n. 2, p. 616 – 622, 2002.

LIANOU, A., STOPFORTH, J. D., YOON, Y., WIEDMANN, M., SOFOS, J. Growth and stress resistance variation in culture broth among *Listeria monocytogenes* strains of various serotypes and origins. **Journal of Food Protection.** v. 69, n. 11, p. 2640-2647, 2006.

NIGHTINGALE, K. K., WINDHAM, K., WIEDMANN, M. Evolution and molecular phylogeny of *Listeria monocytogenes* isolated from human and animal listeriosis cases and food. **Journal of Bacteriology**. vol.187, n.16, p 5537-5551, 2005.

ROCOURT, J. & COSSART, P. 1997. Listeria monocytogenes. IN DOYLE, M. P., BEUCAHT, L. R. & MONTVILLE, T. J. (Eds.) **Food Microbiology undamental and Frontiers.** 337-352 ed. Washington, ASM Press.

RUDI, K., KATLA, T., NATERSTAD, K. Multi locus fingerprinting of *Listeria monocytogenes* by sequence-specific labeling of DNA probes combined with array hybridization. **FEMS Microbiology Letters**. v. 220, p. 9-14, 2003.

SEELIGER, H. P. R. 1961. *Listeriosis*, Basel:Karger.

VASQUEZ-BOLAND, J.A., KUHN, M., BERCHE, P., CHAKRABORTY, T., DORNINIQUEZ-BERNAL, G., GOEBEL, W., GONZALEZ-ZORN, B., EHLAND, J., AND KREFT, J. *Listeria* pathogenesis and molecular virulence determinants. **Clinical Microbiology**. Rev. v.14, p. 584-640, 2001.

WALLS, I., BUCHANAN, R. L. Use of food objectives as a tool for reducting foodborne listeriosis. **Food Control.** (article in press), 2005.