Whole genome sequence analysis; an improved technology that identifies underlying genotypic differences between closely related *Listeria monocytogenes* strains

Edward M. Fox[a], Aidan Casey[b,c], Kieran Jordan[b], Aidan Coffey[c], Cormac G.M. Gahan[d], Olivia McAuliffe[b,*]

[a] *CSIRO Agriculture and Food, Victoria, Australia*
[b] *Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland*
[c] *Department of Biological Sciences, Cork Institute of Technology, Bishopstown, Cork, Ireland*
[d] *School of Microbiology, University College Cork, Cork, Ireland*

ARTICLE INFO

ABSTRACT

As the new technology of whole genome sequencing (WGS) has been shown to have greater discriminatory power in differentiating strains than the much-used pulsed-field gel electrophoresis (PFGE), there is currently a transition from using PFGE to WGS for disease outbreak investigation. Therefore, there is a need for comparison of bacterial isolates using both PFGE and WGS. In this study, two pairs of *L. monocytogenes* strains with geographically diverse sources of isolation but which had indistinguishable or closely related PFGE profiles, were subjected to WGS analysis. Comparative analysis of their genomes showed that one pair of strains which had closely related PFGE profiles in fact differed significantly from one another in terms of their antibiotic and heavy metal stress resistance determinants, and mobile genetic elements. Therefore, this research demonstrated the ability of WGS analysis to differentiate very closely related strains and that WGS analysis represents the most effective tool available for subtyping *L. monocytogenes* isolates.

1. Introduction

Listeria monocytogenes is a Gram-positive foodborne pathogen responsible for the bacterial infection listeriosis in humans and animals. While the disease itself is relatively rare, it is associated with a mortality rate of approximately 30% (Vázquez-Boland et al., 2001), mainly among immunocompromised individuals, and is predominantly acquired through the consumption of contaminated food produce (Swaminathan & Gerner-Smidt, 2007). The ability of *L. monocytogenes* to proliferate in a range of environmental extremes such as refrigeration temperatures (Chan & Wiedmann, 2008), low pH (Metselaar, den Besten, Abee, Moezelaar, & Zwietering, 2013) and in the presence of disinfectants (Casey et al., 2014; Fox, Leonard, & Jordan, 2011) allow this pathogen to survive many common industrial food preservation processes. As a consequence, the presence of *L. monocytogenes* in the processing environment has become an issue of concern to the food industry. This issue has been highlighted by a number of high profile outbreaks of listeriosis that have occurred in recent times, involving different types of ready-to-eat (RTE) food products, including pre-

prepared cantaloupe melon (Laksanalamai et al., 2012), quargel cheese (Rychli et al., 2014), and cold cut meats (Greenberg & Elliott, 2009). Outbreaks of listeriosis can be difficult to detect, given the potentially large geographical spread of disease cases in addition to the prolonged period of time between consumption of contaminated food and physical onset of a systemic listeriosis infection (Fugett, Schoonmaker-Bopp, Dumas, Corby, & Wiedmann, 2007). The incubation period for listeriosis disease manifestation in humans often exceeds 30 days (Swaminathan & Gerner-Smidt, 2007), and in certain circumstances has been determined to be as high as 67 days (Goulet, King, Vaillant, & de Valk, 2013). Therefore, the use of molecular subtyping techniques in establishing links between otherwise sporadic cases of listeriosis can be critical for the rapid detection and curtailment of a disease outbreak.

Pulsed-field gel electrophoresis (PFGE) is a widely used technique for the separation of large molecules of DNA (Herschleb, Ananiev, & Schwartz, 2007), and, prior to the emergence of WGS analysis, was generally regarded as the most discriminatory method for subtyping *L. monocytogenes* isolates (Graves & Swaminathan, 2001). This technique involves initial cleavage of the bacterial DNA of interest

with specific restriction enzymes, followed by separation of DNA fragments in a horizontal agarose gel using an electrical current with a periodically changing direction of flow (Dalmasso & Jordan, 2014). With regard to *L. monocytogenes*, the process of PFGE subtyping of isolates is carried out according to the PulseNet standardized protocol (Graves & Swaminathan, 2001). Following this protocol, the bacterial DNA of interest is digested separately by the *Apa*I and *Asc*I restriction enzymes, which cleave the DNA at specific restriction sites to generate a fragment pattern for each strain. Previous experimentation has indicated that this enzyme combination is particularly effective at discriminating *L. monocytogenes* isolates (Brosch et al., 1996), and that by using two restriction enzymes, the level of *L. monocytogenes* discrimination is higher than by digestion with either enzyme alone (Fox et al., 2012; Fugett et al., 2007). The pattern of DNA fragments that is generated using PFGE represents the "pulsotype" of a particular strain, and the pulsotype can be used as a basis for determining the relatedness of a collection of isolates by following previously defined guidelines (Tenover et al., 1995). The comparison of two PFGE patterns on this basis will classify their relationship into one of four categories: "indistinguishable" (patterns have same number of bands and same apparent band sizes), "closely related" (maximum of 2–3 fragment differences between patterns), "possibly related" (4–6 fragment differences) and "different" (> 6 fragment differences). Indistinguishable isolates are presumed to represent the same strain of the bacterium (Tenover et al., 1995), while closely related isolates may represent strains that only differ from one another on the basis of a single genetic event, such as an insertion, deletion, substitution or rearrangement of DNA, leading to the loss or gain of an enzyme restriction site (Goering, 2010).

The advent of PFGE has contributed greatly to *L. monocytogenes* research in recent times, particularly with regard to its role in the discovery of the "persistence" phenomenon, which is defined by the regular re-isolation of a given strain (on the basis of indistinguishable PFGE pulsotypes) from the same environment over the course of several months or years (Lundén, Autio, Markkula, Hellström, & Korkeala, 2003). PFGE has also proven to be a highly effective tool in the detection and sourcing of a number of listeriosis outbreaks (Gottlieb et al., 2006; Graves et al., 2005), contributing to disease curtailment through the rapid identification and recall of contaminated products (Olsen et al., 2005). While PFGE is widely regarded as the "gold standard" for the assessment of relationships between *L. monocytogenes* isolates (Dalmasso & Jordan, 2014), this technique does have certain limitations. Foremost of these limitations is that two isolates with indistinguishable pulsotypes may not necessarily be identical to one another, as only a small proportion of each genome is subjected to analysis (Tenover, Arbeit, & Goering, 1997). As such, PFGE is limited in that it cannot provide an insight into some of the phenotypic characteristics of a pair of closely related strains, such as their respective virulence potential, stress tolerance or antibiotic resistance ability. As a result, alternative methods for subtyping isolates of *L. monocytogenes* are continuously sought. The consistent reduction in sequencing costs in recent times has led to the emergence of whole genome sequence (WGS) analysis as a viable alternative technique for such purposes. WGS analyses have been previously used to identify several *L. monocytogenes* virulence factors on the basis of their relative absence in the non-pathogenic *L. innocua* (Buchrieser, Rusniok, Kunst, Cossart, & Glaser, 2003; Glaser et al., 2001; Stachowiak et al., 2015), as well as being invaluable in elucidating evolutionary relationships between closely related *Listeria* isolates (Hain et al., 2012). As a result, the Centres for Disease Control and Prevention (CDC) in the United States now favour WGS as their primary subtyping technique for listeriosis outbreaks (CDC, 2015a, b, c), on the basis that in comparison to other techniques such as PFGE, WGS has the capability to provide a clearer distinction between isolates that belong to a given outbreak and those which do not.

The aims of this study were to evaluate the discriminatory power of WGS analysis as a tool for subtyping *L. monocytogenes* isolates, and to compare it to PFGE. To achieve this, two pairs of *L. monocytogenes* strains with highly similar PFGE profiles were subjected to comparison at the genomic level, in order to identify any potential underlying genetic differences between them.

2. Materials and methods

2.1. Pulsed-field gel electrophoresis

PFGE analysis was carried out using International Standard PulseNet protocol (Pulsenet, 2013). DNA was restricted using both *Asc*I (50 U enzyme at 37 °C for 2 h) (Fermentas, UK), and 50 U *Apa*I (30 °C for 18 h) (Roche, UK) restriction endonucleases. DNA fragments were separated on a 1% SeaKem Gold agarose gel for 21 h as described in the PulseNet protocol, using the CHEF Mapper system (Bio-Rad, UK). Gels were visualised after staining in GelRed DNA stain (Biotium, US) for 30 min, followed by destaining in molecular grade water for 30 min. Cluster analysis of restriction profiles was performed using BioNumerics v7.5 software (Applied Maths, Belgium).

2.2. Genome sequencing

DNA was prepared from each of the four *L. monocytogenes* strains using the DNeasy Blood and Tissue Kit (Qiagen, The Netherlands) as per manufacturer's instructions. Following library preparation using the Illumina Nextera XT kit (Illumina, San Diego, CA), 300 bp paired-end sequencing was performed using the Illumina MiSeq platform. Raw reads were pre-processed to remove adapter sequences and low quality reads using the Trimmomatic (v0.22) software (Bolger, Lohse, & Usadel, 2014). *De novo* assembly was performed using the SPAdes (Species Prediction and Diversity Estimation) genome assembler tool v2.5.1 for strains 944 and 2993, or SPAdes v1.0 for strains 198 and 2932. SPAdes is based on an algorithm which employs multisized De Bruijn graphs with K-mer values of '21, 33, 55, 77' to construct the contiguous sequences (Nurk et al., 2013). Open reading frames (ORFs) for all of the genomes were predicted using RAST (Aziz et al., 2008). Annotations were verified and curated using BLASTp (Altschul et al., 1997) and Artemis (Rutherford et al., 2000).

2.3. Whole genome comparative analysis

Comparative genomic analysis of each pair of input strains was undertaken using Mauve (Darling, Mau, Blattner, & Perna, 2004), Artemis (Rutherford et al., 2000) and BRIG (BLAST Ring Image Generator) (Alikhan, Petty, Zakour, & Beatson, 2011). For each pair of genomes assessed, genes that were initially identified to belong to just one genome of the pair were subsequently confirmed to be strain-specific through BLASTp comparisons against the other genome of the pair using RAST (Altschul et al., 1997; Aziz et al., 2008).

2.4. Multilocus Sequence Typing (MLST)

The sequence type of each input strain in this study was determined by following a previously described methodology (Ragon et al., 2008). Gene fragments from seven housekeeping genes, namely *abcZ* (ABC transporter), *bglA* (beta-glucosidase), *cat* (catalase), *dapE* (succinyl diaminopimelate desuccinylase), *dat* (ᴅ-alanine aminotransferase), *ldh* (lactate dehydrogenase), and *lhkA* (histidine kinase), were subjected to whole sequence identity searches against the *L. monocytogenes* MLST database (http://www.pasteur.fr/cgi-bin/genopole/PF8/mlstdbnet.pl?page=oneseq&file=Lmono_profiles.xml) to determine their respective allele numbers. The corresponding allele numbers identified from the database searches were subsequently combined to generate a specific sequence type and clonal complex for each of the four *L. monocytogenes* input strains.
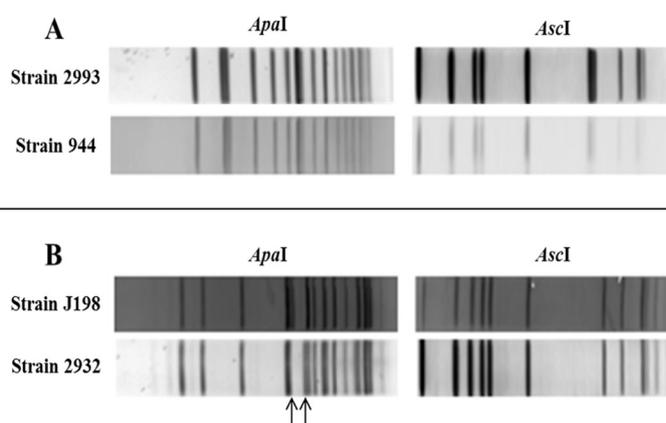
Fig. 1. PFGE restriction profile comparisons of each pair of *L. monocytogenes* isolates, with *Apa*I and *Asc*I restriction profiles of strains 2993 and 944 compared in A, and *Apa*I and *Asc*I restriction profiles of strains 198 and 2932 compared B. Arrows denote locations where restriction profiles differed from one another.

2.5. Detection of prophages, and subsequent linear comparisons

Each of the input genomes were scanned for the presence of prophage DNA using the PHAST software tool (Zhou, Liang, Lynch, Dennis, & Wishart, 2011). Genomic regions of interest for linear comparison were initially prepared in Artemis (Rutherford et al., 2000), then subsequently visualised using the EasyFig software (Sullivan, Petty, & Beatson, 2011).

3. Results and discussion

3.1. Comparative analysis – L. monocytogenes strains 944 and 2993

Despite their geographically distant isolation sources, the serotype 4b *L. monocytogenes* strains 944 (Rep. of Ireland) and 2993 (Australia) were identified *via* PFGE to be indistinguishable from one another when digested with both *Apa*I and *Asc*I restriction enzymes (Fig. 1). *In silico* MLST analysis determined that both isolates belonged to the same sequence type (ST) and clonal complex (CC), ST1 and CC1 (Table 1). A comparison of their respective draft genomes demonstrated that these two strains differ in length by just 275 base pairs (bp) (Table 2). Both genomes had a shared GC content of 37.9% and a similar number of open reading frames (ORFs), with strain 2993 containing a total of 2876 ORFs; one ORF less than the 2877 annotated in strain 944. Comparative genomic analysis determined that only 19 of the 2877 ORFs in strain 2993 were unique to that strain when compared to strain 944 (Table S1), while the corresponding number of ORFs unique to strain 944 totalled 17 (Table S2). The genes unique to each strain were predominantly contained within a single approximately 40 kbp region in each of these genomes, and were subsequently identified by the PHAST software to represent intact prophages (Fig. 2, Table 3). All of the 17 genes unique to strain 944 were contained in this region, and consisted of phage structural proteins or phage-related hypothetical proteins, while this was also the case for 18 of the 19 genes unique to strain 2993 (Fig. S1). Strain 2993 contained one additional novel gene, which BLASTp analysis identified to encode a transposase (*AN919_1197*). The role this transposase plays is not yet fully understood, and the consequence of its absence in strain 944 is not known.

Table 1
Sequence type and clonal complex of each of the *L. monocytogenes* input strains.

Strain	Multi Locus Sequence Type							Allelic profile	Clonal complex
	abcZ	bglA	cat	dapE	dat	ldh	lhkA		
944	3	1	1	1	3	1	3	1	1
2993	3	1	1	1	3	1	3	1	1
2932	6	5	6	4	1	4	1	9	9
198	6	5	6	4	1	4	1	9	9

Table 2
General features of the *L. monocytogenes* input strains.

L. monocytogenes strain	198	2932	944	2993
Country of origin	Ireland	Australia	Ireland	Australia
Serotype	1/2c	1/2c	4b	4b
Chromosome length (bp)	2,992,472[a]	3,005,504[a]	2,897,681	2,897,956
No. of chromosome contigs	19	28	10	13
G + C content (%)	37.9	37.9	37.9	37.9
No. of coding sequences (CDS)	2975	2993	2877	2876
Sequence type	9	9	1	1
No. of plasmids	1	1	–	–
Plasmid size (bp)	58,524	51,980[a]	–	–
Plasmid G + C content (%)	36.6	36.3	–	–
No. of plasmid CDS	69	61	–	–

[a] Genomes are not closed.

The results from the comparative genomic analysis of strains 944 and 2993 demonstrate the effectiveness of PFGE as a tool for accurately subtyping *L. monocytogenes* isolates on the basis of an indistinguishable pulsotype, as these two genomes shared 99% overall nucleotide sequence identity. Furthermore, this analysis also demonstrated the precision of WGS analysis, as clear differences between the two strains were successfully identified despite their extremely close relationship. However, strains of *L. monocytogenes* which have closely related PFGE pulsotypes may not always be as similar as the subtyping technique would suggest, exemplified by the comparative analysis of the other two *L. monocytogenes* strains in this study, namely strains 198 and 2932.

3.2. Comparative analysis – L. monocytogenes strains 198 and 2932

In a similar fashion to strains 944 and 2993, the serotype 1/2c *L. monocytogenes* strains 198 (Rep. of Ireland) and 2932 (Australia) also had geographically distant sources of isolation coupled with closely related PFGE profiles. Digestion of these genomes with the *Asc*I restriction enzyme yielded indistinguishable pulsotypes, while digestion with the *Apa*I enzyme produced pulsotypes that differed from one another in terms of a single DNA fragment (Fig. 1). When subjected to *in silico* MLST analysis, both strains were identified to belong to the same ST and CC (ST9 and CC9) (Table 1). The draft genome of strain 2932 was approximately 13,032 bp larger than that of strain 198, containing 2993 ORFs; a total of 18 more than the 2975 ORFs identified in strain 198. In addition, each strain harboured its own distinct plasmid. A total of 94 genes were identified in the genome and plasmid of strain 198 that were absent from strain 2932 (Table S2), while conversely a total of 91 genes were determined to be unique to strain 2932 (Table S4).

A large proportion of the novel genes in strain 198 were clustered into two specific genomic regions, which were subsequently identified to represent a prophage region (Fig. S1) and a previously described
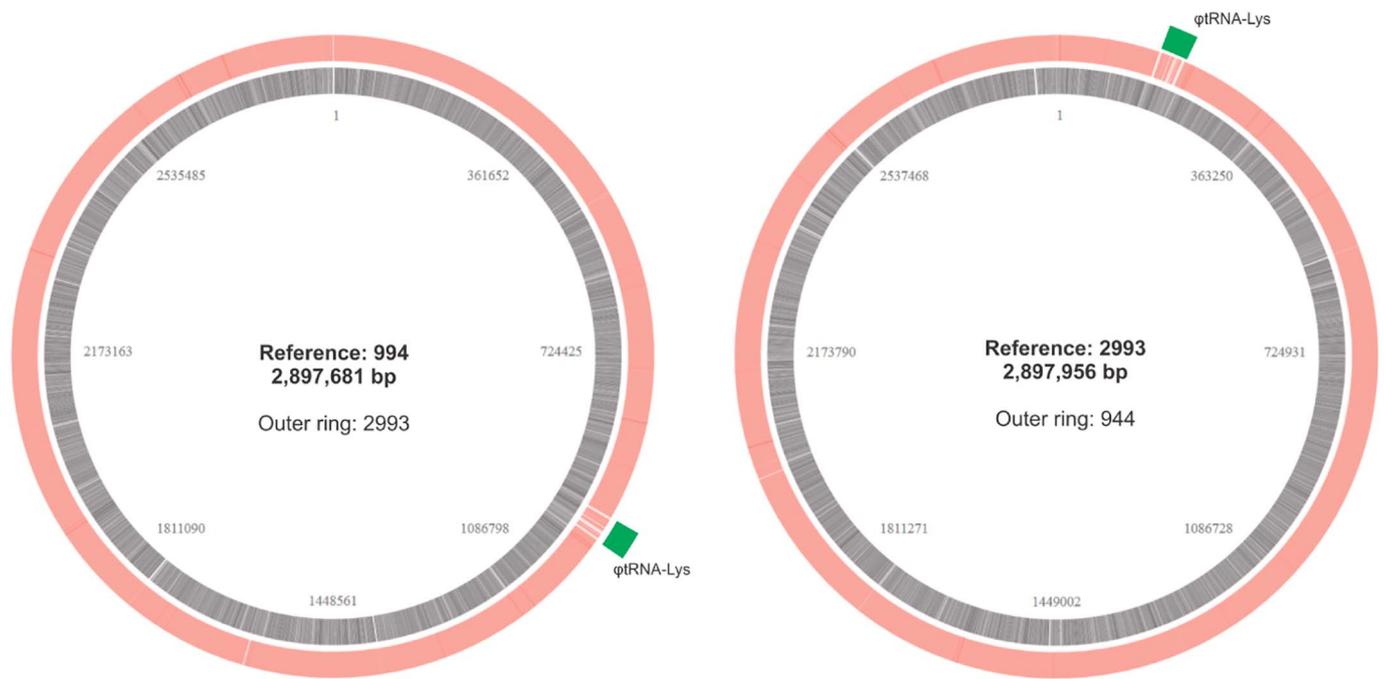
Fig. 2. Whole genome comparisons between strains 944 and 2993 visualised using a local implementation of BRIG with custom scripts (https://github.com/esteinig/brigD3https://github.com/esteinig/brigD3) (B).

genomic island (Fig. 3, Table 3; Gilmour et al., 2010). As a result, the majority of genes unique to strain 198 had phage-related or unknown functions, however they also included genes related to stress resistance. Some of these novel genes had functionally important roles in *L. monocytogenes*. For example, strain 198 contained genes that were absent from the 2932 genome, two of which (*AOA13_1620c* and *AOA13_1622c*) encoded the CpaF and CpaB proteins that are thought to have a role in the assembly of type IV pilus-like surface appendages (Gilmour et al., 2010), while the other (*AOA13_1541*) encoded the cell shape-determining protein MreB (Donaldson, Nanduri, Burgess, & Lawrence, 2009). The absence of these genes from the 2932 genome suggests a reduced capacity for this strain in terms of maintenance of cell morphology, in addition to a potential attenuation of cellular motility and biofilm formation capabilities, considering the role for pili in bacterial "twitching" (Proft & Baker, 2009). Together, these results demonstrate potentially important physical differences between these two strains that had not been previously identified. Subsequently, strains 198 and 2932 were subjected to a closer evaluation of two of their clinically relevant characteristics, namely their relative resistances to antibiotic and heavy metal stresses, and their respective pathogenic potentials.

3.3. Resistance to antibiotics and heavy metal tolerances of strains 198 and 2932

An increased resistance of *L. monocytogenes* to antibiotic treatments has become a common observation in recent times (Khen, Lynch, Carroll, McDowell, & Duffy, 2015; Li, Sherwood, & Logue, 2007), with numerous proposed mechanisms for mediation of that resistance, including the use of multidrug efflux pumps, as well as plasmid conjugation of specific resistance genes (Lungu et al., 2011). When *L. monocytogenes* strains 198 and 2932 were examined, both genomes were found to contain a number of non-specific multidrug efflux genes, as well as genes with putative functions in increasing resistance of the bacterium to lincomycin, quinolone, β-lactams and tetracycline (Table 4). *L. monocytogenes* strain 198 harboured *AOA13_1592c*, a non-specific multidrug resistance protein. In terms of heavy metal tolerance, both of these strains contain a number of non-specific heavy metal transport genes, as well as numerous lead/cadmium/zinc resistance genes (Table 4). Both strains 198 and 2932 contain a cadmium resistance transposon of approximately 6 kb in length harboured on their plasmids. The plasmid of strain 198 also harbours additional copper resistance genes which are absent from strain 2932 (*AOA13_p38c* and *AOA13_p44c*). In addition to this, strain 198 contains the previously

Table 3
Prophages identified in each of the *L. monocytogenes* input genomes.

Strain	Prophage	Location	Status	Size (kb)	Number of CDS	%GC
944	Monocin	Contig 3 (97,097–110,211)	Incomplete	13.1	18	38.2
	φtRNA-Lys	Contig 3 (149,707–189,301)	Intact	40.0	63	37.0
2993	Monocin	Contig 2 (97,487–110,601)	Incomplete	13.1	18	38.2
	φtRNA-Lys	Contig 2 (150,097–190,508)	Intact	40.4	63	36.6
198	Monocin	Contig 4 (38,337–51,450)	Incomplete	13.1	18	38.2
	φcomK	Contig 8 (1–33,143) & contig 2 (1–1698)	Incomplete	34.8[a]	51	36.6
	φtRNA-Arg	Contig 5 (877–38,552) & contig 3 (476,356–477,704)	Intact	39.9[a]	54	35.2
2932	Monocin	Contig 1 (120,591–133,704)	Incomplete	13.1	18	38.2
	φcomK	Contig 8 (1–12,789) & contig 7 (506,085–509,246)	Incomplete	15.7[a]	29	33.7
	φtRNA-Arg	Contig 4 (109,408–115,404) & contig 16 (1–32,821)	Intact	38.8[a]	11	32.9
	φRNA-MT	Contig 10 (13,852–68,405)	Intact	54.6	55	41.1

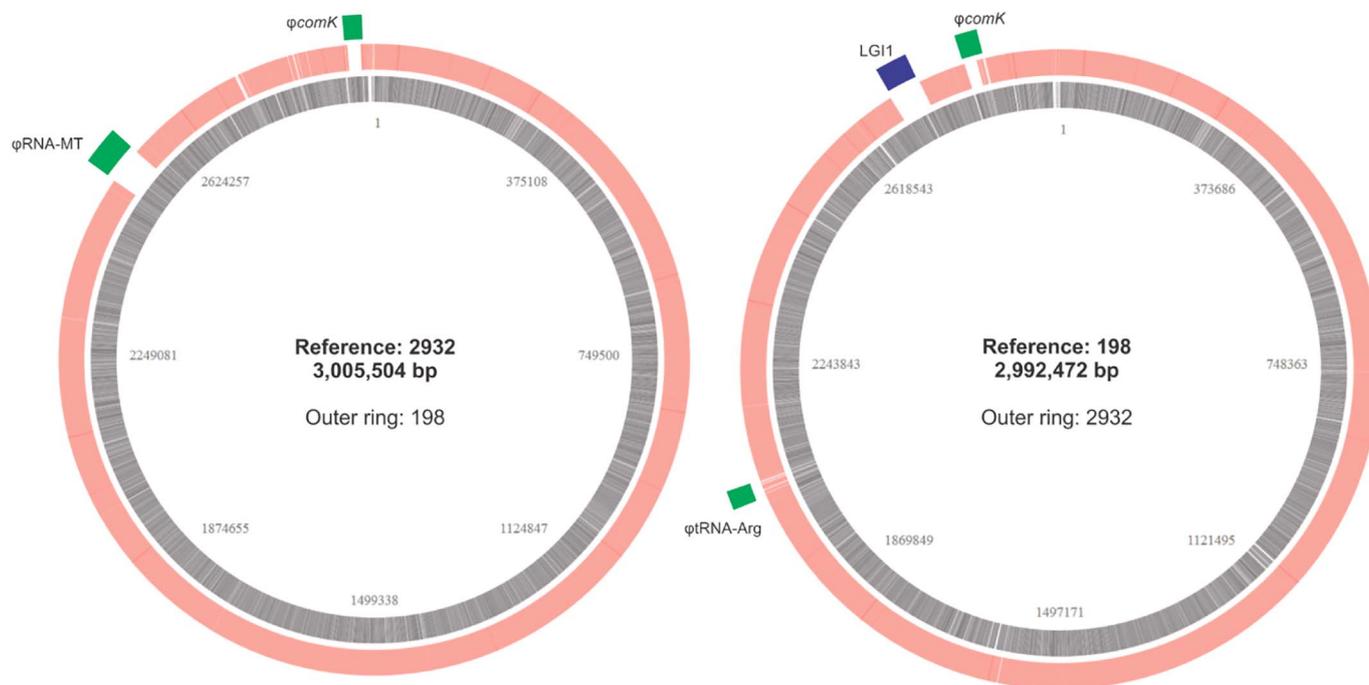[a] Incomplete phage (*i.e.* phage is split over multiple contigs, and size noted refers to known sequence).

Fig. 3. Whole genome comparisons of strains 2932 and 198 visualised in a local implementation of BRIG using custom scripts (https://github.com/esteinig/brigD3https://github.com/esteinig/brigD3). Phage inserts (φRNA-MT, φcomK, φtRNA-Arg) as well as *Listeria* genomic island 1 (LGI1) are marked.

described antimicrobial resistance genomic island LGI1 (Gilmour et al., 2010; Kovacevic et al., 2016). This mobile insert confers resistance to quaternary ammonium compounds (QACs), and also encodes a type IV secretion system. The results from this study suggest that strain 198 may have a considerably higher tolerance to the effects of environmental stresses in comparison to strain 2932, representing a marked potential difference between these two closely related isolates that was successfully identified using WGS analysis.

3.4. Virulence factors and pathogenic potential of strains 198 and 2932

The identification of closely related or indistinguishable isolates *via* PFGE represents a fundamental step both for the detection of "persistent" *L. monocytogenes* strains in a particular environment, and for the successful tracing of a listeriosis outbreak. In both instances, one essential piece of information that is missing is a detailed knowledge of an isolate's pathogenic potential. With regard to persistence, this information is important for determining the associated public health risks posed by a recurrent strain, while similarly the pathogenic potential of a *L. monocytogenes* isolate is a pivotal attribute in determining whether or not it represents the possible origin of a listeriosis outbreak. The pathogenic potential of a given *L. monocytogenes* strain is

determined by the functionality of a large number of genes known as "virulence factors", all of which have different roles in the various stages of the infection cycle.

Some of the main *L. monocytogenes* virulence factors, including the two *Listeria* pathogenicity islands LIPI-1 (Vázquez-Boland et al., 2001) and LIPI-3 (Cotter et al., 2008), as well as the internalin gene family (Bierne, Sabet, Personnic, & Cossart, 2007; Hamon, Bierne, & Cossart, 2006), were examined in strains 198 and 2932 with the aim of evaluating the true extent of their pathogenic potential. This was achieved by comparative analysis with the genomes of other well-annotated *L. monocytogenes* strains (EGDe and F2365), as well as the aforementioned strains 944 and 2993. With regard to LIPI-1, which harbours numerous important virulence genes such as the main transcriptional activator of *L. monocytogenes* virulence (*prfA*) and the vacuolar escape factor (*hly*) (Camejo et al., 2011), strains 198 and 2932 were determined to contain complete and intact islands within their genomes that shared 99–100% amino acid identity with the corresponding island in *L. monocytogenes* strain EGDe (Fig. 4A). However, both strains 198 and 2932 completely lacked the *Listeria* pathogenicity island LIPI-3 (Fig. 4C), which is associated with enhancing the virulence capabilities of *L. monocytogenes* by expressing a haemolytic and cytotoxic factor known as listeriolysin S (Cotter et al., 2008). In addition to these observations for LIPI-3, strains

Table 4
Multidrug, heavy metal and antibiotic resistance genes present in each *L. monocytogenes* serotype 1/2c strain.

Function	Strain 198 locus tag (AOA13_)	Strain 2932 locus tag (AOB47_)
Multidrug resistance	386c, 387, 449, 527, 618, 766c, 775, 1237, 1446c, 1480c, 1481c, 1540, **1592c**, 1959, 2620, 2623c, 2639c, 2643, 2650, 2678, 2761, 2866c, 2867c	03, 109c, 110c, 619c, 621, 682, 760, 852, 1001c, 1010, 1430, 1638c, 1672c, 1673c, 2058, 2690, 2693c, 2709c, 2713, 2720, 2748
β-lactamase & metallo-beta-lactamase proteins	63, 729, 1262c, 1406c, 1443c, 1676, 2044c, 2054c	294, 963, 1455c, 1598c, 1635c, 1776, 2144c, 2154c
Quinolone resistance	1077c, 2720c	1266c, 2791c
Lincomycin resistance	299, 2473c	531, 2545c
Tetracycline resistance	620c	854c
Antibiotic resistance	1484c	1676c
Heavy metal transport	1274c, 1738c, 1948c, p33c, **p38c**, p43c, **p44c**	1466c, 1838c, 2047c
Lead/cadmium/zinc resistance	422, 1466c, 1525, 2118, 2304c, 2458c, p34c, p35c, p39c, p62c	655, 1352, 1353, 1839c, 2218, 2379c, 2531c

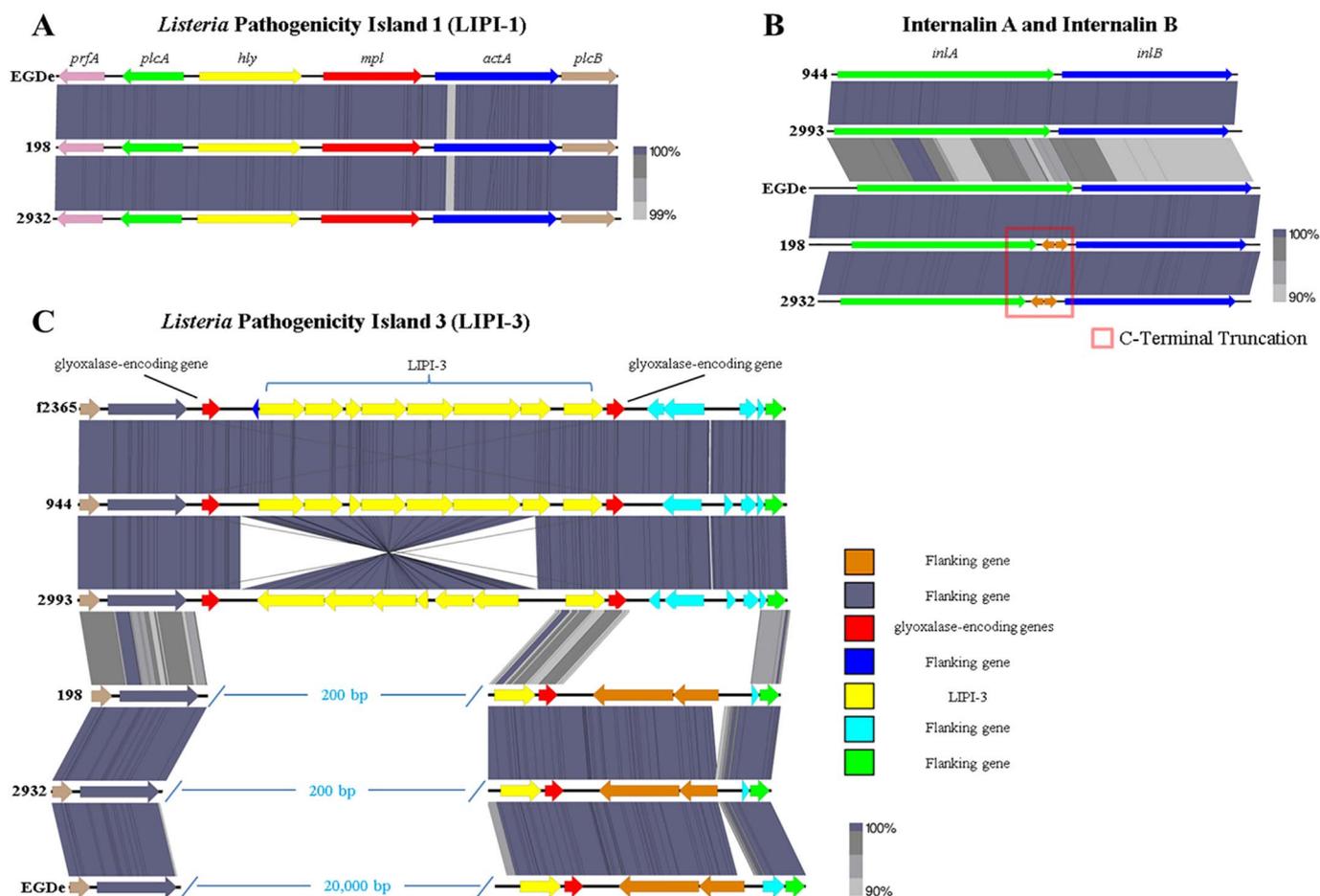Genes highlighted in bold are specifically discussed in the text.

Fig. 4. Linear comparisons between LIPI-1 (A), Internalin A and Internalin B (B) and LIPI-3 (C) for each of the *L. monocytogenes* input strains, visualised in Easyfig.

198 and 2932 both harboured a truncated *inlA* gene within their respective genomes (Fig. 4B), resulting from a point mutation (guanine to adenine) at gene position 2054, which caused a premature stop codon at amino acid residue 685 of the Internalin A (InlA) protein product. InlA is an important cell-surface internalin with a role in early invasive disease, facilitating *L. monocytogenes*' crossing of the intestinal epithelium through interaction with the host-cell receptor E-cadherin (Bierne et al., 2007). This truncation has been previously observed in a number of *L. monocytogenes* isolates (Ragon et al., 2008; Van Stelten, Simpson, Ward, & Nightingale, 2010). As a result of this premature stop codon, the translated InlA protein products in each of strains 198 and 2932 lacked the LPXTG domain required by Sortase A to recognize and anchor the protein to the membrane (Carvalho, Sousa, & Cabanes, 2014). This protein was therefore thought to be secreted or non-functional in strains 198 and 2932, which previous research has demonstrated (Nightingale, Windham, Martin, Yeung, & Wiedmann, 2005; Olier et al., 2003) could greatly reduce the virulence potential of these two strains. The combined results of this analysis indicate that both strains 198 and 2932 may be heavily attenuated in terms of their pathogenic potential, and this additional information about these genomes could prove extremely significant when evaluating their relevance to public health.

Recently, WGS analysis has emerged as the preferred technique for subtyping *L. monocytogenes* isolates associated with listeriosis outbreaks, exemplified by the increased use of this method by the CDC for pathogen detection and tracking in the United States. The results presented in this research justify the emerging preference for WGS analysis, by demonstrating the ability of this method to overcome the limitations of more commonly used techniques such as PFGE and MLST. Similar results to those recorded in this study were observed in other research involving isolates from the 2008 Canadian listeriosis outbreak

(Gilmour et al., 2010), whereby two strains that were initially presumed to be distinct from one another *via* PFGE were identified to in fact be very closely related, differing from one another only in terms of a single prophage insertion, and likely shared a recent common ancestor. Of particular note is the presence of LGI1 first described in that study, which was present in strain 198 in this study, however was absent in strain 2932. In addition to conferring resistance to QACs, this mobile genetic element encodes a Type-IV secretion system and additional genes with functionality in translocation, stress resistance and regulation. Likewise, another study involving isolates from the 2011–2013 Austrian/German listeriosis outbreak (Schmid et al., 2014) demonstrated that WGS was capable of discriminating *L. monocytogenes* clones that were otherwise indistinguishable from one another on the basis of PFGE. The advent of WGS as a viable subtyping tool for *L. monocytogenes* has been made possible by the hardware and software advancements that have been made in sequencing technologies over the last number of years. In addition, the cost of WGS of bacterial genomes continues to decrease, with numerous high-throughput sequencing technologies now allowing for the simultaneous sequencing of multiple *L. monocytogenes* isolates at a rate of less than $100 per isolate (Stasiewicz, den Bakker, & Wiedmann, 2015), making WGS a more viable option, even for small laboratories with stringent financial limitations.

4. Conclusions

The advent of WGS has changed the overarching structure of modern day bacterial research. Over the years, PFGE has established itself as an integral method for the detection and curtailment of listeriosis outbreaks, however with the continued steady advances being

made in the field of bioinformatics, far more sensitive techniques for subtyping *L. monocytogenes* isolates have emerged. Whole genome sequence analysis is one such technique which has recently become the preferred tool of choice for differentiating *L. monocytogenes* strains by the CDC. The results of this research demonstrate the true power of this technique not only for identifying key differences between closely related pairs of *L. monocytogenes* isolates, but also for identifying clinically and environmentally relevant characteristics of such strains, which may have an important relevance to public health. WGS analysis was demonstrated to be proficient in determining important genomic differences between strains 198 and 2932 in terms of their external stress and antibiotic resistances, as well as showing a capacity for detecting strain-specific characteristics such as the pathogenic potential of an isolate (exemplified by the absence of LIPI-3 and a truncated *inlA* gene in strains 198 and 2932 indicating an attenuation in virulence of both strains). Therefore, the findings from this study conclude that WGS is a more powerful method for identifying differences in physical and genotypic characteristics between bacterial strains, and demonstrate the merit of using WGS analysis as the subtyping technique of choice for *L. monocytogenes*, in place of PFGE.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ifset.2017.07.010.

Accessibility of the data

The whole-genome shotgun projects for these *L. monocytogenes* strains have been deposited at DDBJ/EMBL/GenBank under the accession numbers LJPD00000000 (strain 944), LJPE00000000 (strain 2932), LJPF00000000 (strain 2993), and LJOZ00000000 (strain 198). The versions described in this paper are versions LJPD01000000, LJPE01000000, LJPF01000000, and LJOZ01000000, respectively.

References

Alikhan, N.-F., Petty, N. K., Zakour, N. L. B., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics, 12*, 402.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Research, 25*, 3389–3402.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: Rapid annotations using subsystems technology. *BMC Genomics, 9*, 75.

Bierne, H., Sabet, C., Personnic, N., & Cossart, P. (2007). Internalins: A complex family of leucine-rich repeat-containing proteins in *Listeria monocytogenes*. *Microbes and Infection, 9*, 1156–1166.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 15*, 2114–2120.

Brosch, R., Brett, M., Catimel, B., Luchansky, J. B., Ojeniyi, B., & Rocourt, J. (1996). Genomic fingerprinting of 80 strains from the WHO multicenter international typing study of *Listeria monocytogenes* via pulsed-field gel electrophoresis (PFGE). *International Journal of Food Microbiology, 32*, 343–355.

Buchrieser, C., Rusniok, C., Kunst, F., Cossart, P., & Glaser, P. (2003). Comparison of the genome sequences of *Listeria monocytogenes* and *Listeria innocua*: Clues for evolution and pathogenicity. *FEMS Immunology and Medical Microbiology, 35*, 207–213.

Camejo, A., Carvalho, F., Reis, O., Leitão, E., Sousa, S., & Cabanes, D. (2011). The arsenal of virulence factors deployed by *Listeria monocytogenes* to promote its cell infection cycle. *Virulence, 2*, 379–394.

Carvalho, F., Sousa, S., & Cabanes, D. (2014). How *Listeria monocytogenes* organizes its surface for virulence. *Frontiers in Cellular and Infection Microbiology, 4*, 1–22.

Casey, A., Fox, E. M., Schmitz-Esser, S., Coffey, A., McAuliffe, O., & Jordan, K. (2014). Transcriptome analysis of *Listeria monocytogenes* exposed to biocide stress reveals a multi-system response involving cell wall synthesis, sugar uptake, and motility. *Frontiers in Microbiology, 5*, 1–14.

Centre for Disease Control. *Multistate outbreak of listeriosis linked to blue bell creameries products (final update)*. (2015). http://www.cdc.gov/listeria/outbreaks/ice-cream-03-15/index.html (Retrieved from: Accessed 22nd October 2016).

Centre for Disease Control. *multistate outbreak of listeriosis linked to commercially produced, prepackaged caramel apples made from Bidart Bros. Apples (final update)*. (2015). http://www.cdc.gov/listeria/outbreaks/caramel-apples-12-14/index.html (Retrieved from: Accessed 22nd October 2016).

Centre for Disease Control. *Multistate outbreak of listeriosis linked to frozen vegetables (final update)*. (2015). https://www.cdc.gov/listeria/outbreaks/frozen-vegetables-05-16/ (Retrieved from: Accessed 22nd October 2016).

Chan, Y. C., & Wiedmann, M. (2008). Physiology and genetics of *Listeria monocytogenes* survival and growth at cold temperatures. *Critical Reviews in Food Science and Nutrition, 49*, 237–253.

Cotter, P. D., Draper, L. A., Lawton, E. M., Daly, K. M., Groeger, D. S., Casey, P. G., ... Hill, C. (2008). Listeriolysin S, a novel peptide haemolysin associated with a subset of lineage I *Listeria monocytogenes*. *PLoS Pathogens, 4*, e1000144.

Dalmasso, M., & Jordan, K. (2014). *Pulsed-field gel electrophoresis (PFGE) analysis of Listeria monocytogenes*. New York: Springer.

Darling, A. C., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research, 14*, 1394–1403.

Donaldson, J. R., Nanduri, B., Burgess, S. C., & Lawrence, M. L. (2009). Comparative proteomic analysis of *Listeria monocytogenes* strains F2365 and EGD. *Applied and Environmental Microbiology, 75*, 366–373.

Fox, E. M., Garvey, P., McKeown, P., Cormican, M., Leonard, N., & Jordan, K. (2012). PFGE analysis of *Listeria monocytogenes* isolates of clinical, animal, food and environmental origin from Ireland. *Journal of Medical Microbiology, 61*, 540–547.

Fox, E. M., Leonard, N., & Jordan, K. (2011). Physiological and transcriptional characterisation of persistent and non-persistent *Listeria monocytogenes* isolates. *Applied and Environmental Microbiology, 77*, 6559–6569.

Fugett, E. B., Schoonmaker-Bopp, D., Dumas, N. B., Corby, J., & Wiedmann, M. (2007). Pulsed-field gel electrophoresis (PFGE) analysis of temporally matched *Listeria monocytogenes* isolates from human clinical cases, foods, ruminant farms, and urban and natural environments reveals source-associated as well as widely distributed PFGE types. *Journal of Clinical Microbiology, 45*, 865–873.

Gilmour, M. W., Graham, M., Van Domselaar, G., Tyler, S., Kent, H., Trout-Yakel, K. M., ... Nadon, C. (2010). High-throughput genome sequencing of two *Listeria monocytogenes* clinical isolates during a large foodborne outbreak. *BMC Genomics, 11*, 120.

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

Goering, R. V. (2010). Pulsed field gel electrophoresis: A review of application and interpretation in the molecular epidemiology of infectious disease. *Infection, Genetics and Evolution, 10*, 866–875.

Gottlieb, S. L., Newbern, E. C., Griffin, P. M., Graves, L. M., Hoekstra, R. M., Baker, N. L., et al. (2006). Multistate outbreak of listeriosis linked to turkey deli meat and subsequent changes in US regulatory policy. *Clinical Infectious Diseases, 42*, 29–36.

Goulet, V., King, L. A., Vaillant, V., & de Valk, H. (2013). What is the incubation period for listeriosis? *BMC Infectious Diseases, 13*, 11.

Graves, L. M., Hunter, S. B., Ong, A. R., Schoonmaker-Bopp, D., Hise, K., Kornstein, L., ... Swaminathan, B. (2005). Microbiological aspects of the investigation that traced the 1998 outbreak of listeriosis in the United States to contaminated hot dogs and establishment of molecular subtyping-based surveillance for *Listeria monocytogenes* in the PulseNet network. *Journal of Clinical Microbiology, 43*, 2350–2355.

Graves, L. M., & Swaminathan, B. (2001). PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. *International Journal of Food Microbiology, 65*, 55–62.

Greenberg, J., & Elliott, C. (2009). A cold cut crisis: Listeriosis, Maple Leaf Foods, and the politics of apology. *Canadian Journal of Communication, 34*, 189–204.

Hain, T., Ghai, R., Billion, A., Kuenne, C. T., Steinweg, C., Izar, B., et al. (2012). Comparative genomics and transcriptomics of lineages I, II, and III strains of *Listeria monocytogenes*. *BMC Genomics, 13*, 144.

Hamon, M., Bierne, H., & Cossart, P. (2006). *Listeria monocytogenes*: A multifaceted model. *Nature Reviews Microbiology, 4*, 423–434.

Herschleb, J., Ananiev, G., & Schwartz, D. C. (2007). Pulsed-field gel electrophoresis. *Nature Protocols, 2*, 677–684.

Khen, B., Lynch, O., Carroll, J., McDowell, D., & Duffy, G. (2015). Occurrence, antibiotic resistance and molecular characterization of *Listeria monocytogenes* in the beef chain in the Republic of Ireland. *Zoonoses and Public Health, 62*, 11–17.

Kovacevic, J., Ziegler, J., Wałecka-Zacharska, E., Reimer, A., Kitts, D. D., & Gilmour, M. W. (2016). Tolerance of *Listeria monocytogenes* to quaternary ammonium sanitizers is mediated by a novel efflux pump encoded by *emrE*. *Applied and Environmental Microbiology, 82*, 939–953.

Laksanalamai, P., Joseph, L. A., Silk, B. J., Burall, L. S., Tarr, C. L., Gerner-Smidt, P., & Datta, A. R. (2012). Genomic characterization of *Listeria monocytogenes* strains involved in a multistate listeriosis outbreak associated with cantaloupe in US. *PLoS One, 7*, e42448.

Li, Q., Sherwood, J., & Logue, C. (2007). Antimicrobial resistance of *Listeria* spp. recovered from processed bison. *Letters in Applied Microbiology, 44*, 86–91.

Lundén, J., Autio, T., Markkula, A., Hellström, S., & Korkeala, H. (2003). Adaptive and cross-adaptive responses of persistent and non-persistent *Listeria monocytogenes* strains to disinfectants. *International Journal of Food Microbiology, 82*, 265–272.

Lungu, B., O'Bryan, C. A., Muthaiyan, A., Milillo, S. R., Johnson, M. G., Crandall, P. G., & Ricke, S. C. (2011). *Listeria monocytogenes*: Antibiotic resistance in food production. *Foodborne Pathogens and Disease, 8*, 569–578.

Metselaar, K. I., den Besten, H. M., Abee, T., Moezelaar, R., & Zwietering, M. H. (2013). Isolation and quantification of highly acid resistant variants of *Listeria monocytogenes*. *International Journal of Food Microbiology, 166*, 508–514.

Nightingale, K., Windham, K., Martin, K., Yeung, M., & Wiedmann, M. (2005). Select *Listeria monocytogenes* subtypes commonly found in foods carry distinct nonsense mutations in inlA, leading to expression of truncated and secreted internalin A, and are associated with a reduced invasion phenotype for human intestinal epithelial cells. *Applied and Environmental Microbiology, 71*, 8764–8772.

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., et al. (2013). Assembling genomes and mini-metagenomes from highly chimeric reads. *Research in computational molecular biology* (pp. 158–170). Springer-Verlag Berlin Heidelberg.

Olier, M., Pierre, F., Rousseaux, S., Lemaître, J.-P., Rousset, A., Piveteau, P., & Guzzo, J. (2003). Expression of truncated internalin A is involved in impaired internalization of some *Listeria monocytogenes* isolates carried asymptomatically by humans. *Infection*

and Immunity, 71, 1217–1224.

Olsen, S. J., Patrick, M., Hunter, S. B., Reddy, V., Kornstein, L., MacKenzie, W. R., et al. (2005). Multistate outbreak of *Listeria monocytogenes* infection linked to delicatessen turkey meat. *Clinical Infectious Diseases, 40*, 962–967.

Proft, T., & Baker, E. (2009). Pili in Gram-negative and Gram-positive bacteria—Structure, assembly and their role in disease. *Cellular and Molecular Life Sciences, 66*, 613–635.

Pulsenet (2013). *Standard operating procedure for PulseNet PFGE of Listeria monocytogenes*.

Ragon, M., Wirth, T., Hollandt, F., Lavenir, R., Lecuit, M., Le Monnier, A., & Brisse, S. (2008). A new perspective on *Listeria monocytogenes* evolution. *PLoS Pathogens, 4*, e1000146.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.-A., & Barrel, B. (2000). Artemis: Sequence visualization and annotation. *Bioinformatics, 16*, 944–945.

Rychli, K., Müller, A., Zaiser, A., Schoder, D., Allerberger, F., Wagner, M., & Schmitz-Esser, S. (2014). Genome sequencing of *Listeria monocytogenes* "quargel" listeriosis outbreak strains reveals two different strains with distinct in vitro virulence potential. *PloS One, 9*, e89964.

Schmid, D., Allerberger, F., Huhulescu, S., Pietzka, A., Amar, C., Kleta, S., ... Mellmann, A. (2014). Whole genome sequencing as a tool to investigate a cluster of seven cases of listeriosis in Austria and Germany, 2011–2013. *Clinical Microbiology and Infection, 20*, 431–436.

Stachowiak, R., Jagielski, T., Roeske, K., Osińska, O., Gunerka, P., Wiśniewski, J., & Bielecki, J. (2015). Lmo0171, a novel internalin-like protein, determines cell morphology of *Listeria monocytogenes* and its ability to invade human cell lines. *Current*

Microbiology, 70, 267–274.

Stasiewicz, M. J., den Bakker, H. C., & Wiedmann, M. (2015). Genomics tools in microbial food safety. *Current Opinion in Food Science, 4*, 105–110.

Sullivan, M. J., Petty, N. K., & Beatson, S. A. (2011). Easyfig: A genome comparison visualizer. *Bioinformatics, 27*, 1009–1010.

Swaminathan, B., & Gerner-Smidt, P. (2007). The epidemiology of human listeriosis. *Microbes and Infection, 9*, 1236–1243.

Tenover, F. C., Arbeit, R. D., & Goering, R. V. (1997). How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections a review for healthcare epidemiologists. *Infection Control, 18*, 426–439.

Tenover, F. C., Arbeit, R. D., Goering, R. V., Mickelsen, P. A., Murray, B. E., Persing, D. H., & Swaminathan, B. (1995). Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: Criteria for bacterial strain typing. *Journal of Clinical Microbiology, 33*, 2233–2239.

Van Stelten, A., Simpson, J., Ward, T., & Nightingale, K. (2010). Revelation by single-nucleotide polymorphism genotyping that mutations leading to a premature stop codon in inlA are common among *Listeria monocytogenes* isolates from ready-to-eat foods but not human listeriosis cases. *Applied and Environmental Microbiology, 76*, 2783–2790.

Vázquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Dominguez-Bernal, G., Goebel, W., ... Kreft, J. (2001). *Listeria* pathogenesis and molecular virulence determinants. *Clinical Microbiology Reviews, 14*, 584–640.

Zhou, Y., Liang, Y., Lynch, K., Dennis, J. J., & Wishart, D. S. (2011). PHAST: A fast phage search tool. *Nucleic Acids Research.* http://dx.doi.org/10.1093/nar/gkr485.