

Northumbria Research Link

Citation: Palaiodimou, Lydia, Fanning, Séamus and Fox, Edward M. (2021) Genomic insights into persistence of *Listeria* species in the food processing environment. *Journal of Applied Microbiology*, 131 (5). pp. 2082-2094. ISSN 1364-5072

Published by: Wiley-Blackwell

URL: <https://doi.org/10.1111/jam.15089> <<https://doi.org/10.1111/jam.15089>>

This version was downloaded from Northumbria Research Link:
<http://nrl.northumbria.ac.uk/id/eprint/45796/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)



**Northumbria
University**
NEWCASTLE



UniversityLibrary

DR. EDWARD FOX (Orcid ID : 0000-0001-5850-7558)

Article type : Original Article

Journal of Applied Microbiology - ORIGINAL ARTICLE

Genomic insights into persistence of *Listeria* species in the food processing environment

Lydia Palaiodimou¹, Séamus Fanning^{2,3}, Edward M. Fox^{1*}

¹Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK.

²UCD-Centre for Food Safety, School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin D04 N2E5, Ireland.

³Institute for Global Food Security, Queen's University Belfast, Chlorine Gardens, Belfast, BT5 6AG, United Kingdom

Running headline: Persistence of *Listeria* species

***Author for correspondence:**

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/JAM.15089](https://doi.org/10.1111/JAM.15089)

This article is protected by copyright. All rights reserved

Dr Edward M. Fox, edward.fox@northumbria.ac.uk

Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK.

ABSTRACT

AIMS

Listeria species may colonise and persist in food processing facilities for prolonged periods of time, despite hygiene interventions in place. To understand the genetic factors contributing to persistence of *Listeria* strains, this study undertook a comparative analysis of 7 persistent and 6 presumed non-persistent strains, isolated from a single food processing environment, to identify genetic markers correlating to promoting persistence of *Listeria* strains, through whole genome sequence analysis.

METHODS AND RESULTS

A diverse pool of genetic markers relevant to hygiene tolerance were identified, including disinfectant resistance markers *qacH*, *emrC*, and the efflux cassette *bcrABC*. Both persistent and presumed non-persistent cohorts encoded a range of stress resistance markers, including heavy metal resistance, oxidative and pH stress, although trends were associated with each cohort (e.g. *qacH* and *cadA1C* resistance was more frequently found in persistent isolates). Persistent isolates were more likely to contain mutations associated with attenuated virulence, including a truncated *InlA*. Plasmids and transposons were widespread between cohorts.

CONCLUSIONS

Results suggest no single genetic marker identified was universally responsible for a strains ability to persist. Persistent strains were more likely to harbour mutation associated with hypovirulence.

SIGNIFICANCE AND IMPACT OF THE STUDY

This study provides additional insights into the distribution of genetic elements relevant to persistence across *Listeria* species, as well as strain virulence potential.

KEYWORDS: *Listeria*; Food safety; Disinfection; Virulence; Plasmids.

INTRODUCTION

Listeria species comprise an expanding genus of bacteria, which to date includes 21 recorded species, many of which are relatively recently described (Leclercq et al., 2019, Quereda et al., 2020). Of these species, *L. monocytogenes* is of primary concern to public health, although *L. ivanovii* is an important pathogen of animals (Orsi and Wiedmann, 2016). These bacteria can be found ubiquitously in the environment, and may contaminate foods including ready-to-eat foods, vegetable, seafood, meat, eggs, and dairy products; thus, incidence of disease is mainly linked to infections *via* foodborne transmission (Fugett et al., 2007, Scallan et al., 2011, McAuley et al., 2014). Both pathogenic and non-pathogenic species are known to share common niches, and as such non-pathogenic *Listeria* species, such as *L. innocua* and *L. welshimeri*, may be used as index organisms for potential contamination and/or colonisation of *L. monocytogenes* in food processing environments (FPEs). Hygiene regimes, which include cleaning, sanitising and disinfection cycles, are among the primary interventions applied in FPEs to control *Listeria* species. This includes the use of antimicrobial agents such as quaternary ammonium compounds (QACs), as well as other antimicrobial formulations.

The persistence of *L. monocytogenes* has frequently been reported in FPEs, with several studies from various locations reporting re-isolation of the same clone over extended periods up to and exceeding 10 years (Wulff et al., 2006, Chambel et al., 2007, Lomonaco et al., 2009, Fox et al., 2011b). This presents an important challenge to food producers, as these persistent contaminants are associated with an increased likelihood of cross-contamination of food products produced, since they are not eliminated from the FPE. Persistence may be caused, or at least contributed to, by several factors such as: poor hygiene practice and/or ineffective sanitisers; harbourage sites in the FPE, such as damaged equipment of surfaces; the presence of genetic markers providing a competitive advantage to persistent strains; the efficient production of biofilms by persistent strains, or interactions with native microbiota (Fox et al., 2014, Fox et al., 2015, Schmitz-Esser et al., 2015, Harter et al., 2017, Rodríguez-Campos et al., 2019). Although a number of genetic elements have been purported to play a role in colonisation and persistence dynamics, the

nature of the role of these mechanisms to the persistence phenomenon remains poorly understood. This may, in part, be due to the different environmental conditions across sometimes disparate FPEs, which may vary in aspects such as hygiene systems, temperature conditions, and resident microbiota, among other factors. A particular challenge to studying the persistence phenotype, is the difficulty in replicating experimental conditions amenable to its study under laboratory conditions. Similarly, introduction of pathogenic *L. monocytogenes* to an FPE to examine colonisation dynamics represents an unacceptable food safety risk. To further elucidate the potential genetic factors that may promote persistence of *Listeria* strains, and to understand other relevant aspects of interest such as pathogenic potential, this study aimed to characterise 7 persistent and 6 presumed non-persistent strains, isolated from the same FPE, over the same timeframe, to correlate genetic traits across persistent and/or non-persistent cohorts. This would facilitate the comparison of strains experiencing comparable environmental selection, such as antimicrobial agents in use, temperatures, resident microbiota, and surface materials. This study included persistent and presumed non-persistent strains from 3 species (*L. monocytogenes*, *L. welshimeri*, and *L. innocua*), to examine pangenome markers across multiple members of the genus *Listeria*.

MATERIALS AND METHODS

Bacterial isolates in this study.

This study characterized 13 *Listeria* strains collected over a 2-year period from a meat processing facility in Ireland. Seven of them (*L. monocytogenes*: UC DL011, UC DL016, UC DL019, UC DL187; *L. innocua*: UC DL146; and *L. welshimeri*: UC DL122) were characterized as 'persistent' contaminants, representing pulsed-field gel electrophoresis (PFGE) pulsotypes isolated multiple times over 6 months or more. All isolates were subjected to the PulseNet Standard PFGE method for subtyping *Listeria monocytogenes* (PulseNet, 2013), utilising two restriction enzymes (namely *Ascl* and *Apal*). Isolates were classified as the same strain based on an indistinguishable PFGE pulsotype considering

fingerprints of both enzymes (i.e. 100% similarity score). The other six isolates (*L. monocytogenes*: UC DL037, UC DL133, UC DL150, UC DL175; *L. innocua*: UC DL085; and *L. welshimeri*: UC DL063) were designated as 'presumed non-persistent' contaminants, comprising genotypes only identified a single time over the 2-year surveillance period.

Genome assembly and annotation

Genomic DNA was extracted using the QIAGEN DNeasy kit (QIAGEN, Germany). Sample quality was confirmed using a NanoDrop instrument (Thermo Fisher Scientific, Waltham, MA) to confirm 260:280 nm and 260:230 nm ratios between 1.8 and 2.0. Library preparation using genomic DNA of isolates was performed using the Nextera XT library prep kit (illumina, San Diego, CA). Raw read sequences were then generated using 250-bp paired end sequencing on the MiSeq platform (illumina). The raw read quality was assessed with FastQC (version 0.11.8). These raw reads were subsequently processed to remove adapter sequences and low quality reads using Trimmomatic software v0.22 (Bolger et al., 2014). Draft genomes were assembled using SPAdes (Species Prediction and Diversity Estimation) software v2.5.1 based on an algorithm which employs multi-sized De Bruijn graphs with *k*-mer values of '21, 33, 55, 77' to construct the contiguous sequences (Bankevich et al., 2012). All draft genomes were annotated using the RAST online platform tool, and using Prokka algorithms (Aziz et al., 2008, Seemann, 2014).

Molecular subtyping of isolates

The serotype of *L. monocytogenes* isolates was identified using previously described *in silico* schemes (Doumith et al., 2004). Strain MLST type was derived using the seven housekeeping gene targets previously described (Ragon et al., 2008), and referencing the Institut Pasteur BIGSdb-*Lm* database (<https://bigsdb.pasteur.fr/listeria>). Novel alleles were submitted to the Institut Pasteur BIGSdb-*Lm* database for assignment of novel STs.

Genome screening for molecular markers, and comparative visualisation of sequence data

A strain BLAST database was created using the Geneious Prime software platform (Kearse et al., 2012). Additional databases were created comprising genes of interest relating to virulence, stress resistance, or other features such as mobile genetic elements, as detailed in **Supplementary Table 1**. Sequence alignments were performed using MAFFT program (Kato et al., 2002). EasyFig software was utilised to visualise sequence alignment similarities, including transposon and phage alignments (Sullivan et al., 2011). The BLAST ring image generator (BRIG) platform was used to visualise BLAST comparisons using constructed pangenome references (Alikhan et al., 2011).

Pangenome analysis, core SNP analysis and phylogenetic tree construction

Pan-genome analysis was performed utilising the Roary pipeline (Page et al., 2015), and maximum likelihood phylogenetic trees were constructed using RAxML (Stamatakis, 2014). Pangenome interrogation for phage insert regions was performed using the online PHASTER tool (Arndt et al., 2016). Core SNP analysis was conducted utilising the Snippy pipeline (Seemann, 2015). Phylogenetic analysis of core genome alignments was performed using FastTree, with the GTR+CAT model (Price et al., 2009). Plasmid searches were performed using PlasmidFinder 2.1, interrogating against the Gram-Positive database (Carattoli et al., 2014), coupled with BLAST searches of draft genomes utilising published plasmid sequences pLI100, pN1-011A, and p6179 (NCBI accession numbers NC_003383, NC_022045, and NZ_HG813250, respectively).

Draft sequence archiving

Draft genome sequences for strains from this project have been deposited in the NCBI genome database, BioProject PRJNA692370, under the following accessions: UC DL011, SAMN17321004; UC DL016, SAMN17321005; UC DL019, SAMN17321006; UC DL037, SAMN17321007; UC DL063, SAMN17321008; UC DL085, SAMN17321009; UC DL122, SAMN17321010; UC DL133, SAMN17321011; UC DL146, SAMN17321012; UC DL150, SAMN17321013; UC DL162, SAMN17321014; UC DL175, LXR D000000000; UC DL187, SAMN17321015.

RESULTS

Overview of the genomes.

An overview of the genetic subtypes and genome characteristics of isolates in this study are presented in **Table 1**. The *L. monocytogenes* strains' genomes ranged from 2,913,758 bp to 3,080,560 bp, the *L. innocua* from 2,991,782 bp to 3,026,780 bp, and *L. welshimeri* from 2,856,944 bp to 2,946,539 bp. GC content was lowest among *L. welshimeri* (36.1-36.3%), followed by *L. innocua* (37.3%), and highest among *L. monocytogenes* strains (37.8-37.9%). The total pangenome size was 7,669 CDS sequences, of which 1,314 were core to all strains (17%; **Figure 1**). *Listeria monocytogenes* shared a greater number of genes exclusively with *L. innocua* ($n=477$), relative to genes exclusive with *L. welshimeri* ($n=178$); this is supported by the likelihood of these species being closer in evolutionary terms (Orsi and Wiedmann, 2016). Among the *L. monocytogenes* strains, clonal complex 9 (CC9) strains were most common (50%, 4/8); this included 3 ST9 isolates, and a single allele variant (ST622). Other STs identified included ST1, ST121 and ST204. Among the *L. welshimeri* strains was ST168 and ST1084 isolates, along with a novel ST not previously described, designated as ST2688 in the Institut Pasteur BIGSdb-*Lm* database. For *L. innocua*, ST602 and ST1008 were observed.

Stress resistance markers.

Three disinfectant resistance markers were identified among strains in this study: *bcrABC*, *emrC*, and *qacH*. The most common of these genetic markers found was *bcrABC*, present in 4 strains, followed by *emrC* in 3 strains, and *qacH* in 2 strains (**Figure 2**). Interestingly, strains only harboured one of these disinfectant resistance markers. When considering persistent and presumed non-persistent cohorts, 5/7 persisters harboured disinfectant resistance markers (71%), compared with 4/6 non-persisters (67%). The *qacH* marker was only identified in persistent strains (2/7), while the other 2 markers were present in both cohort groups.

Cadmium resistance cassettes were prevalent among both persistent (86%) and presumed non-persistent (83%) cohorts; this included the *cadA1*, *cadA2*, and *cadA3* variants of this resistance system (**Figure 2**). Of these, *cadA1* was more frequent among persisters compared with non-persisters (57% versus 17%), while *cadA4* was only identified in non-persisters.

Listeria Genomic island 2 (LGI2) encodes a large arsenic resistance operon (*arsD1A1R1D2R2A2B1B2*), and was identified in two presumed non-persistent isolates: UC DL037 and UC DL175. Another known arsenic cassette carrying element, a Tn544 transposon, was found in 6 isolates (4 persistent and 2 presumed non-persistent). An additional screen of 100 isolates previously described by Hurley et al. (2019) was then analysed for the presence of *arsA1*, *arsA2*, and the Tn544 resistance cassette, with an overall prevalence of 12%, 2% and 1% identified, respectively.

The *L. monocytogenes* stress survival islets (SSIs) encode genetic mechanisms for resistance to stress conditions such as temperature, pH and osmotic stress (Ryan et al., 2010, Harter et al., 2017). Of the previously described SSIs, SSI-1 was most common; it was found in 8 isolates, including 5/7 persisters (71%) and 3/6 non-persisters (50%; **Figure 2**, **Supplementary Figure 1**). SSI-2 was identified in 3 isolates (UC DL019, UC DL085 and UC DL146).

Virulence markers.

The distribution of a number of important virulence markers is shown in **Figure 3**. None of the virulence genes in **Figure 3** were identified outside of *L. monocytogenes* in this study. All of the 12 internalins, however, were identified among the *L. monocytogenes* strains; of these, *inIA*, *inIB*, *inIC*, *inIE*, *inIF*, *inII*, *inIJ*, and *inIK*, were present in all strains. The *inIG* gene was absent in 2 persistent strains and a presumed non-persistent strain, while *inIC2*, *inID*, and *inIH*, were each absent from 4 strains (2 persistent and 2 presumed non-persistent). Premature stop codons (PMSCs) were identified in *inIA* sequences in 5 *L. monocytogenes* strains (all persisters and 1 non-persister); in addition, UCDL187 harboured PMSCs in *inIC*, in *ascB* and a hypothetical protein in its *inIC2DE* locus, and in the *prfA* virulence regulator (**Supplementary Figures-2 and -3**).

The *Listeria* Pathogenicity Islands (LIPI) -3 and -4 contribute to the pathogenesis of strains, and are associated with increased virulence in mammalian host infections (Cotter et al., 2008, Maury et al., 2016). LIPI-3 was identified in a single presumed non-persistent strain, the *L. monocytogenes* strain UCDL037; LIPI-4 was absent in all isolates in this study.

Mobile genetic elements.

Plasmids were identified in 10 of the strains in this study (77%); with three of these containing 2 plasmids (**Figure 4, Supplementary Figure 4**). Both *L. welshimeri* UCDL063 and UCDL122 contained identical plasmids (pUCDL063-1 and pUCDL122-1); two ST9 strains (UCDL016 and UCDL133) contained similar plasmids (pUCDL016-1 and pUCDL133-1), and a smaller plasmid of 4,265 bp was present in three isolates in this study (UCDL011, UCDL016, and UCDL133).

The ϕ *comK* phage insert was identified in 7 strains in this study (54%); this included 4 persistent (57%) and 3 presumed non-persistent isolates (50%, **Supplementary Figure 5**). This phage was identified in both *L. monocytogenes* and *L. innocua* species, but not in *L. welshimeri*.

The maximum likelihood tree generated from the pangenome analysis of all coding sequences among the strains in this study identified 3 clades, one representing each of the three species (**Supplementary Figure 6**). The *L. monocytogenes*-containing clade included a CC9 sub-clade, supporting their genetic similarity relative to other STs identified. The CC9 sub-clade was investigated through a core SNP analysis, and SNP frequencies supporting diverse strains (ranging from 198 to 1,236 SNPs between isolates (**Supplementary Figure 7**)).

DISCUSSION

This study characterised persistent and presumed non-persistent strains of three *Listeria* species. While a diverse population was noted (**Table 1, Supplementary Figure 6**), for the *L. monocytogenes* strains, CC9 was most common. This CC was represented among both persistent and presumed non-persistent isolates, suggesting it is associated with food-related niches. This finding supports previous similar studies from Ireland and Europe, which found a high incidence of CC9 among food-related sources (Ebner et al., 2015, Henri et al., 2016, Hurley et al., 2019). One each of ST1, ST121 and ST204 were also identified; the latter two STs also being among the more common food-related clonal subgroups (Schmitz-Esser et al., 2015, Jennison et al., 2017). Lineage II was the most common genetic lineage observed among the *L. monocytogenes* strains, which supports the assertion that this lineage has adapted to food-associated niches, compared with lineage I, which is more frequently associated with clinical incidence of disease (Maury et al., 2016, Jennison et al., 2017).

The ability of strains to colonise and persist in FPEs is thought to be a multi-faceted phenomenon. Strains must colonise an environmental niche, where efficient biofilm production is likely to be important (Norwood and Gilmour, 1999, Rodríguez-Campos et al., 2019), although a direct correlation with persistence has not always been identified (Djordjevic et al., 2002). Subsequently, these bacteria must tolerate a plethora of environmental stressful conditions, many which may be unfavourable or antagonistic to their

Accepted Article

survival. Exposure to disinfectants imposes a continual stress on bacterial species colonising FPEs, as these are central to hygiene efforts and food safety. Previous studies have suggested disinfectant resistance may be a feature of strains encountered, or persisting, in FPE's; this is thought to contribute to the dominance of ST121 and ST204 clonally related strains (Schmitz-Esser et al., 2015, Fox et al., 2016), however a positive association with persistence of specific strains in FPEs varies (Kastbjerg and Gram, 2009, Fox et al., 2011a, Rodríguez-Campos et al., 2019). A variety of disinfectant resistance markers have been identified in *Listeria* species, typically comprising efflux pump systems; these include the *bcrABC* cassette, *emrC*, *emrE*, *qacA*, *qacC*, *qacH*, and *qacEΔ1* determinants. These systems are typically associated with resistance to QACs, with this class of disinfectant in use at the food processing facility where these strains were isolated. Comparative analysis of the aforementioned genetic markers between persistent and presumed non-persistent isolates did not suggest a correlation with either phenotype group. Although 3 different markers were identified among persisters (*bcrABC*, *emrC*, and *qacH*) relative to 2 among non-persisters (*bcrABC* and *emrC*), the overall prevalence was similar (**Figure 2**); 5/7 persisters harboured disinfectant resistance markers (71%), compared with 4/6 non-persisters (67%). This suggests that the presence of known disinfectant resistant markers was not the sole causative mechanism for persistence. It does, however, support the previous associations of Tn6188 with strains associated with food environments, and their propensity to be associated with survival and/or persistent contamination dynamics, since this transposon, including the *qacH* gene, was only identified among persistent strains in this study (Muller et al., 2013, Ortiz et al., 2015, Hurley et al., 2019).

Heavy metal resistance has been frequently observed in *Listeria* species, principally to cadmium and arsenic; the associated genetic resistance markers are among the more commonly found stress resistance markers associated with mobile genetic elements across the genus (Parsons et al., 2018). Cadmium resistance is generally mediated through the *cadAC* cassette system in *Listeria* species (Lebrun et al., 1994a, Lebrun et al., 1994b), with 6 *cadA* variants (*cadA1-A6*) described to date (Chmielowska et al., 2020). Of these variants, *cadA4* is thought to provide the lowest relative tolerance to cadmium, permitting growth up to

approximately 50 µg/ml (Parsons et al., 2017); *cadA1* and *cadA2*, however, facilitate growth at concentrations >140 µg/ml. While a direct link to persistence and cadmium resistance has not been demonstrated, there is growing evidence that the prevalence of cadmium resistance is higher among clones showing recurrent contamination patterns in FPEs compared with their sporadically contaminating counterparts (Harvey and Gilmour, 2001, Parsons et al., 2020). Results of this study suggest that high frequencies of known cadmium resistance cassettes were present among both persistent (86%) and presumed non-persistent (83%) cohorts. Although found at high incidence among strains in this study, results suggest *cadA1* is more common in persisters, while *cadA4*, which provides lower tolerance than *cadA1*, was only carried in non-persisters.

Arsenic resistance is typically associated with higher prevalence among serotype 4b strains of *L. monocytogenes* (McLauchlin et al., 1997, Mullapudi et al., 2008); this study only included a single 4b isolate, UCDL037, harbouring Listeria Genomic island 2 (LGI2), which carries a large arsenic resistance operon (*arsD1A1R1D2R2A2B1B2*). LGI2 was also present in UCDL175, and this operon encodes both the *arsA1* and *arsA2* ATP transporters, as well as the membrane transporters *arsB1* and *arsB2*. Interestingly in this study, lineage II *L. monocytogenes* had a relatively high rate of carriage of arsenic resistance determinants, with 86% (6/7) of these strains encoding an arsenic transporter. This was primarily due to the presence of a Tn544 resistance transposon containing an *arsCDABR* cassette (Kuenne et al., 2013). This prevalence is higher than previously noted by McLauchlin et al. (1997) or Mullapudi et al. (2008), who both reported lineage II resistance rates of 3%. An extended wider analysis of 100 isolates was conducted to further investigate if the higher arsenic resistance prevalence from the facility in this study, was also observed in other facilities in Ireland. A set of 100 isolates previously described by Hurley et al. (2019) was analysed for carriage of arsenic resistance markers (*arsA1*, *arsA2*, and the Tn544 resistance cassette); although the carriage rate among the lineage II isolates was higher than the previously mentioned studies at 14%, it was still lower than the prevalence observed in this study. In both cases, the Tn544 cassette was the most common resistant determinant. Interestingly, Pasquali et al. (2018) noted a high carriage rate of LGI2-associated *ars* operon among ST14

isolates; however, this was absent in ST121 isolates collected from the same environment.

The reason for the higher prevalence of arsenic resistance in the present study is not clear, but may elude to introduction of resistant isolates from ingredient suppliers, and/or horizontal gene transfer (HGT) events at the facility. The high carriage rate of Tn544-mediate resistance among persisters (50%), coupled with carriage across different species, supports the likelihood of HGT dynamics.

A number of broad spectrum stress survival islets of *L. monocytogenes* and/or *L. innocua*, denoted as SSIs (SSI-1 and SSI-2), have been described; these provide benefits to growth and/or survival under sub-optimal or stress conditions, such as low pH (SSI-1), alkaline pH (SSI-2), or oxidative stress conditions (both islets) (Ryan et al., 2010, Harter et al., 2017). Their carriage is typically overrepresented among food isolates, and has been implicated in persistence of clonally-related groups, such as ST121; serotype 4b isolates often lack SSI-1 or SSI-2, but instead harbour a 549 bp hypothetical protein CDS (referred to as the variation 'SSI-V' islet in this study). In line with previous studies, UCDL019 (an ST121 strain) and both *L. innocua* isolates, harboured SSI-2; the other lineage II isolates harboured SSI-1 (UCD011, UCDL016, UCDL133, UCDL150, UCDL175, and UCDL187), while the ST1 isolate UCDL037 contained SSI-V (**Figure 2, Supplementary Figure 1**). Ryan et al. (2010) noted *L. welshimeri* strain SLCC5334 lacked any genes in the SSI insertion hotspot; interestingly, in our study both persistent *L. welsimeri* isolates harboured SSI-1, while the non-persistent isolates had an absence of any insert in the SSI locus. This may allude to the possible contribution of the SSI inserts to persistence of *Listeria* strains in FPEs, and should be further investigated among other persistent and presumed non-persistent clones to provide additional insights.

The internalin family of proteins comprise 25 members with characteristic leucine-rich repeat domains, and have demonstrated roles in virulence and host pathogen interactions (Radoshevich and Cossart, 2017). The most well-characterised of these, InIA, mediates entry to host cells through binding of the E-cadherin host cell receptor (Gaillard et al., 1991, Mengaud et al., 1996). A number of mutations have been reported in the coding gene, *inIA*, which lead to production of truncated InIA variants (Van Stelten et al., 2010). These variants

typically lack the LPXGT motif at the C-terminal end, and are not bound to the bacterial cell wall by the sortase enzyme. Associated strains of *L. monocytogenes* lacking *inlA* are generally attenuated in their pathogenicity (Olier et al., 2005). In this study, of the 8 *L. monocytogenes* isolates (each harbouring *inlA*), 5 contained mutations leading to a premature stop codon (PMSC) in the gene sequence, which would produce a truncated InlA lacking the LPXTG sequence motif. This included all the persistent strains (4/4; 100%), and 1 of the presumed non-persistent strains (1/4; 25%). This suggests that all the persistent isolates included in this study would be associated with reduced virulence *in vivo*. Persistent isolate UCDL185 also contained a single nucleotide insertion in a polyA region at the N-terminal end of *inlC* (nucleotide positions 7-15), causing a frameshift mutation and leading to a downstream PMSC (Supplementary Figure 2). This mutation may also have a negative impact on the virulence of this strain.

The *ascB-dapE* internalin cluster includes variable combinations of *inlC2*, *inlD*, *inlE*, *inlG*, and *inlH*, and has been suggested as of potential use as a marker for sublineage classification (Chen et al., 2012). Sublineages IIA, IIB and IIC were noted in this study. However, strain UCDL019 included another sublineage II variant, similar to that described by Dramsi et al. (1997). Another interesting feature of this locus was noted in UCDL187, where the flanking *ascB* β -glucosidase gene, as well as one of the hypothetical proteins in the locus, were identified as pseudogenes (Supplementary Figure 3).

The expression of most key *L. monocytogenes* virulence factors identified to date is under the control of *prfA*, the main virulence regulator; this regulator is responsible for the switch to *in vivo* pathogenesis, when the bacterium enters its mammalian host (Chakraborty et al., 1992, Freitag et al., 2009). One persistent isolate in this study, UCDL185, harboured a 7 nucleotide insertion in *prfA*, causing a downstream PMSC at amino acid position 185 (A185*). This mutation has been associated with attenuated virulence *in vivo* (Roche et al., 2005, López et al., 2013).

Apart for LIPI-1 and LIPI-2 (the former encoding the main virulence gene locus in *L. monocytogenes*, and the latter encoding virulence factors in *L. ivanovii*), two additional

pathogenicity island of note have been described: LIPI-3 and LIPI-4. The LIPI-3 pathogenicity island encodes listeriolysin S, which is associated with increased strain virulence. This has been associated with functionality as a bacteriocin when expressed in the intestinal microenvironment, positively contributing to strains' capacity to colonise this niche, and with a role as an alternative haemolysin/cytolysin (Cotter et al., 2008, Quereda et al., 2017). In this study, only a single isolate harboured LIPI-3: UC DL037, a presumed non-persistent strain. The LIPI-4 pathogenicity island is associated with hypervirulence in a subset of *L. monocytogenes* genetic clones, encoding a putative phosphotransferase system (Maury et al., 2016). No isolates in this study harboured LIPI-4. Overall, these results suggest that these additional pathogenicity islands are not common among food isolates, and does not correlate persistence to increased/hypervirulence.

Taken together, our results suggest a lower virulence potential of persistent isolates in this study, due to the widespread prevalence of truncated InIA among the persistent *L. monocytogenes* strains, the lack of additional virulence factors such as LIPI-3 and LIPI-4, and the other notable mutations such as that of *prfA* in persistent strain UC DL187. Since persistence of pathogenic bacteria in FPEs can be associated with an increased risk to public health, due to an ongoing risk of cross-contamination of food products associated with the colonised environment, the attenuated virulence observed among persistent isolates in this study is positive from a food safety perspective. These results also suggest in the facility studied, persistent strains were likely to be less virulent than other transient strains found in the same environment. None of the non-*monocytogenes* species in this study contained homologs of any of the virulence genes shown in **Figure 3**.

The presence of mobile genetic elements typically gives rise to diverse functional variation in the *L. monocytogenes* accessory genome, although species-specific differences have also been noted (Glaser et al., 2001, Hain et al., 2006, Fox et al., 2016). Plasmids were found to contribute to variation across strains of the same species, as well as inter-species, with at least one plasmid present in 10 of the isolates included in this study (77%); of these, three isolates contained 2 plasmids (**Supplementary Figure 4**). These plasmids

encoded a number of genetic markers related to stress resistance, as illustrated in **Figure 4**. This included determinants related to disinfectant resistance (*bcrABC* and *emrC*), heavy metal resistance systems (including cadmium and copper resistance), and other stress resistance markers with roles in oxidative and temperature stress, such as *clpB*, *clpL* and NADH peroxidase. Interestingly, homologs of the same plasmid were found in multiple strains: for example, both *L. welshimeri* UC DL063 and UC DL122 strains contained identical plasmids (pUC DL063-1 and pUC DL122-1, **Supplementary Figure 4**); similarly, two ST9 strains (UC DL016 and UC DL133) contained similar plasmids (pUC DL016-1 and pUC DL133-1), and a smaller plasmid previously described in ST6 strains (Kropac et al., 2019), encoding *emrC*, was present in three isolates in this study (UC DL011, UC DL016, and UC DL133). Interestingly, all strains harbouring this small 4,265 bp plasmid also harboured larger plasmids, but these larger plasmids did not encode *bcrABC*, nor did these strains carry the *qacH*-containing transposon Tn6188. The carriage of extrachromosomal plasmid DNA confers an associated fitness cost to strains; whether the presence of either disinfectant resistance plasmid leads to exclusion or unstable carriage of the other, requires further investigation.

The *comK* gene is a known phage insertion hotspot in *L. monocytogenes*, and may contain the variant of A118, or other, phage (Loessner et al., 2000, Orsi et al., 2008, Fox et al., 2016). Interestingly, the presence of a *comK* phage insertion ($\phi comK$), has been suggested to play a role in colonisation and persistence by functioning as a rapid adaptation island through recombination events (Verghese et al., 2011). Analysis of isolates in this study found $\phi comK$ variants among both persistent (57%) and presumed non-persistent isolates (50%), suggesting the presence of an insert does not predispose a strain to persistence. The $\phi comK$ genotypes also suggested multiple recombination events had occurred, and no clear genotype responsible for persistence was apparent (**Supplementary Figure 5**). Interestingly among the isolates in this study, no *L. welshimeri* had a $\phi comK$ insert. To further investigate this, we examined the phage *attP* attachment site and the corresponding *attB* bacterial site. The *comK* phage *attP* site is unusual in that the core insert sequence is just 3 nucleotides in length (-GGA-). When comparing the insert site sequence

Accepted Article
across isolates in this study, all three *L. welshimeri* isolates contained a SNP in their *attB* (GGT, **Figure 5**). To further elaborate this, we compared this region in four additional *L. welshimeri* strains (**Supplementary Figure 8**); all of these also harboured the 'GGT' variant. Taken together, these results may indicate that the *comK* gene in *L. welshimeri* may not be an insertion hotspot due to *attP/attB* sequence variation. This could be further investigated as more *L. welshimeri* genomes become available.

To investigate if similar genotypes were shared by persistent isolates, we conducted a pangenome analysis and constructed an associated maximum likelihood phylogenetic analysis of the strains in this study (**Supplementary Figure 6**). No clear segregation was noted based on pangenome genotype; similarly, considering plasmid carriage and ϕ *comK* status, no clear clustering was observed. In addition, further analysis of the CC9 sub-clade was undertaken, comparing core SNPs across the 4 associated strains. This sub-clade included both persistent and presumed non-persistent isolates, and the SNP differences supported a diverse strain cohort, with no clear segregation between persistent and presumed non-persistent strains, again reinforcing the observations of the pangenome phylogeny.

This study sought to further investigate persistence of *Listeria* species in FPEs by comparing cohorts of persistent and presumed non-persistent isolates collected from the same environment. This facilitated evaluation of related molecular mechanisms, in the context of an environment exerting similar selective pressures on associated strains of *Listeria*. Taken together, the insights provided in this study do not point to a single genetic mechanism driving the persistence of *Listeria* strains in the FPE of their isolation. Persistent strains of *L. monocytogenes* were more likely to harbour mutations associated with hypovirulence and less invasive disease. While disinfectant resistance markers were found in both persistent and presumed non-persistent strains, *qacH* was only identified among the persistent cohort. Persistent *L. welshimeri* strains harboured SSI-1, and strains of this species may be less prone to *comK* phage insertion due to *attB* site mutation.

ACKNOWLEDGEMENTS

The authors wish to sincerely thank the food business involved in this study. We thank the Institut Pasteur teams for the curation and maintenance of BIGSdb-Pasteur databases at <http://bigsdbs.pasteur.fr/>.

CONFLICT OF INTEREST

No conflict of interest declared.

REFERENCES

- Chmielowska, C., Korsak, D., Szymulowska, B., Krop, A., Lipka, K., Krupińska, M. & Bartosik, D. (2020) Genetic carriers and genomic distribution of cadA6-a novel variant of a cadmium resistance determinant identified in *Listeria* spp. *Int J Mol Sci*, **21**.
- Quereda, J. J., Leclercq, A., Moura, A., Vales, G., Gómez-Martín, Á., García-Muñoz, Á., Thouvenot, P., Tessaud-Rita, N., Bracq-Dieye, H. & Lecuit, M. (2020) *Listeria valentina* sp. nov., isolated from a water trough and the faeces of healthy sheep. *Int J Syst Evol Microbiol*, **70**, 5868-5879.
- Parsons, C., Lee, S. & Kathariou, S. (2020) Dissemination and conservation of cadmium and arsenic resistance determinants in *Listeria* and other Gram-positive bacteria. *Mol Microbiol*, **113**, 560-569.
- Leclercq, A., Moura, A., Vales, G., Tessaud-Rita, N., Aguilhon, C. & Lecuit, M. (2019) *Listeria thailandensis* sp. nov. *Int J Syst Evol Microbiol*, **69**, 74-81.
- Hurley, D., Luque-Sastre, L., Parker, C. T., Huynh, S., Eshwar, A. K., Nguyen, S. V., Andrews, N., Moura, A., Fox, E. M., Jordan, K., Lehner, A., Stephan, R. & Fanning, S. (2019) Whole-genome sequencing-based characterization of 100 *Listeria monocytogenes* isolates collected from food processing environments over a four-year period. *mSphere*, **4**, e00252-19.
- Kropac, A. C., Eshwar, A. K., Stephan, R. & Tasara, T. (2019) New insights on the role of the pLMST6 plasmid in *Listeria monocytogenes* biocide tolerance and virulence. *Frontiers in Microbiology*, **10**.

- Rodríguez-Campos, D., Rodríguez-Melcón, C., Alonso-Calleja, C. & Capita, R. (2019) Persistent *Listeria monocytogenes* isolates from a poultry-processing facility form more biofilm but do not have a greater resistance to disinfectants than sporadic strains. *Pathogens*, **8**, 13.
- Pasquali, F., Palma, F., Guillier, L., Lucchi, A., De Cesare, A. & Manfreda, G. (2018) *Listeria monocytogenes* sequence types 121 and 14 repeatedly isolated within one year of sampling in a rabbit meat processing plant: persistence and ecophysiology. *Frontiers in Microbiology*, **9**.
- Parsons, C., Lee, S. & Kathariou, S. (2018) Heavy metal resistance determinants of the foodborne pathogen *Listeria monocytogenes*. *Genes*, **10**, 11.
- Radoshevich, L. & Cossart, P. (2017) *Listeria monocytogenes*: towards a complete picture of its physiology and pathogenesis. *Nat Rev Microbiol*.
- Jennison, A. V., Masson, J. J., Fang, N.-X., Graham, R. M., Bradbury, M. I., Fegan, N., Gobius, K. S., Graham, T. M., Guglielmino, C. J., Brown, J. L. & Fox, E. M. (2017) Analysis of the *Listeria monocytogenes* Population Structure among Isolates from 1931 to 2015 in Australia. *Frontiers in Microbiology*, **8**.
- Harter, E., Wagner, E. M., Zaiser, A., Halecker, S., Wagner, M. & Rychli, K. (2017) Stress Survival Islet 2, predominantly present in *Listeria monocytogenes* strains of sequence type 121, is involved in the alkaline and oxidative stress responses. *Appl Environ Microbiol*, **83**.
- Parsons, C., Lee, S., Jayeola, V. & Kathariou, S. (2017) Novel cadmium resistance determinant in *Listeria monocytogenes*. *Appl Environ Microbiol*, **83**.
- Quereda, J. J., Meza-Torres, J., Cossart, P. & Pizarro-Cerdá, J. (2017) Listeriolysin S: A bacteriocin from epidemic *Listeria monocytogenes* strains that targets the gut microbiota. *Gut Microbes*, **8**, 384-391.
- Orsi, R. H. & Wiedmann, M. (2016) Characteristics and distribution of *Listeria* spp., including *Listeria* species newly described since 2009. *Appl Microbiol Biotechnol*, **100**, 5273-5287.
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y. & Wishart, D. S. (2016) PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, **44**, W16-W21.
- Henri, C., Félix, B., Guillier, L., Leekitcharoenphon, P., Michelon, D., Mariet, J.-F., Aarestrup, F. M., Mistou, M.-Y., Hendriksen, R. S. & Roussel, S. (2016) Population genetic structure of *Listeria monocytogenes* strains as determined by pulsed-field gel electrophoresis and multilocus sequence typing. *Appl Environ Microbiol*, **82**, 5720-5728.

- Accepted Article
- Maury, M. M., Tsai, Y.-H., Charlier, C., Touchon, M., Chenal-Francisque, V., Leclercq, A., Criscuolo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E. P. C., Brisse, S. & Lecuit, M. (2016) Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. *Nat Genet*, **48**, 308-313.
- Fox, E. M., Allnutt, T., Bradbury, M. I., Fanning, S. S. & Chandry, P. S. (2016) Comparative genomics of the *Listeria monocytogenes* ST204 subgroup. *Frontiers in Microbiology*, **7**.
- Fox, E. M., Wall, P. G. & Fanning, S. (2015) Control of *Listeria* species food safety at a poultry food production facility. *Food Microbiol*, **51**, 81-86.
- Schmitz-Esser, S., Müller, A., Stessl, B. & Wagner, M. (2015) Genomes of sequence type 121 *Listeria monocytogenes* strains harbor highly conserved plasmids and prophages. *Frontiers in Microbiology*, **6**.
- Ortiz, S., López-Alonso, V., Rodríguez, P. & Martínez-Suárez, J. V. (2015) The Connection between persistent, disinfectant-resistant *Listeria monocytogenes* strains from two geographically separate Iberian pork processing plants: evidence from comparative genome analysis. *Appl Environ Microbiol*, **82**, 308-317.
- Ebner, R., Stephan, R., Althaus, D., Brisse, S., Maury, M. & Tasara, T. (2015) Phenotypic and genotypic characteristics of *Listeria monocytogenes* strains isolated during 2011–2014 from different food matrices in Switzerland. *Food Control*, **57**, 321-326.
- Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. G., Fookes, M., Falush, D., Keane, J. A. & Parkhill, J. (2015) Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, **31**, 3691-3693.
- Seemann, T. 2015. snippy: fast bacterial variant calling from NGS reads. <https://github.com/tseemann/snippy>.
- Mcauley, C. M., Mcmillan, K., Moore, S. C., Fegan, N. & Fox, E. M. (2014) Prevalence and characterization of foodborne pathogens from Australian dairy farm environments. *J Dairy Sci*, **97**, 7402-12.
- Fox, E. M., Solomon, K., Moore, J. E., Wall, P. G. & Fanning, S. (2014) Phylogenetic profiles of in-house microflora in drains at a food production facility: comparison and biocontrol implications of *Listeria*-positive and -negative bacterial populations. *Appl Environ Microbiol*, **80**, 3369-74.
- Bolger, A. M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114-2120.
- Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England)*, **30**, 1312-1313.

Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., Møller Aarestrup, F. & Hasman, H. (2014) In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial agents and chemotherapy*, **58**, 3895-3903.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, **30**, 2068-9.

Kuenne, C., Billion, A., Mraheil, M. A., Strittmatter, A., Daniel, R., Goesmann, A., Barbuddhe, S., Hain, T. & Chakraborty, T. (2013) Reassessment of the *Listeria monocytogenes* pan-genome reveals dynamic integration hotspots and mobile genetic elements as major components of the accessory genome. *BMC Genomics*, **14**, 47.

Pulsenet. 2013. *Standard operating procedure for PulseNet PFGE of Listeria monocytogenes* [Online].

Available: <http://www.cdc.gov/pulsenet/PDF/listeria-pfge-protocol-508c.pdf> [Accessed].

Muller, A., Rychli, K., Muhterem-Uyar, M., Zaiser, A., Stessl, B., Guinane, C. M., Cotter, P. D., Wagner, M. & Schmitz-Esser, S. (2013) Tn6188 - a novel transposon in *Listeria monocytogenes* responsible for tolerance to benzalkonium chloride. *PLoS One*, **8**, e76835.

López, V., Navas, J. & Martínez-Suárez, J. V. (2013) Low potential virulence associated with mutations in the *inlA* and *prfA* genes in *Listeria monocytogenes* isolated from raw retail poultry meat. *J Food Prot*, **76**, 129-32.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A. & Pevzner, P. A. (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, **19**, 455-477.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647-1649.

Chen, J., Fang, C., Zhu, N., Lv, Y., Cheng, C., Bei, Y., Zheng, T. & Fang, W. (2012) Genetic organization of *ascB-dapE* internalin cluster serves as a potential marker for *Listeria monocytogenes* sublineages IIA, IIB, and IIC. *J Microbiol Biotechnol*, **22**, 575-84.

Fox, E. M., Leonard, N. & Jordan, K. (2011a) Physiological and transcriptional characterization of persistent and nonpersistent *Listeria monocytogenes* isolates. *Appl Environ Microbiol*, **77**, 6559-69.

- Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., Jones, J. L. & Griffin, P. M. (2011) Foodborne Illness Acquired in the United States—Major Pathogens. *Emerging Infect Dis*, **17**, 7-15.
- Verghese, B., Lok, M., Wen, J., Alessandria, V., Chen, Y., Kathariou, S. & Knabel, S. (2011) *comK* prophage junction fragments as markers for *Listeria monocytogenes* genotypes unique to individual meat and poultry processing plants and a model for rapid niche-specific adaptation, biofilm formation, and persistence. *Appl Environ Microbiol*, **77**, 3279-92.
- Sullivan, M. J., Petty, N. K. & Beatson, S. A. (2011) Easyfig: a genome comparison visualizer. *Bioinformatics*, **27**, 1009-1010.
- Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. (2011) BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, **12**, 1-10.
- Fox, E., Hunt, K., O'brien, M. & Jordan, K. (2011b) *Listeria monocytogenes* in Irish Farmhouse cheese processing environments. *Int J Food Microbiol*, **145 Suppl 1**, S39-45.
- Van Stelten, A., Simpson, J. M., Ward, T. J. & Nightingale, K. 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. *Appl Environ Microbiol*, **76**, 2783-90.
- Ryan, S., Begley, M., Hill, C. & Gahan, C. G. (2010) A five-gene stress survival islet (SSI-1) that contributes to the growth of *Listeria monocytogenes* in suboptimal conditions. *J Appl Microbiol*, **109**, 984-95.
- Price, M. N., Dehal, P. S. & Arkin, A. P. (2009) FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Mol Biol Evol*, **26**, 1641-1650.
- Lomonaco, S., Decastelli, L., Nucera, D., Gallina, S., Manila Bianchi, D. & Civera, T. (2009) *Listeria monocytogenes* in Gorgonzola: subtypes, diversity and persistence over time. *Int J Food Microbiol*, **128**, 516-20.
- Kastbjerg, V. G. & Gram, L. (2009) Model systems allowing quantification of sensitivity to disinfectants and comparison of disinfectant susceptibility of persistent and presumed nonpersistent *Listeria monocytogenes*. *J Appl Microbiol*, **106**, 1667-81.
- Freitag, N. E., Port, G. C. & Miner, M. D. (2009) *Listeria monocytogenes* - from saprophyte to intracellular pathogen. *Nat Rev Microbiol*, **7**, 623-8.

- Ragon, M., Wirth, T., Hollandt, F., Lavenir, R., Lecuit, M., Le Monnier, A. & Brisse, S. (2008) A new perspective on *Listeria monocytogenes* evolution. *PLoS Pathog*, **4**, e1000146.
- Aziz, R. K., Bartels, D., Best, A. A., Dejongh, M., Disz, T., Edwards, R. A., Formis, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A. & Zagnitko, O. (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, **9**, 75-75.
- Mullapudi, S., Siletzky, R. M. & Kathariou, S. (2008) Heavy-metal and benzalkonium chloride resistance of *Listeria monocytogenes* isolates from the environment of turkey-processing plants. *Appl Environ Microbiol*, **74**, 1464-1468.
- Cotter, P. D., Draper, L. A., Lawton, E. M., Daly, K. M., Groeger, D. S., Casey, P. G., Ross, R. P. & Hill, C. (2008) Listeriolysin S, a novel peptide haemolysin associated with a subset of lineage I *Listeria monocytogenes*. *PLoS Pathog*, **4**, e1000144.
- Orsi, R. H., Borowsky, M. L., Lauer, P., Young, S. K., Nusbaum, C., Galagan, J. E., Birren, B. W., Ivy, R. A., Sun, Q., Graves, L. M., Swaminathan, B. & Wiedmann, M. (2008) Short-term genome evolution of *Listeria monocytogenes* in a non-controlled environment. *BMC Genomics*, **9**, 539.
- Chambel, L., Sol, M., Fernandes, I., Barbosa, M., Zilhao, I., Barata, B., Jordan, S., Perni, S., Shama, G., Adriaio, A., Faleiro, L., Requena, T., Pelaez, C., Andrew, P. W. & Tenreiro, R. (2007) Occurrence and persistence of *Listeria* spp. in the environment of ewe and cow's milk cheese dairies in Portugal unveiled by an integrated analysis of identification, typing and spatial-temporal mapping along production cycle. *Int J Food Microbiol*, **116**, 52-63.
- 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. *J Clin Microbiol*, **45**, 865-73.
- Wulff, G., Gram, L., Ahrens, P. & Vogel, B. F. (2006) One group of genetically similar *Listeria monocytogenes* strains frequently dominates and persists in several fish slaughter- and smokehouses. *Appl Environ Microbiol*, **72**, 4313-22.
- Hain, T., Steinweg, C. & Chakraborty, T. (2006) Comparative and functional genomics of *Listeria* spp. *J Biotechnol*, **126**, 37-51.

- Olier, M., Garmyn, D., Rousseaux, S., Lemaître, J.-P., Piveteau, P. & Guzzo, J. (2005) Truncated internalin A and asymptomatic *Listeria monocytogenes* carriage: in vivo investigation by allelic exchange. *Infection and immunity*, **73**, 644-648.
- Roche, S. M., Gracieux, P., Milohanic, E., Albert, I., Virlogeux-Payant, I., Témoin, S., Grépinet, O., Kerouanton, A., Jacquet, C., Cossart, P. & Velge, P. (2005) Investigation of specific substitutions in virulence genes characterizing phenotypic groups of low-virulence field strains of *Listeria monocytogenes*. *Appl Environ Microbiol*, **71**, 6039-6048.
- Doumith, M., Buchrieser, C., Glaser, P., Jacquet, C. & Martin, P. (2004) Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. *J Clin Microbiol*, **42**, 3819-22.
- Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 3059-3066.
- Djordjevic, D., Wiedmann, M. & Mclandsborough, L. A. (2002) Microtiter plate assay for assessment of *Listeria monocytogenes* biofilm formation. *Appl Environ Microbiol*, **68**, 2950-2958.
- Harvey, J. & Gilmour, A. (2001) Characterization of recurrent and sporadic *Listeria monocytogenes* isolates from raw milk and nondairy foods by pulsed-field gel electrophoresis, monocin typing, plasmid profiling, and cadmium and antibiotic resistance determination. *Appl Environ Microbiol*, **67**, 840-7.
- Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P., Chakraborty, T., Charbit, A., Chetouani, F., Couve, E., De Daruvar, A., Dehoux, P., Domann, E., Dominguez-Bernal, G., Duchaud, E., Durant, L., Dussurget, O., Entian, K. D., Fsihi, H., Garcia-Del Portillo, F., Garrido, P., Gautier, L., Goebel, W., Gomez-Lopez, N., Hain, T., Hauf, J., Jackson, D., Jones, L. M., Kaerst, U., Kreft, J., Kuhn, M., Kunst, F., Kurapkat, G., Madueno, E., Maitournam, A., Vicente, J. M., Ng, E., Nedjari, H., Nordsiek, G., Novella, S., De Pablos, B., Perez-Diaz, J. C., Purcell, R., Rimmel, B., Rose, M., Schlueter, T., Simoes, N., Tierrez, A., Vazquez-Boland, J. A., Voss, H., Wehland, J. & Cossart, P. (2001) Comparative genomics of *Listeria* species. *Science*, **294**, 849-52.
- Loessner, M. J., Inman, R. B., Lauer, P. & Calendar, R. (2000) Complete nucleotide sequence, molecular analysis and genome structure of bacteriophage A118 of *Listeria monocytogenes*: implications for phage evolution. *Mol Microbiol*, **35**, 324-340.
- Norwood, D. E. & Gilmour, A. (1999) Adherence of *Listeria monocytogenes* strains to stainless steel coupons. *J Appl Microbiol*, **86**, 576-82.

Mclauchlin, J., Hampton, M. D., Shah, S., Threlfall, E. J., Wieneke, A. A. & Curtis, G. D. W. (1997) Subtyping of *Listeria monocytogenes* on the basis of plasmid profiles and arsenic and cadmium susceptibility. *J Appl Microbiol*, **83**, 381-388.

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

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

Lebrun, M., Audurier, A. & Cossart, P. (1994a) Plasmid-borne cadmium resistance genes in *Listeria monocytogenes* are present on Tn5422, a novel transposon closely related to Tn917. *J Bacteriol*, **176**, 3049-3061.

Lebrun, M., Audurier, A. & Cossart, P. (1994b) Plasmid-borne cadmium resistance genes in *Listeria monocytogenes* are similar to *cadA* and *cadC* of *Staphylococcus aureus* and are induced by cadmium. *J Bacteriol*, **176**, 3040-3048.

Chakraborty, T., Leimeister-Wächter, M., Domann, E., Hartl, M., Goebel, W., Nichterlein, T. & Notermans, S. (1992) Coordinate regulation of virulence genes in *Listeria monocytogenes* requires the product of the *prfA* gene. *J Bacteriol*, **174**, 568.

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

TABLE 1 Summary of the genetic subtypes and genomes of strains in this study.

Strain and Molecular Subtype						Genome		
Isolate	Species	Sequence	Clonal	Serotype	Lineage	Size (bp)	GC%	CDS
		Type (ST)	Complex (CC)					
UCDL011	<i>L. monocytogenes</i>	ST9	CC9	1/2c, 3c	II	3,073,215	37.8	3,058
UCDL016	<i>L. monocytogenes</i>	ST9	CC9	1/2c, 3c	II	3,063,336	37.8	3,035
UCDL019	<i>L. monocytogenes</i>	ST121	CC121	1/2a, 3a	II	3,057,649	37.8	3,018
UCDL187	<i>L. monocytogenes</i>	ST31	CC31	1/2a, 3a	II	3,080,560	37.8	3,094
UCDL122	<i>L. welshimeri</i>	ST168	CC168	ND	-	2,856,944	36.3	2,787
UCDL162	<i>L. welshimeri</i>	ST2688	-	ND	-	2,832,429	36.1	2,817
UCDL146	<i>L. innocua</i>	ST602	-	ND	-	3,026,780	37.3	3,054
UCDL037	<i>L. monocytogenes</i>	ST1	CC1	4b, 4e, 4d	I	2,913,758	37.9	2,836
UCDL133	<i>L. monocytogenes</i>	ST9	CC9	1/2c, 3c	II	3,009,680	37.8	2,977
UCDL150	<i>L. monocytogenes</i>	ST622	CC9	1/2c, 3c	II	3,030,065	37.9	3,007
UCDL175	<i>L. monocytogenes</i>	ST204	CC204	1/2a, 3a	II	3,020,537	37.8	2,951
UCDL063	<i>L. welshimeri</i>	ST1084	CC1084	ND	-	2,946,539	36.3	2,999
UCDL085	<i>L. innocua</i>	ST1008	-	ND	-	2,991,782	37.3	3,002

ND, Not determined. -, Not defined.

Accepted Article

FIGURE 1 Pangenome analysis of the isolates in this study, grouped by species. Numbers represent gene coding loci associated with one or more species.

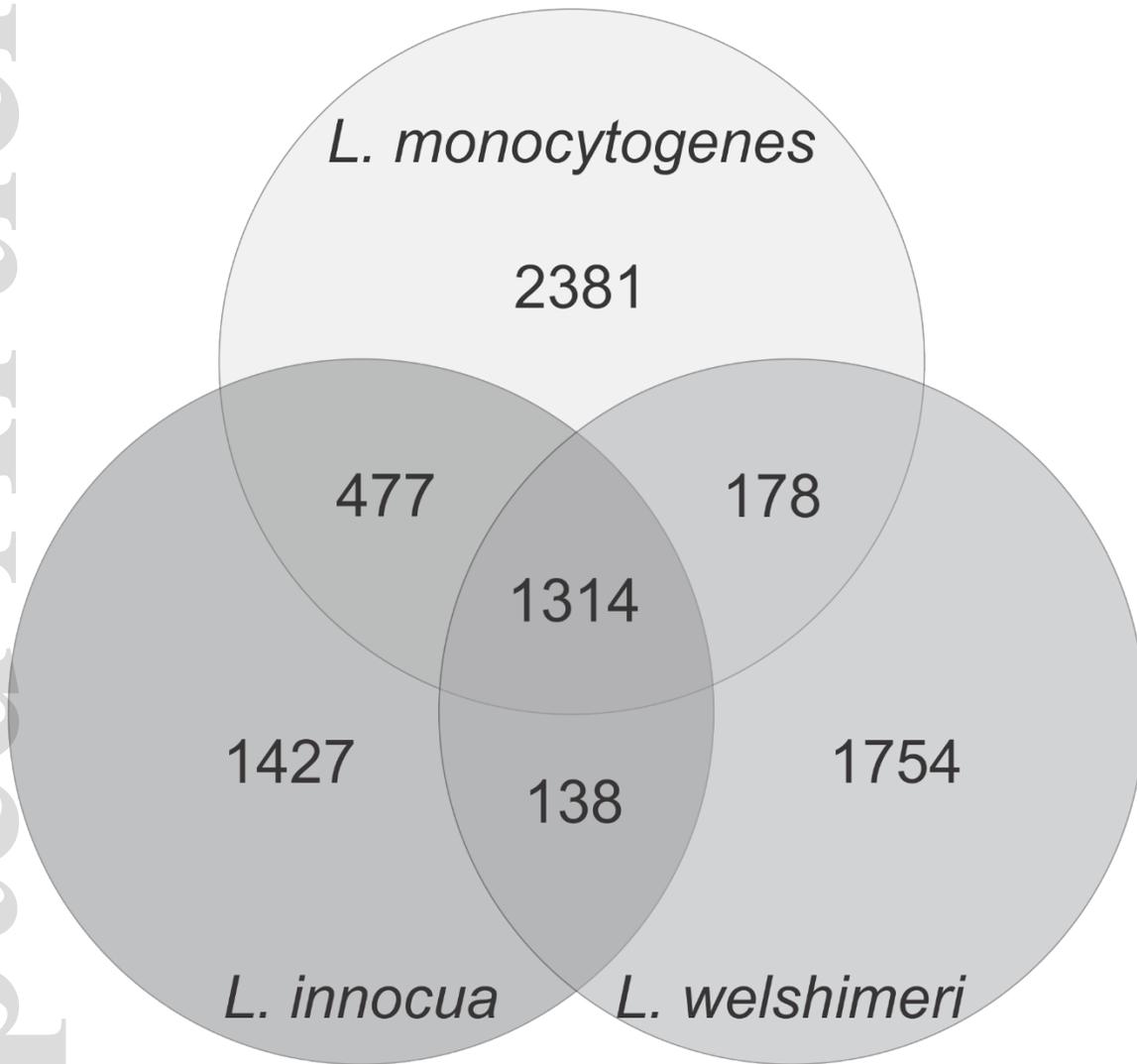


FIGURE 2 Presence of selected stress resistance markers among isolates in this study.

Black boxes indicate presence of gene, white indicate absence. Persistent strains above the grey line, presumed non-persistent strains below.

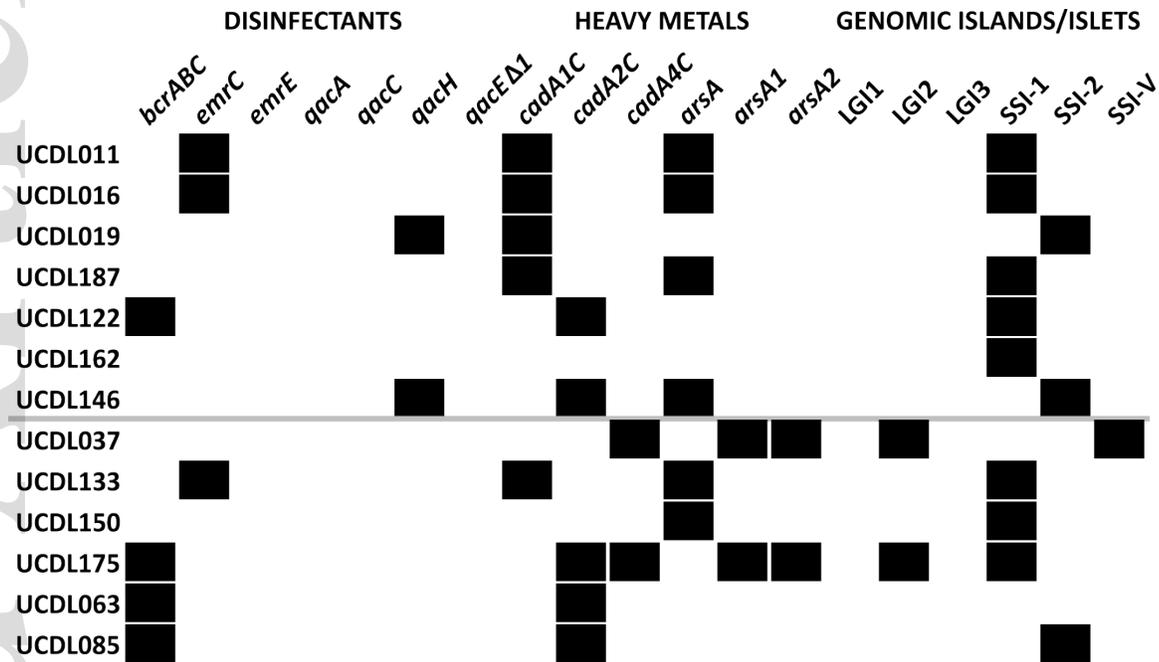


FIGURE 3 Presence of selected stress virulence markers among *L. monocytogenes* isolates in this study. Black boxes indicate presence of gene/genetic island, grey indicate the presence of premature stop codon in the gene sequence (leading to a predicted truncated protein product), and a white box indicates absence of a gene/genetic island. Persistent strains above the grey line, presumed non-persistent strains below.

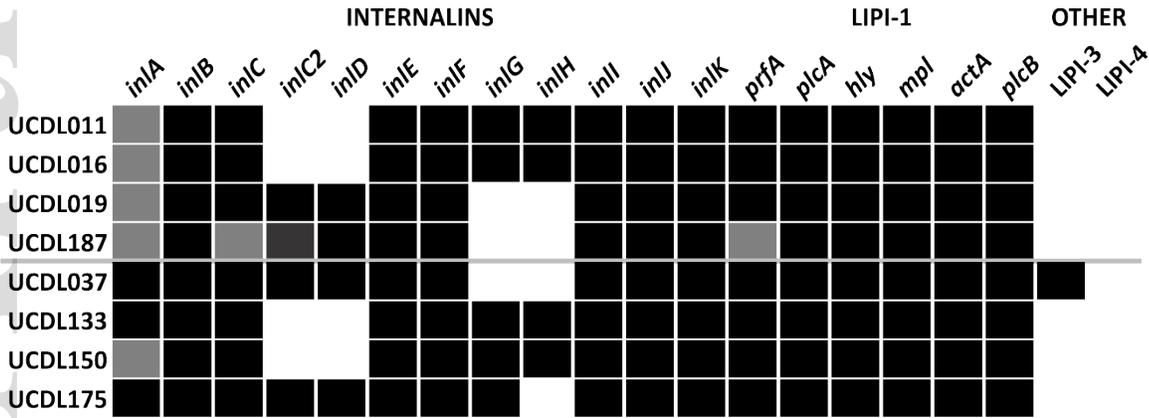


FIGURE 4 BRIG comparative analysis of selected plasmid-borne genes of the *Listeria* isolates. Each ring represents a different plasmid. The black circle in the middle separates the persistent (innermost circles) from the presumed non-persistent (outermost circles) isolates. Numbered genes: 7, 8, 10, heavy metal transporting ATPase; 9, cadmium resistance protein; 11, 12, multicopper oxidase; 13, 14, copper transport negative regulator; 15, multi-antimicrobial extrusion family transporter; 16, NADH peroxidase; 17, NADH quinone oxidoreductase; 18, Na(+)/H(+) antiporter. Gene colours: orange, disinfectant resistance; purple, heavy metal resistance; red, antimicrobial resistance; green, other stress resistance.

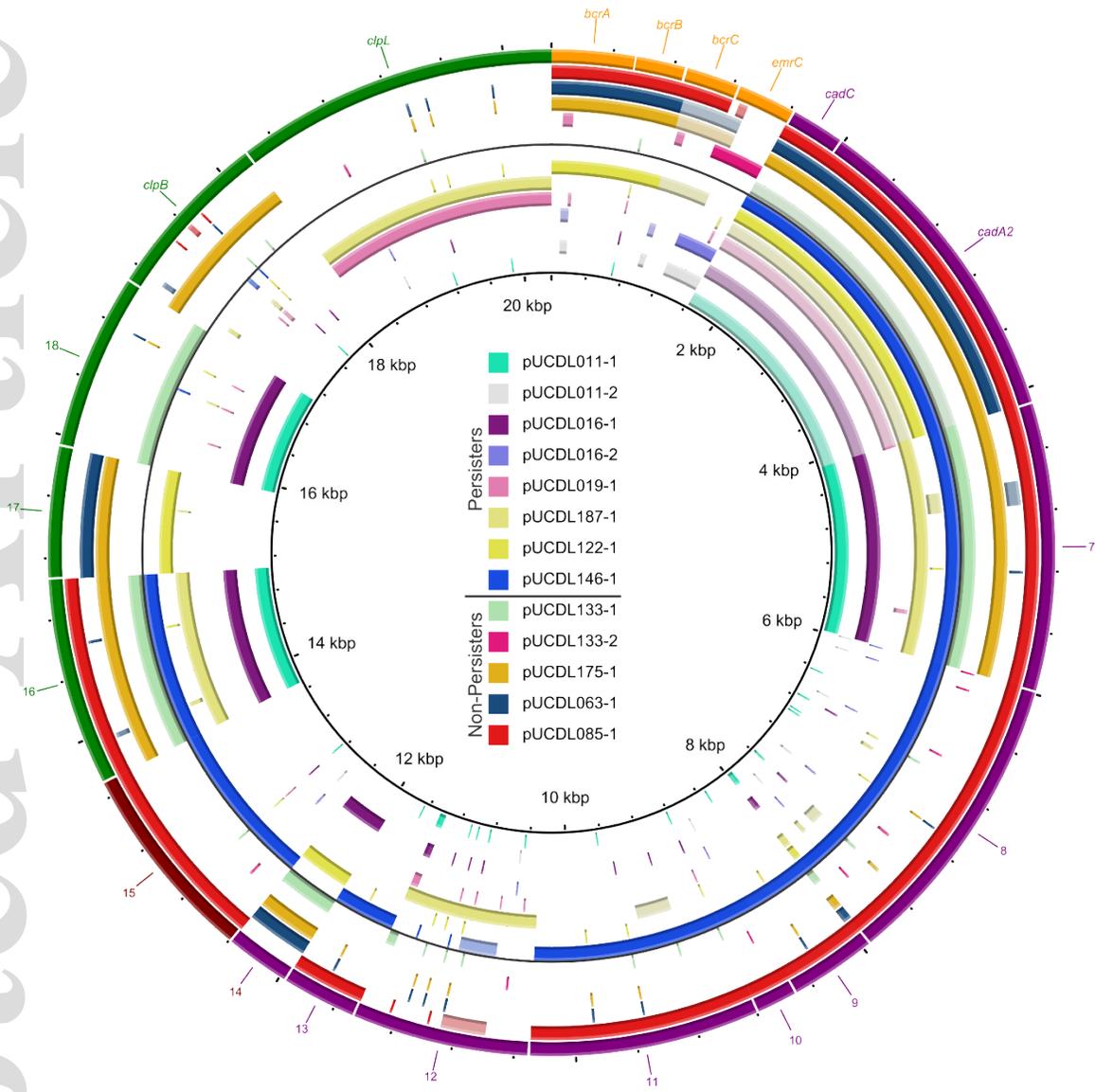
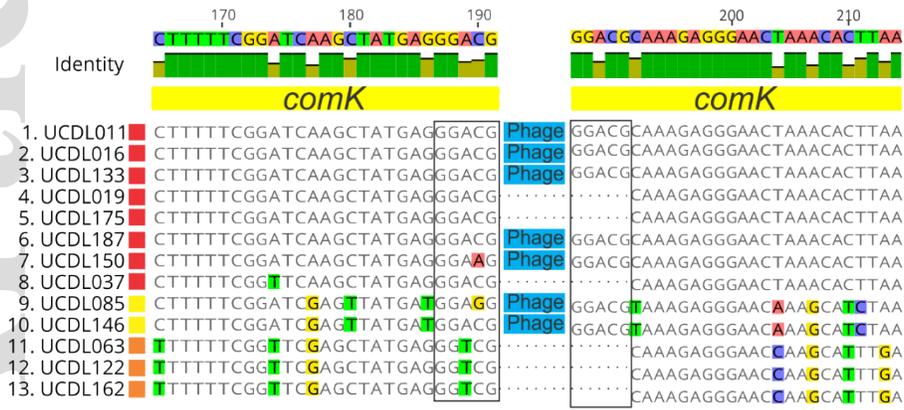


FIGURE 5 Comparative analysis of the *comK* gene and associated phage insert site among isolates in this study. The *attP/attB* site is indicated by the boxes. Species: Red square, *L. monocytogenes*; yellow square, *L. innocua*; orange square, *L. welshimeri*. Number track indicates nucleotide position in wild type *comK* gene sequence.



SUPPLEMENTARY TABLE 1 List of genes used for BLAST analysis, and associated NCBI accession numbers.

SUPPLEMENTARY FIGURE 1 Stress Survival Islet insert site and associated genes among strains in this study. The 'SSI-1' genotype was found in strains UC DL011, UC DL016, UC DL133, UC DL150, UC DL175, UC DL187. The 'SSI-2' genotype was present in strains UC DL019, UC DL085 and UC DL146.

SUPPLEMENTARY FIGURE 2 Comparative analysis of the *inlC* gene showing the wild type sequence (UC DL019), and the UC DL187 sequence harbouring a single nucleotide insertion in the poly(A) tract from nucleotide positions 7-15. This insert results in a premature stop codon at amino acid position 6.

SUPPLEMENTARY FIGURE 3 Comparative analysis of the *ascB-dapE* internalin cluster among strains in this study. Pseudogenes present in UC DL187 are indicated by a dotted line arrow.

SUPPLEMENTARY FIGURE 4 BLAST Ring comparison of plasmids identified among isolates in this study, showing sequence homology between different plasmids. Each ring represents an individual plasmid, as per the Figure legend. The inner black circle represents a constructed plasmid pangenome, incorporating all plasmid sequences of strains in this study.

SUPPLEMENTARY FIGURE 5 Comparative analysis of the *comK* gene and associated ϕ *comK* phage inserts. The *comK* gene is coloured black, and the integrase gene is blue.

SUPPLEMENTARY FIGURE 6 Maximum likelihood analysis of strains in this study, based on a comparative pangenome analysis. From inside to out: the inner ring shows species, then persistent or presumed non-persistent status, then the presence or absence of a ϕ *comK* phage insert, with the outermost ring indicating the presence or absence of a plasmid.

SUPPLEMENTARY FIGURE 7 Core SNP analysis of the CC9 strains in this study. A, a maximum-likelihood phylogenetic tree generated from alignments of core genome nucleotide SNPs. B, number of SNPs between different CC9 strains in this study.

SUPPLEMENTARY FIGURE 8 Alignment of the *comK* gene sequence from *L. welshimeri* strains in this study, as well as an additional 4 strains taken from the NCBI genome database (strains CDPHFDLB, F4083, NCTC 11857, and SLCC5334).