DIVERSE MOBILE GENETIC ELEMENTS SUPPORT THE PERSISTENCE OF *LISTERIA MONOCYTOGENES* ON DAIRY FARMS

Hanna Castro[1], Francois Douillard[1], Hannu Korkeala[1] and Miia Lindström[1]

[1]Department of Food Hygiene and Environmental Health, Faculty of Veterinary Medicine, University of Helsinki

Corresponding author:

Hanna K. Castro

Abstract

Listeria monocytogenes is a food-borne pathogen and a resilient environmental saprophyte. Dairy farms are a reservoir of *L. monocytogenes* and strains can persist on farms for years. Here, we sequenced the genomes of 250 *L. monocytogenes* isolates to investigate the persistence and mobile genetic elements of *Listeria* inhabiting dairy farms. We found that prophages and other mobile elements were significantly more numerous among persistent than sporadically occurring strains. We identified a remarkable diversity of mobile elements among farm isolates, including a novel group of plasmids infecting hypervirulent subtypes of *L. monocytogenes* and occasionally carrying biocide resistance determinants *bcrABC* or *qacH*. Resistance genes against bacitracin, arsenic and cadmium were significantly more prevalent among persistent than sporadic strains. Several of the mobile elements in *Listeria* were identical to the mobile elements of *Enterococci*, indicative of recent transfer between these genera. Finally, we demonstrated that the CRISPR-*cas* IIa system and a type II restriction-modification system were negatively associated with persistence on farms. Our findings suggest that mobile elements support the persistence of *L. monocytogenes* on dairy farms and that *L. monocytogenes* inhabiting the agroecosystem is a potential reservoir of mobile elements harbouring resistance genes against antimicrobials, biocides, and heavy metals.

Introduction

Listeria monocytogenes leads a double life: in one it is a potentially lethal, zoonotic foodborne pathogen and in the other, a ubiquitous environmental saprophyte (Gray et al., 2006). Agroecosystems provide a favourable habitat for *L. monocytogenes* and the pathogen is especially prevalent on dairy farms (Nightingale et al., 2004; Esteban et al., 2009). *L. monocytogenes* strains can inhabit dairy farms for years and be widely distributed in the farm environment, leading to the frequent contamination of milk (Ho et al., 2007; Castro et al., 2018). Raw milk and animals destined for slaughter are a major contamination source for the food industry (Samelis & Metaxopoulos 1999; Fox et al., 2009; Hellström et al., 2010). Knowledge of the pathogens ecology on farms is essential for controlling the spread of *L. monocytogenes* from farms to the food industry.

L. monocytogenes is extremely resilient and can tolerate various stresses used to control the pathogen in the food industry (Lundén et al., 2003; Aarnisalo et al., 2007). These phenotypic traits enable *L. monocytogenes* to survive in food processing environments for years, a phenomenon known as persistence (Lundén et al., 2000; Keto-Timonen et al., 2007; Stasiewicz et al., 2015; Pasquali et al., 2018; Hurley et al., 2019). Mobile genetic elements are common among *L. monocytogenes* isolates from food processing environments (Harvey & Gilmour, 2001; Pasquali et al., 2018; Hurley et al., 2019) and may harbour genes mediating tolerance to heat shock (Pöntinen et al., 2017), salt and acid stress (Naditz et al., 2019; Hingston et al., 2019) and biocides (Müller et al., 2013; Meier et al., 2017). These findings led us to the hypothesis that mobile genetic elements play a key role in the environmental adaptation and persistence of *L. monocytogenes*.

Although dairy farms are considered a reservoir of *L. monocytogenes* (Nightingale et al., 2004), and are known to harbour hypervirulent strains (Maury et al., 2019), the era of next

generation sequencing has witnessed very few efforts to illuminate the pathogen's ecology in the farm environment. How *L. monocytogenes* adapts to life in the farm ecosystem, and to what extent the farm environment acts as source of mobile genetic elements for *L. monocytogenes* persisting in food processing environments, are key issues to explore. Such insights would be instrumental in developing novel strategies to reduce contamination on farms and to prevent the emergence and spread of persistent strains.

Here, we sequenced the genomes of 250 *L. monocytogenes* isolates obtained from three Finnish dairy farms during 2013–2016 (Castro et al., 2018) to investigate the persistence and mobile genomic elements of *L. monocytogenes* in the farm environment. We found that prophages and other mobile genetic elements were significantly more abundant among persistent than sporadic strains. We identified a remarkable diversity of mobile elements among the dairy farm isolates and discovered several novel mobile elements that are relevant to food safety. We demonstrated that resistance genes against bacitracin, arsenic and cadmium were significantly more prevalent among persistent than sporadic strains. Finally, we found a negative association between persistence and three putative defence systems against invading prophages and plasmids (Garneau et al, 2010; Lee et al., 2012). Taken together, our findings suggest that prophages and mobile genetic elements confer an ecological advantage for persistence on farms, and that *L. monocytogenes* inhabiting the farm environment constitutes a reservoir of diverse mobile genetic elements having the potential to carry resistance genes against heavy metals, antimicrobials and biocides across the food chain.

Results

77 **Persisting clades of *L. monocytogenes* were detected on all three farms**

78 Whole genome sequencing and subsequent *in silico* subtyping of 250 *Listeria monocytogenes*
79 isolates, collected from three dairy farms during 2013–2016 (Castro et al., 2018), yielded 25
80 unique STs (Fig. 1a, see Supplementary Data S1). The most frequently detected subtype was
81 ST20, which represented 28% of all sequenced isolates. We identified in total 14 persistent
82 clades (Table 1). Persistent clades represented 71% of all sequenced isolates and all persistent
83 clades belonged to serogroup 1/2a. Clade C4 contained isolates from two different farms,
84 suggesting that strains of *L. monocytogenes* can spread between farms faster than the rate of
85 genomic diversification.

86

87 **Mobiles genetic elements were more common among persistent clade isolates than**
88 **singleton isolates**

89 Prophages and other mobile genetic elements were significantly more numerous (Independent
90 Samples Median Test; $p<.01$) among persistent clade isolates than singleton isolates (Fig. 1b-
91 c). Overall, non-phage mobile genetic elements were detected in 72% of isolates in persistent
92 clades and 66% of singleton isolates. All isolates contained at least one prophage, known as
93 the monocin (Zink et al., 1995), and additional prophages were detected among 95% of
94 isolates from persistent clades and 85% of singleton isolates.

95

96 **Dairy farms harboured a novel plasmid that can infect hypervirulent STs and carry**
97 **biocide resistance genes**

98 Plasmids were detected among 10% of *L. monocytogenes* isolates in persistent clades and
99 11% singleton isolates. We detected three previously identified plasmids (pCFSAN010068,
100 pLM58, pLMR479a), and three novel plasmids, labelled pHC143, pHC192 and pHC195-2

(Fig. 2a, see Supplementary Data S1). These plasmids were 55.5 – 86.7 Kb in size, except for pHC192, which was only 4.6 Kb. A Maximum-Likelihood phylogenetic analysis based on RepA grouped the large plasmids into three distinct groups (G1 – G3), which appear to be specific to the genus *Listeria* (Fig. 3a). In the small plasmid pHC192 the *repA* gene was truncated so it was excluded from the analysis.

Plasmid groups G1 and G2 composed of several well-characterized *L. monocytogenes* reference plasmids (Kuenne et al., 2010; Hingston et al., 2019; Nadiz et al., 2019). G3 represented a phylogenetically distinct, novel group of *Listeria* plasmids that all closely resemble pHC143, a plasmid we detected in five isolates of this study (see Supplementary Data S1). Visualisation of assembly graphs indicated that the plasmid was successfully assembled into a single 55.8 Kb contig in all five isolates. We identified three previously unreported variants of pHC143 among short-read sequence assembles deposited in GenBank, which all contained resistance genes against biocides (Fig. 3b, see Supplementary Fig. S1). Specifically, the plasmid of the *L. monocytogenes* ST1 environmental isolate FDA550584-30 (SAMN02923676) contained a benzalkonium chloride resistance cassette (*bcrABC*), and the plasmids of the ST6 human outbreak isolate YA00079283 (SAMN08970420) and the ST20 food isolate 967535 (SAMN15680309) contained a resistance gene against quaternary ammonium compounds (*qacH*). Additionally, pFDA550584-30 contained a mercuric resistance (*mer*) operon, and pYA00079283 contained a Tn554-family transposon carrying an arsenic resistance operon (*arsABCD*). This Tn554-family transposon was identified previously in the chromosome of *L. monocytogenes* ST9 (Kuenne et al., 2013), a subtype especially adapted to the food processing environment (Maury et al., 2019). All G3 plasmids contained a predicted fimbrial adhesin (WP_061691480.1), suggestive of a role associated with attachment and host colonisation (Ageorges et al., 2020). Unlike G1 and G2 plasmids, which primarily infect serogroups 1/2a and 1/2b (Hingston et al., 2019), G3 plasmids appear

126 to primarily infect 4b, including *L. monocytogenes* STs 1 and 6, which are considered
127 hypervirulent (Maury et al., 2019). In this study, pHC143 was detected among ST6 and
128 ST149 isolates.

129

130 **Persistent clade C7 acquired a novel plasmid with a potentially broad host species range**

131 The 4.6 Kb plasmid pHC192 represented a separate phylogenetic group (G4) than the
132 previously reported 4.7 Kb *Listeria* plasmid pLMST6 (Kremer et al., 2017), which carries the
133 benzalkonium chloride resistance gene *ermC* (Fig. 4a). Phylogenetically, pHC192 clustered
134 closely with plasmids from *Lactobacillus*. Indeed, RepB of pHC192 (WP_035147907.1) was
135 also detected in *Lactobacilli* and *Brochothrix*, suggestive of a broad host range for this
136 plasmid. The closest relative of pHC192 in *Listeria* was the plasmid of the strain *L.*
137 *monocytogenes* strain CFIAFB20130002, which possesses the lincosamide resistance gene
138 *lnuA* (WP_001829870.1). Similarly, pLMST6 appears to also have a broad host range, as its
139 RepB (WP_061092472.1) homologues were detected in *Listeria, Salmonella* and
140 *Enterococcus*. Taken together, these findings suggest that unlike the large (>50 Kb) *Listeria*
141 plasmids, the small (<10 Kb) plasmids of *Listeria* have a broad host range and several
142 phylogenetically unrelated small plasmids have been acquired by *Listeria* through distinct
143 transfer events across host species.

144 The plasmid pHC192 contains a putative *tauE/safE* -family sulphite exporter gene
145 (WP_016896343.1) (Fig.4b) that is not typically present in *Listeria* plasmids (Hingston et al.,
146 2019). The sequencing depth of coverage for pHC192 was approximately five times that of
147 the chromosome, suggesting that pHC192 is a high copy number plasmid. This plasmid
148 became increasingly prevalent among persistent clade C7 isolates during the sampling period
149 and was detected in all isolates at the end of the study (Fig. 4c). An additional plasmid,

pHC195-2, was detected in several isolates in the latter part of the study period. The pHC195-2 plasmid belonged to the phylogenetic group G2 (Fig. 3a) and closely resembled the reference plasmid pLMR479a (see Supplementary Fig. S2).

Transposons harbouring bacitracin and heavy metal resistance genes were more prevalent among persistent clades than singleton isolates

Among the 250 dairy farm isolates we identified six transposable elements: the *L. monocytogenes* IS3-like element (Kuenne et al., 2013); *Listeria* Genomic Island 2 (LGI-2) (Lee et al., 2017); Tn5801_B23 (León-Sampedro et al., 2016); and three novel transposons, which were submitted to the Transposon Registry (Tansirichaiya et al., 2019), and assigned the labels Tn7101, Tn7103 and Tn7104. The conjugative transposons LGI-2 and Tn5801_B23 were significantly more prevalent among persistent clade isolates than singleton isolates (Fisher's Exact Test, $p<.01$) (Fig. 2b). LGI-2 is a conjugative transposon carrying cadmium and arsenic resistance cassettes and two multidrug transporters (see Supplementary Fig. S3). Identical (100% nucleotide identity) LGI-2 were present among all ST14 and ST145 isolates of this study. Moreover, BLASTn search identified identical LGI-2 in 11 *L. monocytogenes* and two *Enterococcus faecalis* complete genomes (see Supplementary Table S2), suggestive of recent transfer between these species.

The conjugative transposon Tn5801_B23 was detected in a subset of ST20 isolates, including the persistent clades C9 – C12 (see Supplementary Data S1). Tn5801_B23 detected in this study shared 97% identity with the Tn5801_B23 of the *E. faecalis* strain JH2-2 (see Supplementary Fig. S4). Tn5801_B23 is a conjugative transposon that contains an operon of resistance genes against the antimicrobial bacitracin (*bcrABC*) and a two-component system (*baeSR*) putatively involved in the regulation of the *bcrABC* operon (León-Sampedro et al.,

174 2016). Unlike Tn5801_B23, other Tn5801-like transposons mediate tetracycline resistance in
175 *Enterococcus, Listeria* and several other firmicute species (León-Sampedro et al., 2016). In
176 *L. monocytogenes* ST20, Tn5801_B23 was inserted downstream of *guaA* (*lmo1096*), which is
177 also the insertion site of the related conjugative transposon ICELm1 of the *L. monocytogenes*
178 strain EGD-e, harbouring cadmium resistance genes (Kuenne et al., 2013).

179

180 **Three novel transposons were identified among the farm isolates**

181 The novel transposon Tn7101 was detected in the ST155 singleton isolate HC258, where it
182 was inserted between *lmo2596* and *lmo2597* (see Supplementary Fig. S3). This transposon
183 contains resistance genes against cadmium (*cadA, cadC*) and an arsenate reductase (*arsC*).
184 Through a BLAST search we identified a variant of the Tn7101 containing a seven gene
185 arsenic resistance cassette. This variant, labelled Tn7102, was detected in several *L.*
186 *monocytogenes* and *Enterococcus* genomes deposited in GenBank (see Supplementary Fig.
187 S2). The Tn7101 and Tn7102 of *Listeria* and *Enterococcus* were identical (100% nucleotide
188 identity), suggestive of recent promiscuity between the two genera. Arsenic resistance genes
189 in Tn7102 were distantly related (≥67% identity) to the arsenic resistance cassette of LGI-2
190 (see Supplementary Fig. S3).

191 The novel conjugative transposon Tn7103 was detected in the ST119 singleton isolate
192 HC183, where it was inserted between *lmo0810* and *lmo0811*. This transposon contained
193 putative virulence genes encoding an InlJ-like internalin and a bacterial immunoglobulin
194 (Big)-like protein (see Supplementary Fig. S5). A BLAST search confirmed the presence of
195 Tn7103 in other *L. monocytogenes* strains, including N12-2532 (SAMN09947958), but we
196 did not identify this element in other species.

The novel conjugative transposon Tn7104 was detected in the ST391 singleton isolate HC187 and was inserted between *lmo1786* and *lmo1787*. This transposon contained a putative type I restriction-modification system (see Supplementary Fig. S6). Tn7104 was identified in several other *L. monocytogenes* strains deposited in GenBank, including the *L. monocytogenes* ST391 strain SHL013 (SAMN03265960), but we did not identify this element in other species.

Prophages were more prevalent among persistent clades than singleton isolates

All 250 dairy farm isolates from this study contained the *L. monocytogenes* monocin (Zink et al., 1995), and 0 – 3 additional prophages, which were detected at eight insertion sites (Fig. 2c). Prophages inserted into tRNA-Arg(tct) were significantly more prevalent among persistent clusters, and prophages inserted into tRNA-Lys(ctt) were significantly more prevalent among singleton isolates (Fisher's Exact Test, $p<.05$).

Prophages harbouring cadmium resistance genes were detected in persistent clade C8

OPTSIL taxonomic clustering assigned prophages from this study into six genera. Prophages inserted into *comK* and tRNA genes were assigned to genera of *Siphoviridae* that are known to only infect *Listeria*. Surprisingly, in the isolate HC189, a 67 Kb *Myovirus* was inserted into *comK*, a site usually occupied by *Siphoviridae* (Pasechnek et al., 2020).

Prophages inserted between the *rlmCD* (*lmo1703*) and *fosX* (*lmo1702*) genes were not related to any of the *Listeria* specific phage genera, but instead represented a separate genus that infects a broad range of *Firmicute* species (Fig. 5). Many of the phages in this genus harbour antimicrobial and heavy metal resistance cassettes (see Supplementary Fig. S7). In this study, phages inserted between the *rlmCD* and *fosX* were detected among all isolates of persistent

clade C8 and among three singleton isolates (see Supplementary Data S1). Among isolates of persistent clade C8, prophages inserted between *rlmCD* and *fosX* all harboured a cadmium resistance cassette (see Supplementary Fig. S7). In contrast, in the singleton isolates prophages inserted between *rlmCD* and *fosX* harboured no cadmium or antimicrobial resistance genes. Within *Listeria* genomes deposited in GenBank, we identified prophages inserted between *rlmCD* and *fosX* that carried resistance genes against cadmium (*cadA*), macrolides (*mefA*, *msrD*), tetracycline (*tetM*) and streptogramin (*vatA*).

Acquired heavy metal resistance genes were more prevalent among persistent clades than singleton isolates

The frequent occurrence of heavy metal resistance genes among *L. monocytogenes* persisting in food processing environments has led to the speculation that these determinants promote the environmental persistence of *L. monocytogenes* (Parsons et al., 2020). Indeed, heavy metal resistance genes were significantly more prevalent among persistent clade isolates than singleton isolates (Fig. 2d). Among the 250 sequenced isolates, we detected mobile genetic elements harbouring resistance genes against arsenic and cadmium (see Supplementary Fig. S8). Resistance genes against biocides and antimicrobials other than bacitracin were not detected in this study. However, variants of the mobile elements harbouring genes against biocides and antimicrobials were detected among farm and farm animal isolates of *L. monocytogenes* among genomes deposited in GenBank (see Supplementary Fig. S8 and Table S3). Taken together, the results of this study suggest that *L. monocytogenes* inhabiting the dairy farm environment constitutes a reservoir of diverse mobile genetic elements that have the potential to harbour resistance genes against antimicrobials, biocides, and heavy metals.

Three systems conveying immunity against invading DNA were negatively associated with persistence on dairy farms

The CRISPR-*cas* type IIA system, and the type II restriction-modification system LmoJ3 (Lee et al., 2012) were negatively associated with persistence in the dairy farm environment among *L. monocytogenes* Lineage II isolates (see Supplementary Table S4). CRISPR-*cas* systems and restriction-modification systems can act in synchrony to protect the host against invading prophages and other mobile elements (Price et al., 2016). Additionally, a putative recombination and DNA strand exchange inhibitor protein (WP_03166494.1) was negatively associated with persistence. These findings agree with the lower prevalence of mobile genetic elements and prophages among singleton isolates than persistent clade isolates and suggest that systems involved with inhibiting invading DNA are detrimental for the persistence of *L. monocytogenes* in the dairy farm environment.

Among the genes that were positively associated with persistence on dairy farms was a gene associated with biofilm formation (*bapL*) and a peptidoglycan hydrolase (*murA*). *L. monocytogenes* hypervariable hotspots 1 and 8 (Kuenne et al., 2013) contained genes both positively and negatively associated with persistence. Indeed, most genes associated with persistence belonged to *L. monocytogenes* hypervariable hotspots or prophages, suggesting that the role these components play in *Listeria* niche adaptation deems further study.

Discussion

Whole genome sequencing and subsequent analyses of 250 *L. monocytogenes* isolates from dairy farms illustrated that dairy farm isolates are hosts to a diversity of mobile genetic elements that carry, or have the potential to carry, resistance genes against antimicrobials, biocides and heavy metals. We found that prophages and other mobile genomic elements

were significantly more numerous among isolates belonging to persistent clades than among singleton isolates. Moreover, systems that provide immunity against invading mobile genetic elements (Garneau et al, 2010; Lee et al., 2012; Price et al., 2016), namely the CRISPR-*cas* IIA system, the type II restriction-modification system LmoJ3, and a putative recombination and DNA strand exchange inhibitor protein, were negatively associated with persistence. Taken together, our findings suggest that for *L. monocytogenes* inhabiting the farm environment, the benefits of being receptive to potentially advantageous mobile elements outweigh the benefits protecting against elements that are potentially deleterious. Many of the mobile elements we identified carried genes encoding phenotypes that promote the survival of *L. monocytogenes* on farms, such as antimicrobial resistance genes or virulence factors. Moreover, genes responsible for the conjugation of mobile elements may have a dual role in promoting biofilm formation and invasion of the mammalian host (Barrios et al., 2005; Ageorges et al., 2020), thereby supporting the survival of *L. monocytogenes* on farms.

We identified a surprising diversity of mobile genetic elements encoding heavy metal resistance genes among the dairy farm isolates and acquired heavy metal resistance genes were encountered more frequently among persistent clade isolates that singleton isolates. Similarly, heavy metal resistance genes are more prevalent among persistent than sporadic *L. monocytogenes* subtypes from foods and food processing environments (Harvey & Gilmour, 2001; Pasquali et al., 2018). Whether heavy metal resistance genes contribute to persistence or merely co-occur with other fitness enhancing determinants remains unclear (Parsons et al., 2020). Nevertheless, heavy metal resistance genes may represent useful markers to aid the detection of *L. monocytogenes* strains with fitness-enhancing mobile genetic elements.

In the present study we found two novel plasmids, namely pHC143 (plasmid group G3) and pHC192 (plasmid group G4), that were unrelated to the previously described plasmids of *Listeria*. The comparison of these plasmids against the short-read sequence data deposited in

GenBank led to the identification of three novel G3 plasmid variants carrying genes against biocides and heavy metals. Importantly, G3 plasmids can infect hypervirulent ST1 and ST6 strains, potentially enabling these STs to adapt to the food processing environment through the acquisition of biocide and heavy metal resistance. Indeed, we discovered a G3 plasmid harbouring the biocide resistance gene *qacH* and the arsenic resistance cassette *arsABCD* in the ST6 outbreak isolate YA00079283, associated with the largest listeriosis outbreak known to date (Smith et al., 2019).

We identified four transposons in *Listeria*, namely LGI-2, Tn5801_B23, Tn7101 and Tn7102, that closely resembled transposons in *Enterococci*, suggestive of recent transfer between the two genera. The co-occurrence of genomic elements in *Enterococci* to *Listeria* was unsurprising, as both genera are highly prevalent in animal faeces and farms (Franz et al., 1999; Hellström et al., 2010; Castro et al., 2018). Transfer of conjugative elements has been demonstrated both from *Enterococci* to *Listeria* and vice versa (Jahan & Holley, 2016; Haubert et al., 2018) indicating that both genera are potential donors. The extent to which *Enterococci* and other *Firmicutes* contribute to the horizontal spread of mobile elements and their associated antimicrobial, biocide and heavy metal resistance determinants in *Listeria* has implications for food safety and should be explored through further study.

Bacitracin resistance genes, mediated by Tn5801_B23, were significantly more prevalent among isolates in persistent clades than singleton isolates. Moreover, *L. monocytogenes* isolates harbouring Tn5801_B23 were detected frequently on all three farms investigated. It is plausible that the widespread use of bacitracin as a growth promoter in animal feeds during the mid-twentieth century facilitated the expansion of bacitracin resistance in *Listeria*. Although animal feeds are frequently contaminated by *L. monocytogenes* (Hellström et al., 2010; Castro et al., 2018), the frequent detection of Tn5801_B23 on the dairy farms

318 investigated remains curious, as feed supplementation with bacitracin has been banned in the
319 EU since the 1990's.

320 We found that prophages were more prevalent among persistent clade isolates than singleton
321 isolates. It is unclear whether the higher number of prophages among persistent clade isolates
322 represents a beneficial role for these elements or is a side effect of being receptive to foreign
323 DNA. There is increasing evidence that prophages can mediate beneficial phenotypes for
324 their host. Phages mediate resistance or virulence properties in numerous bacterial species
325 (Harrison & Brockhurst, 2017), and in *Listeria*, *Siphoviruses* inserted into *comK* were found
326 to regulate the gene in a symbiotic manner (Pasechnek et al. 2020). Here, we discovered
327 phage-mediated carriage of cadmium resistance and various antimicrobials in *Listeria*,
328 suggesting that prophages contribute to the spread of phenotypes supporting persistence.
329 Moreover, we noted that these phages belonged to a genus of *Siphovirus* with an apparently
330 broad host species range that were introduced to *Listeria* through several distinct transfer
331 evets. Host species jumps have the potential accelerate the transfer of novel resistance
332 determinants between *Listeria* and other *Firmicutes*.

333 In conclusion, our study indicates that *L. monocytogenes* inhabiting the dairy farm
334 environment are receptive to a diversity of prophages and mobile genetic elements. We
335 suggest that mobile elements enable *L. monocytogenes* to adapt to the stresses encountered in
336 the farm ecosystem and in general improve the fitness of the pathogen on farms, thereby
337 supporting persistence. Given the abundance of *L. monocytogenes* on farms farms
338 (Nightingale et al., 2004; Esteban et al., 2009; Castro et al., 2018) and the apparent exchange
339 of mobile genetic elements between *Listeria* and other *Firmicute* species, *L. monocytogenes*
340 occurring in agroecosystems should be viewed as a reservoir of mobile genetic elements.
341 Importantly, many of these elements have the potential to carry and spread antimicrobial,
342 biocide, and heavy metal resistance genes. The spread of mobile genetic elements and

resistance determinants from primary production to *Listeria* in the food processing environments has important food safety implications and should be explored further. The present study represents a leap forward in this effort and in our understanding of listerial ecology in the agroecosystem.

Methods

Whole genome sequencing

Altogether 250 *L. monocytogenes* isolates obtained from three Finnish dairy cattle farms during 2013–2016 (Castro et al., 2018) were selected for whole genome sequencing in the present study (see Supplementary Data S1). DNA was extracted from overnight cultures using the guanidium thiocyanate extraction method (Pitcher et al. 1989). DNA samples were standardized to a concentration of 10 ng/µl using the dsDNA BR Assay Kit (Thermo Fisher Scientific; Waltham, MA, USA) using the Qubit Fluorometer (Thermo Fisher Scientific). Genomic libraries were constructed form the DNA samples using the Nextera XT DNA Sample Preparation Kit (Illumina; San Diego, CA, USA), and paired-end sequencing (2×250 bp) was performed using the Illumina HiSeq platform.

Genome assembly, pangenome construction, subtyping

Following the removal of adapter sequences and low-quality reads using Trimmomatic 0.36 (Bolger et al., 2014), draft genomes were assembled using SPAdes 3.9 with K-mer values 55, 77, 99, 113 and 127 (Bankevich et al., 2012). Assembly quality was assessed using QUAST 4.0 (Gurevich et al., 2013) and taxonomic assignment was performed using Kraken (Wood &

Salzberg, 2014). The assemblies were annotated using Prokka 1.12 (Seemann, 2014). The pangenome of the sequenced isolates was constructed using Roary 3.8.0 (Page et al., 2015) with the protein identity cut-off value set at 90%. Multi Locus Sequence Types (ST) were determined *in silico* from the assembled genomes using the BIGSdb-*Lm* database (Moura et al., 2017). Genome assemblies were deposited in GenBank under the BioProject number PRJNA704814 (see Supplementary Data S1).

Maximum-Likelihood Phylogenomic Analysis

Phylogenomic reconstruction of the *L. monocytogenes* isolates was performed using the Lyve-SET 1.1.4f pipeline (Katz et al., 2017). For each ST, a draft assembly from the present study with the best quality statistics, i.e. the highest N50 value and lowest number of contigs (see Supplementary Data S1), was used as a reference genome. The Lyve-SET pipeline was run using *Listeria monocytogenes* pre-sets (Katz et al., 2017), with additional options --mask-phages, --mask-cliffs, and --read_cleaner CGP. In brief, the pipeline generated genome alignments by mapping quality-filtered reads to a reference genome. To improve the accuracy of phylogenomic inference, putative prophage genes were removed from the reference genome prior to mapping. Mapping was followed by the detection of high-quality SNPs, having ≥ 10x depth of coverage and ≥ 75% consensus among reads. Recombinant sites within the genome alignments generated by Lyve-SET were identified and removed using (Croucher et al., 2015). PhyML 3.3 (Guindon et al., 2010) was used to infer Maximum-Likelihood - phylogeny of each ST using a general time reversible model (GTR) with 100 bootstrap replicates. Persistent clades of *L. monocytogenes* were defined monophyletic clades of isolates with PWDs < 20 SNPs (Pightling et al., 2018) that were isolated from the same farm from ≥3 samples during ≥6 months.

390

391 **Detection and analysis of plasmids**

392 Putative plasmid sequences were assembled from read files using SPAdes with option --
393 plasmid and the assembled contigs were annotated using Prokka 1.12. Additionally, whole-
394 genome assembly graphs generated by SPAdes were visualized using Bandage 0.8.1 (Wick et
395 al., 2015) and extrachromosomal elements were inspected manually. Similarity of the
396 plasmid sequences we identified to plasmids deposited in GenBank was assed using the
397 BLAST suite (http://www.ncbi.nlm.nih.gov/blast). Maximum-Likelihood phylogeny of the
398 plasmids, based on the translated sequence alignments of the *repA* gene, were generated with
399 MEGA7 (Kumar et al., 2016), using the Jones-Taylor-Thornton substitution model with 100
400 bootstraps. Alignments of the translated sequences of the *repB* gene were used to compare
401 plasmids in which *repA* was truncated or absent. Plasmid alignments were generated and
402 visualized using BRIG 0.95 (Alikhan et al., 2018) and EasyFig 1.2 (Sullivan et al., 2011).

403

404 **Detection and analysis of transposons**

405 The occurrence transposons ICELm1 (Kuenne et al., 2013), LGI-2 (Lee et al., 2017), Tn5422
406 (Lebrun et al., 1994), Tn6188 (Müller et al., 2013), and the IS3-like and TN554-like
407 transposons of *L. monocytogenes* (Kuenne et al., 2013), among isolates from this study was
408 assed using BLAST. Additionally, the pangenome was searched for annotations including
409 "recombinase", "integrase", "transposase", "transposon", "cadmium", "arsenic", "mercuric",
410 "*ardA*", "*ftsK*", "P60" and "*iap*" and hits were inspected manually. EasyFig 1.2 was used to
411 align and visualize the identified transposons and their occurrence among genomes deposited
412 in GenBank was assessed using BLAST.

413

Detection and analysis of prophages

Prophages inserted into the *L. monocytogenes* genomes were identified using PHASTER (Arndt et al., 2016) and the insertion sites were inspected manually. Phylogeny and taxonomic clustering of prophages classified by the PHASTER algorithm as "intact" were inferred using VICTOR (Meier-Kolthoff & Göker, 2017). Nineteen additional *Listeria* phage genomes and one Streptococcal phage genome obtained from GenBank were included in the analyses for reference (see Supplementary Table S1). In brief, VICTOR applies the Genome-BLAST Distance Phylogeny (GBDP) method (Meier-Kolthoff et al., 2013) to obtain pairwise distances, from which balanced minimum evolution trees are inferred. VICTOR utilizes OPTSIL (Göker et al., 2009) to obtain taxonomic clustering. Duplicate phage genomes are removed from the analysis. Trees generated by VICTOR were visualized using FigTree 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). BLAST was used to identify phages inserted between *rlmCD* and *fosX* in the genomes of *Listeria* and other bacterial species deposited in GenBank, and hits were inspected manually. Phylogeny and taxonomic clustering of prophages inserted between *rlmCD* and *fosX* were inferred using VICTOR.

Identification of genes associated with predominance

Scoary 1.6.16 (Brynildsrud et al., 2016) was used to identify genes significantly associated with occurrence in persistent clade isolates versus singleton isolates. Scoary was executed using default options, using the gene_presence_abence.csv -file generated by Roary as input. Associations with a Bonferroni corrected $p<.05$ were considered significant. As all predominant clades belonged to Lineage II, the analysis was limited to the 233 Lineage II isolates of this study to reduce noise arising from population structure bias.

437 **References**

Aarnisalo, K., Lundén, J., Korkeala, H. & Wirtanen, G. (2007) Susceptibility of *Listeria monocytogenes* strains to disinfectants and chlorinated alkaline cleaners at cold temperatures. *LWT-Food Sci. Technol.* **40**, 1041–1048.

Ageorges, V., Monteiro, R., Leroy, S., Burgess, C. M., Pizza, M., Chaucheyras-Durand, F., & Desvaux, M. (2020) Molecular determinants of surface colonisation in diarrhoeagenic *Escherichia coli* (DEC): from bacterial adhesion to biofilm formation. *FEMS Microbiol. Rev.* **44**, 314–350.

Alikhan, N. F., Petty, N. K., Zakour, N. L. & Beatson, S. A. (2011) BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*. **12**, 1–10.

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 Res*. **44**, W16–W21.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S. *et al*. (2012) SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. **19**, 455–477.

Barrios, A. F., Zuo, R., Ren, D. & Wood, T. K. (2006) Hha, YbaJ, and OmpA regulate *Escherichia coli* K12 biofilm formation and conjugation plasmids abolish motility. *Biotechnol Bioeng.* **93**, 188–200.

Bolger, A. M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. **30**, 2114–2120.

Brynildsrud, O., Bohlin, J., Scheffer, L. & Eldholm, V. (2016) Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol*. **17**, 1–9.

Castro, H., Jaakkonen, A., Hakkinen, M., Korkeala, H. & Lindström, M. (2018) Occurrence,

persistence, and contamination routes of *Listeria monocytogenes* genotypes on three Finnish dairy cattle farms: a longitudinal study. *Appl. Environ. Microbiol.* 84, e02000-17.

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D. *et al.* (2015) Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res*. 43, e15–e15.

Esteban, J. I., Oporto, B., Aduriz, G., Juste, R. A. & Hurtado, A. (2009) Faecal shedding and strain diversity of *Listeria monocytogenes* in healthy ruminants and swine in Northern Spain. *BMC Vet. Res.* **5**, 1–10.

Fox, E., O'Mahony, T., Clancy, M., Dempsey, R., O'Brien, M. & Jordan, K. (2009) *Listeria monocytogenes* in the Irish dairy farm environment. *J. Food Protect.* **72**, 1450–1456.

Franz, C. M., Holzapfel, W. H. & Stiles, M. E. (1999) *Enterococci* at the crossroads of food safety?. *Int. J. Food Microbiol*. **47**, 1–24.

Garneau, J. E., Dupuis, M. È., Villion, M., Romero, D. A., Barrangou, R., Boyaval, P. *et al*. (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. **468**, 67–71.

Göker, M., García-Blázquez, G., Voglmayr, H., Tellería, M. T. & Martín, M. P. (2009) Molecular taxonomy of phytopathogenic fungi: a case study in *Peronospora*. *PloS One*. **4**, e6319.

Gray, M. J., Freitag, N. E. & Boor, K. J. (2006) How the bacterial pathogen *Listeria monocytogenes* mediates the switch from environmental Dr. Jekyll to pathogenic Mr. Hyde. *Infect. Immun*. **74**, 2505–2512.

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol*. **59**, 307–321.

Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. (2013) QUAST: quality assessment tool

for genome assemblies. *Bioinformatics*. **29**, 1072–1075.

Harrison, E. & Brockhurst, M. A. (2017) Ecological and evolutionary benefits of temperate phage: what does or doesn't kill you makes you stronger. *BioEssays*. **39**, 1700112.

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**, 840–847.

Haubert, L., da Cunha, C. E., Lopes, G. V. & da Silva, W. P. (2018) Food isolate *Listeria monocytogenes* harboring *tetM* gene plasmid-mediated exchangeable to *Enterococcus faecalis* on the surface of processed cheese. *Food Res. Int.* **107**, 503–508.

Hellström, S., Laukkanen, R., Siekkinen, K. M., Ranta, J., Maijala, R., & Korkeala, H. (2010) *Listeria monocytogenes* contamination in pork can originate from farms. *J Food Protect.* **73**, 641–648.

Hingston, P., Brenner, T., Truelstrup Hansen, L. & Wang, S. (2019). Comparative analysis of *Listeria monocytogenes* plasmids and expression levels of plasmid-encoded genes during growth under salt and acid stress conditions. *Toxins*. **11**, 426.

Ho, A. J., Lappi, V. R. & Wiedmann, M. (2007) Longitudinal monitoring of *Listeria monocytogenes* contamination patterns in a farmstead dairy processing facility. *J. Dairy Sci.* **90**, 2517–2524.

Hurley, D., Luque-Sastre, L., Parker, C. T., Huynh, S., Eshwar, A. K., Nguyen, S. V., *et al*. (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.

Jahan, M. & Holley, R. A. (2016) Transfer of antibiotic resistance from *Enterococcus faecium* of fermented meat origin to *Listeria monocytogenes* and *Listeria innocua*. *Lett. Appl.*

Microbiol. **62**, 304–310.

Katz, L. S., Griswold, T., Williams-Newkirk, A. J., Wagner, D., Petkau, A., Sieffert, C., *et al*. (2017) A comparative analysis of the Lyve-SET phylogenomics pipeline for genomic epidemiology of foodborne pathogens. *Front Microbiol.* **8**, 375.

Keto-Timonen, R., Tolvanen, R., Lunden, J. & Korkeala, H. (2007) An 8-year surveillance of the diversity and persistence of *Listeria monocytogenes* in a chilled food processing plant analyzed by amplified fragment length polymorphism. *J. Food Protect.* **70**, 1866–1873.

Kremer, P. H. C., Lees, J. A., Koopmans, M. M., Ferwerda, B., Arends, A. W. M., Feller, M. M. *et al*. (2017) Benzalkonium tolerance genes and outcome in *Listeria monocytogenes* meningitis. *Clin. Microbiol. Infect.* **23**, 265.e1–265.e7.

Kuenne, C., Voget, S., Pischimarov, J., Oehm, S., Goesmann, A., Daniel, R. *et al*. (2010) Comparative analysis of plasmids in the genus *Listeria*. *PloS One*. **5**, e12511.

Kuenne, C., Billion, A., Mraheil, M. A., Strittmatter, A., Daniel, R., Goesmann, A., *et al*. (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**, 1–19.

Kumar, S., Stecher, G. & Tamura, K. (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*. **33**, 1870–1874.

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

Lee, S., Ward, T. J., Jima, D. D., Parsons, C. & Kathariou S. (2017) The arsenic resistance-associated *Listeria* genomic island LGI2 exhibits sequence and integration site diversity and a propensity for three *Listeria monocytogenes* clones with enhanced virulence. *Appl. Environ. Microbiol*. **83**, e01189-17.

Lee, S., Ward, T. J., Siletzky, R. M. & Kathariou, S. (2012) Two novel type II restriction-modification systems occupying genomically equivalent locations on the chromosomes of *Listeria monocytogenes* strains. *Appl. Environ. Microbiol.* **78**, 2623–2630.

León-Sampedro, R., Novais, C., Peixe, L., Baquero, F. & Coque, T. M. (2016) Diversity and evolution of the Tn5801-*tet*(M)-like integrative and conjugative elements among *Enterococcus*, *Streptococcus*, and *Staphylococcus*. *Antimicrob. Agents Chemother*. **60**, 1736–1746.

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. *Int. J. Food Microbiol.* **82**, 265–272.

Maury, M. M., Bracq-Dieye, H., Huang, L., Vales, G., Lavina, M., Thouvenot, P., *et al*. (2019). Hypervirulent *Listeria monocytogenes* clones' adaption to mammalian gut accounts for their association with dairy products. *Nat. Commun.* **10**, 2488.

Meier, A. B., Guldimann, C., Markkula, A., Pöntinen, A., Korkeala, H., & Tasara, T. (2017) Comparative phenotypic and genotypic analysis of Swiss and Finnish *Listeria monocytogenes* isolates with respect to benzalkonium chloride resistance. *Front. Microbiol*. **8**, 397.

Meier-Kolthoff, J. P., Auch, A. F., Klenk, H.P. & Göker, M. (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*. **14**, 1–4.

Meier-Kolthoff, J. P., Göker, M. (2017) VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics.* **33**, 3396–3404.

Miettinen, M. K., Björkroth, K. J. & Korkeala, H. J. (1999) Characterization of *Listeria monocytogenes* from an ice cream plant by serotyping and pulsed-field gel electrophoresis. *Int. J. Food Microbiol.* **46**, 187–192.

Moura, A., Criscuolo, A., Pouseele, H., Maury, M. M., Leclercq, A., Tarr, C. *et al*. (2016)

Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat. Microbiol*. **2**, 16185.

Müller, A., Rychli, K., Muhterem-Uyar, M., Zaiser, A., Stessl, B., Guinane, C. M. *et al*. (2013) Tn6188 -a novel transposon in *Listeria monocytogenes* responsible for tolerance to benzalkonium chloride. *PloS One*. **8**, e76835.

Naditz, A. L., Dzieciol, M., Wagner, M. & Schmitz-Esser, S. (2019) Plasmids contribute to food processing environment–associated stress survival in three *Listeria monocytogenes* ST121, ST8, and ST5 strains. *Int. J. Food Microbiol*. **299**, 39–46.

Nightingale, K. K., Schukken, Y. H., Nightingale, C. R., Fortes, E. D., Ho, A. J., Her, Z. *et al*. (2004) Ecology and transmission of *Listeria monocytogenes* infecting ruminants and in the farm environment. *Appl. Environ. Microbiol.* **70**, 4458–4467 (2004).

Page, A. J., Cummins, C. A., Hunt, M., Wong, V. K., Reuter, S., Holden, M. T. *et al*. (2015) Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*. **31**, 3691–3693.

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.

Pasechnek, A., Rabinovich, L., Stadnyuk, O., Azulay, G., Mioduser, J., Argov, T. *et al*. (2020) Active lysogeny in *Listeria monocytogenes* is a bacteria-phage adaptive response in the mammalian environment. *Cell Rep*. **32**, 107956.

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. *Front. Microbiol.* **9**, 596.

Pightling, A. W., Pettengill, J. B., Luo, Y., Baugher, J. D., Rand, H., & Strain, E. (2018) Interpreting whole-genome sequence analyses of foodborne bacteria for regulatory applications and outbreak investigations. *Frontiers Microbiol*. **9**, 1482.

Pitcher, D. G., Saunders, N. A. & Owen, R. J. (1989) Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett. Appl. Microbiol.* **8**, 151–156.

Pöntinen, A., Aalto-Araneda, M., Lindström, M. & Korkeala, H. (2017) Heat resistance mediated by pLM58 plasmid-borne ClpL in *Listeria monocytogenes*. *Msphere*. **2**, e00364-17.

Price, V. J., Huo, W., Sharifi, A. & Palmer, K. L. (2016) CRISPR-Cas and restriction-modification act additively against conjugative antibiotic resistance plasmid transfer in *Enterococcus faecalis*. *Msphere*. **1**, e00064-16.

Samelis, J. & Metaxopoulos, J. (1999) Incidence and principal sources of *Listeria* spp. and *Listeria monocytogenes* contamination in processed meats and a meat processing plant. *Food Microbiol.* **16**, 465–477.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. **30**, 2068–2069.

Smith, A. M., Tau, N. P., Smouse, S. L., Allam, M., Ismail, A., Ramalwa, N. R. *et al*. (2019) Outbreak of *Listeria monocytogenes* in South Africa, 2017–2018: laboratory activities and experiences associated with whole-genome sequencing analysis of isolates. *Foodborne Pathog. Dis.*16, 524–530.

Stasiewicz, M. J., Oliver, H. F., Wiedmann, M. & den Bakker, H. C. (2015) Whole-genome sequencing allows for improved identification of persistent *Listeria monocytogenes* in food-associated environments. *Appl. Environ. Microbiol.* **81**, 6024–6037.

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

Tansirichaiya, S., Rahman, M. A. & Roberts, A.P. (2019) The transposon registry. Mobile DNA. **10**, 1–6.

Wick, R. R., Schultz, M. B., Zobel, J. & Holt, K. E. (2015) Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics*. **31**, 3350–3352

Wood, D. E., & Salzberg, S. L. (2014) Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*. **15**, 1–2.

Zink, R., Loessner, M. J., Scherer, S. (1995) Charaterization of cryptic prophages (monocins) in *Listeria* and sequence analysis of a holin/endolysin gene. *Microbiology*. **141**, 2577–2584.

438

Acknowledgements

This research was supported by the Finnish Ministry of Agriculture and Forestry (grant number 618/03.01.02/2017) and by the Walter Ehrström Foundation. The authors want to acknowledge Esa Penttinen for his assistance with DNA extraction.

Author Contributions

H.C. conducted the analyses and wrote the manuscript. H.C. and M.L. provided the study materials and designed the study. F.D., H.K. and M.L. supervised the study. All authors participated in data interpretation and reviewed the manuscript.

Competing Interests

The authors have no competing interests to declare.

Figure legends

Fig. 1. Prophages and other mobile genetic elements (MGE) were more numerous among persistent clade isolates than singleton isolates.

a *L. monocytogenes* isolates of this study represented 25 unique STs and persistent clades were detected among the six most prevalent STs. Each circle represents a unique ST and the area of the circle corresponds to the number of isolates. Doughnut charts illustrate the proportion of persistent clades (pink) and singleton isolates (aquamarine) for each STs in which persistent clades were detected.

b, **c** Distribution of persistent clade isolates and singleton isolates according to the number of non-phage MGE (**b**) and prophages (**c**) in each genome. The average number of the elements per genome is also given.

Fig. 2. Occurrence of mobile genetic elements and heavy metal resistance genes among persistent clades and singleton isolates.

Occurrence of plasmids (**a**), transposons (**b**), prophages (**C**) and cadmium and arsenic resistance genes (**d**) among persistent clade isolates and singleton isolates. Significant differences between persistent clade isolates and singleton isolates, identified using the Fisher's Exact Test, are denoted by asterisks: *: $p<.05$; *: $p<.05$; ***: $p<.001$. IS3: *Listeria* IS3-like transposon; LGI-2: *Listeria* Genomic Island 2. Prophages are categorized by insertion site.

Fig. 3. pHC249 represents a novel *Listeria* plasmid group G3.

a Maximum-Likelihood phylogenetic analysis of the >50 Kb plasmids detected in the present

study and other plasmids of *Listeria* and other *Firmicutes* based on the RepA protein sequences. The analysis employed the Jones-Taylor-Thornton substitution model with 100 bootstraps and was performed using the MEGA7 software. Bootstrap support values above 70 are shown and the scale bar depicts similarity. Plasmids represented three phylogenetic clades (plasmid groups G1-G3). Tip labels correspond to plasmid names and host genera; plasmids from this study are labelled in blue.

b G3 plasmids of the *L. monocytogenes* strains HC193 (this study), HC374 (this study) and FDA550584-30 (SAMN02923676) aligned with >95% identity across the entire length of pHC143 from this study; plasmids of the *L. monocytogenes* strains 967535 (SAMN15680309) and YA00079283 (SAMN08970420) aligned with >95% identity to most of pHC143. The alignment was generated using the BRIG 0.95. For pHC143, plasmid length in base pairs (bp) is given.

Fig. 4. Plasmid pHC192 was potentially acquired from a different *Firmicute* species and became increasingly prevalent among persistent clade C7.

a Maximum-Likelihood phylogenetic analysis of the <10 Kb plasmids detected in the present study and other related plasmids the RepB protein sequences. Plasmids other than pHC192 were identified and obtained from GenBank using BLASTp. The analysis employed the Jones-Taylor-Thornton substitution model with 100 bootstraps and was performed using the MEGA7 software. Bootstrap support values above 70 are shown and the scale bar depicts similarity. Tip labels correspond to plasmid names and host genera; plasmids from this study are labelled in blue. Tip shapes depict harbourage of resistance genes against antimicrobials (AMR) and benzalkonium chloride (BCR).

b The 4.5 Kbp plasmid pHC192, carrying a putative SafE/TauE family sulphite exporter (WP_016896343.1). The figure was constructed using the BRIG 0.95. Plasmid length in base pairs (bp) is given.

c Number of samples containing no plasmid, pHC192, or both pHC192 and pHC195-2 among persistent clade C7 during each month of sampling. Plasmid prevalence in C7 increased over the one-year sampling period.

Fig. 5. Prophages inserted between *rlmCD* and *fosX* belonged genus of *Siphovirus* having a broad host species range and a tendency to harbour antimicrobial or heavy metal resistance genes.

Minimum evolutionary tree and taxonomic clustering of six *Listeria* specific phages (genera 2-6); two prophages from this study that were inserted between *rlmCD* and *fosX* genes (genus 1, blue); and prophages from *Listeria* and other *Firmicutes* obtained from GenBank (genus 1, black). Phylogenetic analyses and clustering were generated with the Victor online tool (https://victor.dsmz.de), using the model D_6 and 100 bootstrap replicates. The tree was visualized using FigTree 1.4.4. Bootstrap support values above 70 are shown.

515 **Table 1.** Pairwise distances within persistent clusters of *L. monocytogenes* from dairy farms
516 A – C.

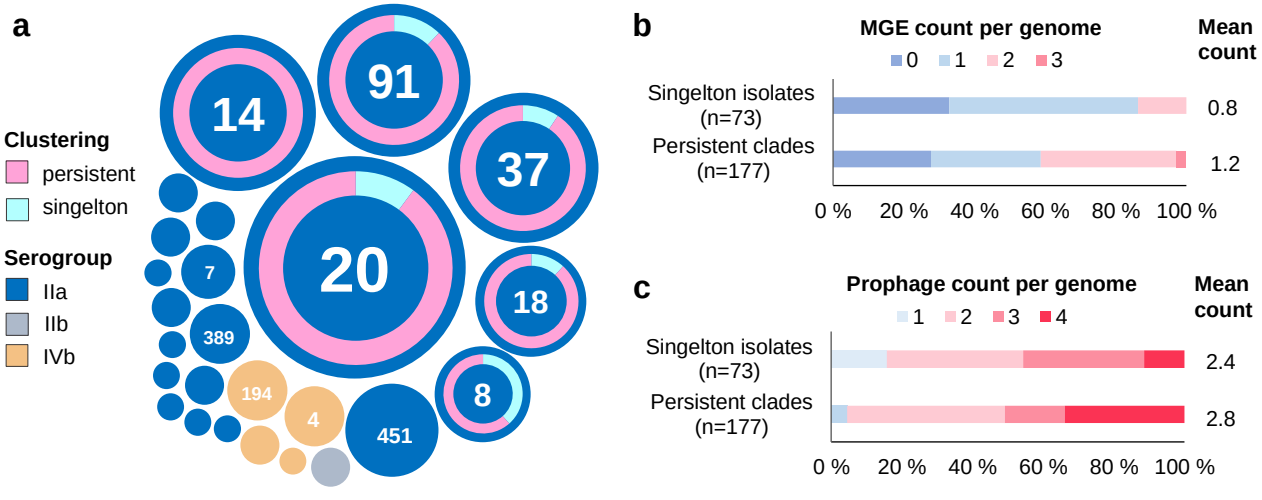
Cluster	CC[a]	ST[b]	N[c]	Farm	Pairwise distances		
					Mean	Minimum	Maximum
C1	8	8	8	A	1.5	0	4
C2	14	14	34	A	3.6	0	12
C3	14	91	18	A	2.6	0	7
C4	14	91	8	A, B	2	0	6
C5	18	18	8	B	3.3	0	8
C6	18	18	6	B	1.7	0	10
C7	20	20	32	C	2.4	0	7
C8	20	20	5	B	3.5	0	9
C9	20	20	6	A	3.9	0	10
C10	20	20	5	B	2.4	0	6
C11	20	20	9	C	5.8	0	11
C12	20	20	4	B	1.5	0	6
C13	37	37	20	A	2.6	0	7
C14	37	37	7	A	2.2	0	6
All clusters					2.8	0	8

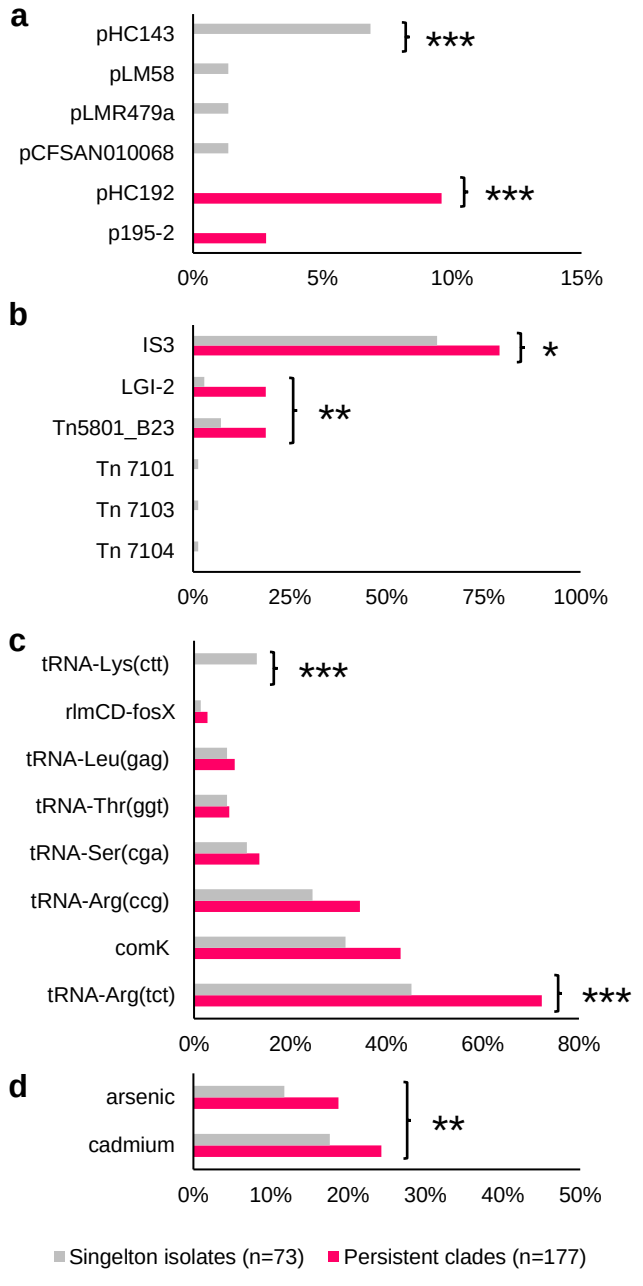
517 [a] CC: Clonal complex

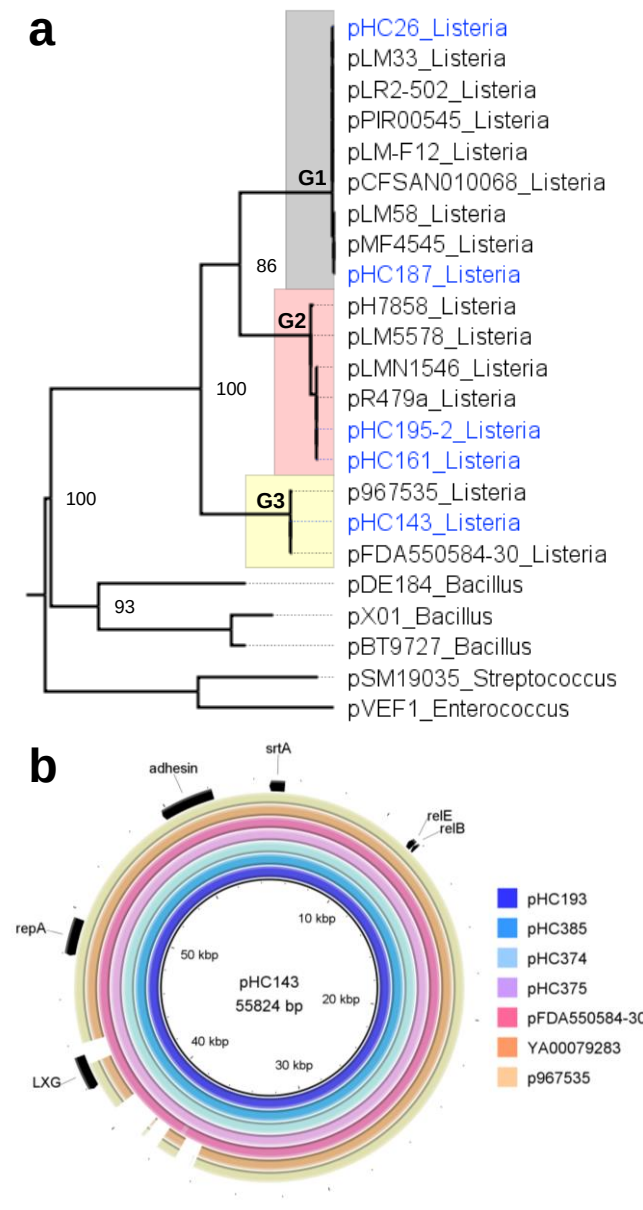
518 [b] ST: Sequence type

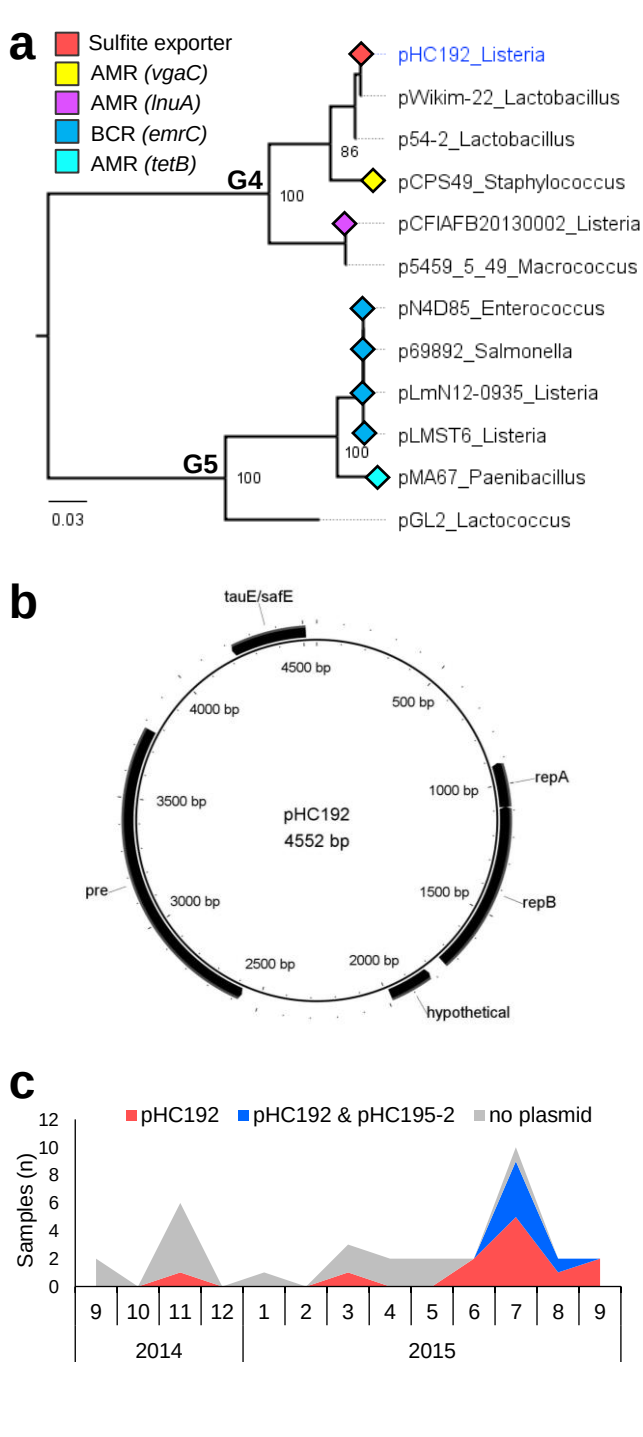
519 [c] N: Number of isolates in the persistent cluster.

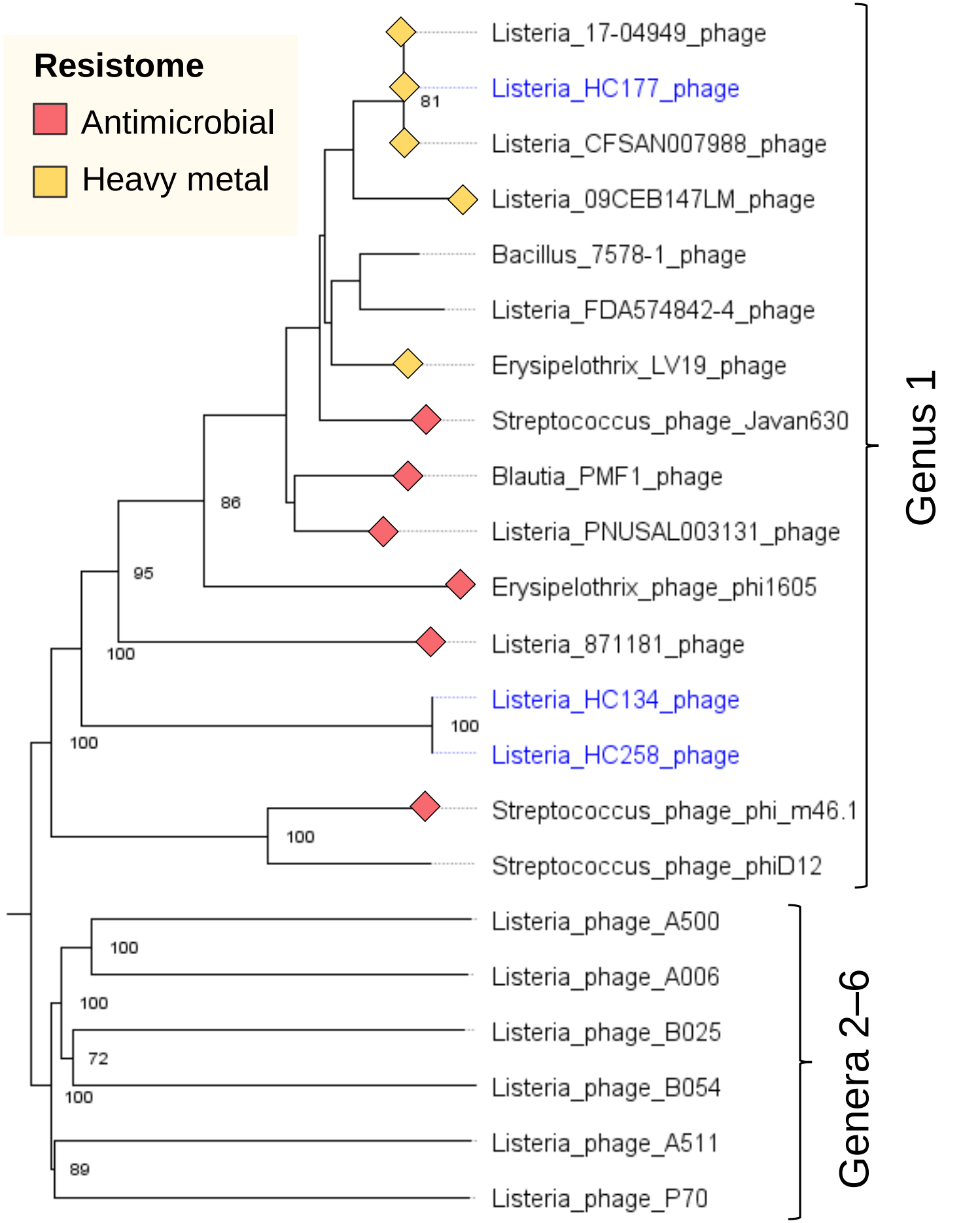
520


a
Clustering
persistent
singelton
Serogroup
IIa
IIb
IVb
14
91
37
20
7
18
389
194
4
8
451
b
MGE count per genome
Mean count
0
1
2
3
Singelton isolates (n=73)
Persistent clades (n=177)
0.8
1.2
0 %
20 %
40 %
60 %
80 %
100 %
c
Prophage count per genome
Mean count
1
2
3
4
Singelton isolates (n=73)
Persistent clades (n=177)
2.4
2.8
0 %
20 %
40 %
60 %
80 %
100 %


a
pHC143
pLM58
pLMR479a
pCFSAN010068
pHC192
p195-2
***
***
0%
5%
10%
15%
b
IS3
LGI-2
Tn5801_B23
Tn 7101
Tn 7103
Tn 7104
*
**
0%
25%
50%
75%
100%
c
tRNA-Lys(ctt)
rlmCD-fosX
tRNA-Leu(gag)
tRNA-Thr(ggt)
tRNA-Ser(cga)
tRNA-Arg(ccg)
comK
tRNA-Arg(tct)
***
***
0%
20%
40%
60%
80%
d
arsenic
cadmium
**
0%
10%
20%
30%
40%
50%
Singelton isolates (n=73)
Persistent clades (n=177)


a
pHC26_Listeria
pLM33_Listeria
pLR2-502_Listeria
pPIR00545_Listeria
pLM-F12_Listeria
G1
pCFSAN010068_Listeria
pLM58_Listeria
pMF4545_Listeria
86
pHC187_Listeria
pH7858_Listeria
G2
pLM5578_Listeria
pLMN1546_Listeria
100
pR479a_Listeria
pHC195-2_Listeria
pHC161_Listeria
100
G3
p967535_Listeria
pHC143_Listeria
pFDA550584-30_Listeria
pDE184_Bacillus
93
pX01_Bacillus
pBT9727_Bacillus
pSM19035_Streptococcus
pVEF1_Enterococcus
b
adhesin
srtA
relE
relB
repA
10 kbp
50 kbp
pHC143
55824 bp
20 kbp
40 kbp
30 kbp
LXG
pHC193
pHC385
pHC374
pHC375
pFDA550584-30
YA00079283
p967535


a
Sulfite exporter
AMR (vgaC)
AMR (lnuA)
BCR (emrC)
AMR (tetB)
G4
G5
86
100
100
100
0.03
pHC192_Listeria
pWikim-22_Lactobacillus
p54-2_Lactobacillus
pCPS49_Staphylococcus
pCFIAFB20130002_Listeria
p5459_5_49_Macrococcus
pN4D85_Enterococcus
p69892_Salmonella
pLmN12-0935_Listeria
pLMST6_Listeria
pMA67_Paenibacillus
pGL2_Lactococcus
b
tauE/safE
4500 bp
500 bp
4000 bp
1000 bp
repA
3500 bp
pHC192
4552 bp
pre
1500 bp
repB
3000 bp
2500 bp
2000 bp
hypothetical
c
pHC192
pHC192 & pHC195-2
no plasmid
Samples (n)
0
2
4
6
8
10
12
9
10
11
12
1
2
3
4
5
6
7
8
9
2014
2015


Resistome
Antimicrobial
Heavy metal
Listeria_17-04949_phage
Listeria_HC177_phage
Listeria_CFSAN007988_phage
Listeria_09CEB147LM_phage
Bacillus_7578-1_phage
Listeria_FDA574842-4_phage
Erysipelothrix_LV19_phage
Streptococcus_phage_Javan630
Blautia_PMF1_phage
Listeria_PNUSAL003131_phage
Erysipelothrix_phage_phi1605
Listeria_871181_phage
Listeria_HC134_phage
Listeria_HC258_phage
Streptococcus_phage_phi_m46.1
Streptococcus_phage_phiD12
Listeria_phage_A500
Listeria_phage_A006
Listeria_phage_B025
Listeria_phage_B054
Listeria_phage_A511
Listeria_phage_P70
81
86
95
100
100
100
100
100
72
100
89
Genus 1
Genera 2–6