Large-scale phenotypic and genomic characterization of *Listeria monocytogenes* susceptibility to quaternary ammonium compounds

Mirena Ivanova[1,#], Martin Laage Kragh[2], Judit Szarvas[1], Elif Seyda Tosun[1], Natacha Friis Holmud[1], Alexander Gmeiner[1], Corinne Amar[3], Claudia Guldimann[4], TuAnh N. Huynh[5], Renáta Karpíšková[6], Carmen Rota García[7], Diego Gomez[7], Eurydice Aboagye[8], Andrea Etter[8], Patrizia Centorame[9], Marina Torresi[9], Maria Elisabetta De Angelis[9], Francesco Pomilio[9], Anders Hauge Okholm[10], Yinghua Xiao[10], Sylvia Kleta[11], Stefanie Lueth[11], Ariane Pietzka[12], Jovana Kovacevic[13], Franco Pagotto[14], Kathrin Rychli[15], Irena Zdovc[16], Bojan Papić[16], Even Heir[17], Solveig Langsrud[17], Trond Møretrø[17], Roger Stephan[18], Phillip Brown[19], Sophia Kathariou[19], Taurai Tasara[18], Frank Aarestrup[1], Patrick Murigu Kamau Njage[1], Annette Fagerlund[17], Lisbeth Truelstrup Hansen[2], Pimlapas Leekitcharoenphon[1]

[1]Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Kgs Lyngby, Denmark
[2]Research Group for Food Microbiology and Hygiene, National Food Institute, Technical University of Denmark, Kgs Lyngby, Denmark
[3] UK Health Security Agency, Gastrointestinal Bacteria Reference Unit, London, UK
[4]Chair for Food Safety and Analytics, Ludwig-Maximilians-University Munich, Munich, Germany
[5]University of Wisconsin-Madison, Madison, WI, USA
[6]Masaryk University, Medical Faculty, Department of Public Health, Brno, Czech Republic
[7]University of Zaragoza, Zaragoza, Spain
[8]University of Vermont, Burlington, VT, USA
[9]Istituto Zooprofilattico Sperimentale dell 'Abruzzo e del Molise "Giuseppe Caporale", Teramo, Italy
[10]Arla Foods Amba, Arla Innovation Center, Aarhus N, Denmark
[11]German Federal Institute for Risk Assessment (BfR), National Reference Laboratory for *Listeria monocytogenes* (NRL-Lm), Berlin, Germany, Vienna, Austria
[12]Austrian Agency for Health and Food Safety (AGES), Institute of Medical Microbiology and Hygiene, National Reference Laboratory for *Listeria monocytogenes*, Graz, Austria
[13]Food Innovation Center, Oregon State University, Portland, OR, USA
[15]Listeriosis Reference Service, Food Directorate, Bureau of Microbial Hazards, Ottawa, ON, Canada
[16]Unit for Food Microbiology, Institute for Food Safety, Food Technology and Veterinary Public Health, University of Veterinary Medicine, Vienna, Austria
[17]Institute of Microbiology and Parasitology, Veterinary Faculty, University of Ljubljana, Slovenia
[18]Nofima, The Norwegian Institute of Food, Fisheries and Aquaculture Research, Ås, Norway
[19]Institute for Food Safety and Hygiene, Vetsuisse Faculty, University of Zürich, Zurich, Switzerland
[20]North Carolina State University, Raleigh, NC, USA

Running Head: *L. monocytogenes* susceptibility to QACs

Abstract word count: 250

[#]Address correspondence to Mirena Ivanova, e-mail: mirivan@food.dtu.dk

ABSTRACT

Listeria monocytogenes is a significant concern for the food industry due to its ability to persist in the food processing environment. Decreased susceptibility to disinfectants is one of the factors that contribute to the persistence of *L. monocytogenes*. The objective of this study was to explore the diversity of *L. monocytogenes* susceptibility to quaternary ammonium compounds (QACs) using 1,671 *L. monocytogenes* isolates. This was used to determine the phenotype-genotype concordance and characterize genomes of the QAC sensitive and tolerant isolates for stress resistance, virulence and plasmid replicon genes. Distribution of QAC tolerance genes among 37,897 publicly available *L. monocytogenes* genomes were also examined. The minimum inhibitory concentration to QACs was determined by the broth microdilution method and non-sequenced isolates (n=1,244) were whole genome sequenced. Genotype-phenotype concordance was 99% for benzalkonium chloride, DDAC and a commercial QAC based sanitizer. Prevalence of QAC tolerance genes was 23% and 28% in our *L. monocytogenes* collection and in the global dataset, respectively. *qacH* was the most prevalent gene in our collection (61%), with 19% prevalence in the global dataset. Notably, *bcrABC* was most common (72%) globally, while 25% in our collection. Prevalence of *emrC* and *emrE* was comparable in both datasets, 7% and 2%, respectively. Replicon genes, indicative of plasmid harborage, were detected in 44% of the isolates and associated with the QAC tolerant phenotype. The presented analysis is based on the biggest *L. monocytogenes* collection in diversity and quantity for characterization of the *L. monocytogenes* QAC tolerance at both phenotypic and genomic levels.

IMPORTANCE

Contamination of *Listeria monocytogenes* within the food processing environment is of concern to the food industry due to challenges in eradicating the pathogen once it becomes persistent in the environment. Genetic markers associated with increased tolerance to disinfectants have been identified, which alongside factors favor the persistence of *L. monocytogenes* in the production environment. By employing a comprehensive large-scale phenotypic testing and genomic analysis our study significantly

enhances the understanding of the prevalence of quaternary ammonium compound (QAC) tolerant *L. monocytogenes* and the genetic determinants associated with the increased tolerance. Furthermore, we report on the prevalence of QAC tolerance genes among 37,897 publicly available *L. monocytogenes* sequences and their distribution within clonal complexes, isolation sources and geographical locations. As the propagation of QAC tolerance showed not be evenly distributed globally this highlights that understanding the development of *L. monocytogenes* disinfectant tolerance can be monitored using publicly available WGS data.

INTRODUCTION

Listeria monocytogenes is a free-living saprophyte and an intracellular pathogen responsible for listeriosis, a serious foodborne infection with life-threatening complications in "at-risk populations" such as individuals with weakened immune system, elderly, neonates, as well as miscarriages in pregnant women (1). Unlike campylobacteriosis and salmonellosis, which are the two leading bacterial causes of foodborne illnesses, listeriosis has a low incidence rate but a high mortality rate (2). Contaminated ready-to-eat (RTE) foods have been identified as the major cause of several large outbreaks, including foods such as polony in South Africa (3), cold cuts in Canada and Denmark in 2008 and 2014, respectively (4,5), and cheese made with unpasteurized milk in USA (6).

Food contamination may occur in the food production facilities where *L. monocytogenes* can enter with e.g., raw materials, personnel or equipment and get established and persist for decades in the food processing environment (FPE) despite ongoing cleaning and disinfection programs (7), or it can be constantly introduced into FPEs with incoming raw materials. Both biotic (e.g., biofilm formation ability, growth at refrigeration temperatures, increased tolerance to sanitizers and disinfectants, heavy metals and desiccation, etc.) and abiotic (e.g., improper cleaning, biocide residuals, poor hygienic design, difficult to clean areas, etc.) factors are involved in successful survival and persistence of *L. monocytogenes* in FPEs, likely leading to contamination with *L. monocytogenes* clones with increased

95 persistence potential (8). Some *L. monocytogenes* clonal complexes (CCs) have been established as
96 successful persisters within FPEs, such as CC121 and CC9 that are also associated with hypovirulent
97 strains, while others have been strongly implicated with clinical sources and hypervirulence (CC1, CC2,
98 CC4 and CC6) (9)). Certain hypervirulent CCs (e.g., CC2) can also persist in FPE (Guidi et al., 2021).

99 To help control *L. monocytogenes* in FPEs, effective and robust cleaning and disinfection programs are
100 needed. Within cleaning and disinfection programs, three major biocide (disinfectants or sanitizers
101 according to EU and US definition, respectively) categories are widely applied in the food industry,
102 namely quaternary ammonium compounds (QACs) with main representatives benzalkonium chloride
103 (BC) and didecyldimethylammonium chloride (DDAC); peroxygens, (e.g., peracetic acid (PAA) and
104 hydrogen peroxide); and halogen-releasing agents, such as sodium hypochlorite (NaClO) (13–15). To
105 date, few specific genetic determinants associated with increased tolerance to oxidative agents such as
106 PAA and NaClO have been described, possibly due to the multiple modes of action of these agents
107 (14,16,17). In contrast, genetic determinants associated with increased tolerance to QACs have been
108 reported and include the efflux pumps genes *bcrABC, emrC*, *emrE* and *qacH* that provide *L.*
109 *monocytogenes* with increased survival in the presence of low concentrations of biocides (18–21). The
110 minimum inhibitory concentrations (MICs) of the QAC-based disinfectants towards *L. monocytogenes*
111 have been reported to vary between studies, ranging from 0.6 to >20 mg/L (17,22–25), while the
112 recommended concentrations in FPEs are 200–1000 mg/L facilities (27). Therefore, it has been
113 accepted that the term "tolerance" is more appropriate to describe the reduced *L. monocytogenes*
114 susceptibility to disinfectants than "resistance" (26).

115 Reduced *L. monocytogenes* susceptibility to food industry disinfectants, specifically to QACs, has been
116 the focus of many studies (17,22,27–29), including co-occurrence with antibiotic resistance genes
117 (22,30). However, the lack of a standardized disinfectant susceptibility method does not allow
118 comparison between the results reported by these studies. Variations in methods and factors such as
119 inoculum level, disinfectant dilutions and concentrations, media used, and incubation conditions can
120 contribute to variability of the MIC values without having a real connection to genetic variation. At the
121 same time, various efflux pump genes have been shown to play a role in increased tolerance to QAC

disinfectants in *L. monocytogenes*, including the *bcrABC* cassette, *emrE*, *emrC*, *qacH* (carried by Tn*6188*), *mdrL*, and *lde* genes (31–33). Here, we aimed at elucidating the genotype-phenotype concordance by performing disinfectant susceptibility testing of 1,671 *L. monocytogenes* isolates to two QAC substances, BC and DDAC, and a commercial QAC based disinfectant and to whole-genome sequence 1,244 isolates that have not been previously sequenced. We analyzed the genomes of the QAC sensitive and tolerant isolates for stress resistance genes, virulence factors and plasmid content. Additionally, the global dissemination of *bcrABC*, *qacH*, *emrC* and *emrE* genes, which were found exclusively in the tolerant isolates in this study, was examined in 37,897 publicly available raw sequencing data, deposited in the European Nucleotide Archive (ENA) as of April 2021. The distribution of QAC-associated genes in the global dataset among CCs, isolation sources and geographical locations was also investigated.

RESULTS

Characterization of the *L. monocytogenes* strain collection and QAC tolerance

To capture the full potential diversity of *L. monocytogenes* to QACs isolates from twenty collaborating laboratories were combined to create a diverse collection of isolates in relation to isolation year, environmental and geographical origin. In total 1,671 *L. monocytogenes* were included from >19 countries collected within a time span of 98 years, from 1924 to 2021 with >100 isolates from before 1990 (Fig. 1, Table S1). The isolates were of different origin with isolates sampled from human (n = 83; 5%), animal (n = 122; 7%), food (n = 839; 50%), FPEs (n = 488; 29%), feed (n = 4; <1%), natural environment (n = 32; 2%), farm environment (n = 66; 4%) and unknown sources (n = 22; 1%) (Fig. 1A). Isolates from different environments and countries were included as the stresses (environmental, antimicrobial, human etc.) faced by *L. monocytogenes* vary, which could affect the QAC tolerance and adaptation of the isolates from different niches and countries (Fig 1B). All *L. monocytogenes* isolates from collaborating partners were included in the strain collection and screened for QAC tolerance before whole genome sequencing. Minimum inhibitory concentration (MIC) towards BC were

148 determined for all isolates and 1283 isolates (77%) were defined as BC sensitive isolates with a MIC <
149 1.25 mg/L, while 388 isolates (23%) were defined as BC tolerant with MIC > 1.25 mg/L.
150
151 **Genomic characteristics and phylogenetic analysis of the *L. monocytogenes* isolates.**
152 Genomes of the 1,671 *L. monocytogenes* isolates ranged from 2.75 to 3.32 Mbp in size with a G+C
153 content of 37.7 to 38.3% (Table S2). The isolates were assigned to four phylogenetic lineages based on
154 the genome alignment of 2,146 core genes present in ≥ 99% of the isolates – lineage I (LI, n = 589,
155 35%), lineage II (LII, n = 1064, 64%), lineage III (LIII, n = 17, < 1%) and lineage IV (LIV, n = 1, <1%)
156 (Fig. S1). Within the isolate set, there were six serogroups (1/2a (n = 875, 52%), 1/2b (n = 262, 16%),
157 1/2c (n = 165, 10%), 4b (n = 319, 19%), IVb-v1 (n = 10, 0.6%), L (n = 19, 1%), atypical (n = 1, <1%)
158 and undetermined (n = 20, 1.2%)), and 75 CCs and 20 singleton sequence types (STs). Eleven CCs
159 predominated, accounting for 67% (n = 1,117) of all the examined isolates. The distribution of the
160 sources of isolates for the 11 most predominant CCs is presented in Fig. 2.
161
162 **Phenotype-genotype concordance of the *L. monocytogenes* QAC susceptibility.**
163 The genomes of the 1,671 *L. monocytogenes* isolates were screened for the presence of QAC tolerance
164 genes (i.e., *bcrABC*, *emrC*, *emrE* and *qacH*) and the genotypes were associated with the respective
165 phenotype (BC MIC values). The genes *bcrABC*, *emrC*, *emrE* and *qacH* were detected only in isolates
166 with MIC ≥ 1.25 mg/L. Based on that, an MIC cut-off of ≥ 1.25 mg/L for BC tolerance was established
167 (Fig. 3) and isolates were classified as BC sensitive (77%, n=1283) having MIC < 1.25 mg/L and BC
168 tolerant (23%, n=388) with MIC ≥ 1.25 mg/L. Twenty (5%) of the BC tolerant isolates did not carry
169 any of the four QAC tolerance genes associated with the tolerant phenotype. Moreover, no BC sensitive
170 isolates with a QAC tolerance gene, nor BC tolerant isolates carrying more than one QAC tolerance
171 gene were found. A randomly selected subset of isolates was tested for sensitivity to DDAC (n=247)
172 and Mida San 360 OM (cQAC) (n=155), respectively. There was a clear distribution of the MIC values
173 for DDAC and cQAC and the cut-offs for tolerance were drawn at MIC > 0.4 mg/L and ≥ 0.63 mg/L
174 for DDAC and cQAC, respectively, based on presence/absence of the four QAC tolerance genes (Fig.

3). All BC tolerant isolates were also tolerant to DDAC and cQAC, except one isolate (414a-1), which was tolerant to BC but not to cQAC.

The most prevalent QAC tolerance gene found within our isolate collection was *qacH* (61% of the BC tolerant isolates; n=237), followed by *bcrABC* (25%, n=96), *emrC* (7%, n=27) and *emrE* (2%, n=8) (Fig. S2A). The majority of the *qacH*-harbouring isolates belonged to CC121 (83%) and CC9 (12%). The isolates carrying *bcrABC* genes most commonly belonged to CC9 (30%), CC5 (26%) and CC321 (16%). The *emrC*-harbouring isolates belonged to 11 CCs, and among them CC6 (41%), CC14 (11%) and CC403 (11%) were the most prevalent CCs. Meanwhile, *emrE* was detected among CC8 isolates. Notably, the BC-tolerant isolates without an identified QAC tolerance gene were genetically diverse, belonging to 12 different CCs (Fig. 4A). Overall, the majority of isolates harboring QAC tolerance genes belonged to LII (85%), with 60% and 20% belonging to CC121 and CC9, respectively. QAC tolerance genes were not detected among the LIII and LIV isolates. The majority of the QAC tolerance genes were found in isolates from food (56%) and FPE (38%), and only 2.6% were from clinical sources (Fig. 4B). All *bcrABC* and *emrC* genes were detected on plasmid contigs. The majority of the *qacH* genes (n = 237) were located on contigs with no identified replicon gene and with lengths greater than the largest identified *Listeria spp*. plasmid (>152 kbp, CP022021.1), assuming chromosomal origin.

Identification of plasmids among the QAC tolerant and sensitive isolates

Genetic determinants of the QAC tolerance are often plasmid-borne (e.g., *bcrABC*, *emrC*), therefore an analysis to identify plasmid content in the tolerant and sensitive isolates was performed. Overall, 44% of the isolates (n=735) were determined to carry plasmids. Of these, 712 carried one replicon gene, 24 carried two replicon genes, and three carried plasmids with unknown replication system. Ninety-seven percent of the plasmid-harboring isolates (714/735) carried *repA* genes. In total 728 *repA* genes were subjected to a phylogenetic analysis, presented in Fig 5A. The majority of the *repA* genes belonged to phylogenetic groups G1 (n=326; 20%) and G2 (n=390; 23%), and 10 genes were allocated to G4. Two *repA* variants clustered independently of any known RepA group and were most closely related to G3; they were named G13 (Fig. 5A). These two G13 isolates (SKB398 and SKB102) belonged to CC912

and ST1365, originating from black bears. Additionally, 20 and three of the 735 isolates with identified plasmids carried *repB* genes and plasmids with an unknown replication system, respectively. The *repB* genes were found on 4 kb contigs, except for isolates N195 and LIS08. In these two isolates, the *repB* genes were co-located with *repA* genes on plasmid contigs that were of length 53 and 81 kbp, respectively. The majority of the *repA* genes were identified in isolates from food (55%) and FPE (39%), while only 3.4% and 0.8% were seen in clinical, and farm and natural environment isolates, respectively (Fig. 5A). Food and FPE isolates were evenly distributed among the plasmid phylogenetic groups in this study. Prevalence of isolates determined as sensitive or tolerant to QAC were in addition not significantly more associated with food or FPE as only significant differences were seen for human clinical samples where QAC sensitive isolates were more prevalent (Fig. 5B). It is worth mentioning that 91% of the *repA* genes in the phylogenetic group G2 belonged to LII isolates, while G1 was equally distributed within LI and LII. Notably, 98% of the CC121 isolates carried *repA* genes belonging to G2, while G1 was the predominant phylogenetic group (88%) in the CC9 isolates. Similarly, other CCs with high frequency of *repA* had different distributions of the phylogenetic groups with CC3 (91% G1), CC5 (50% G1, 50% G2) and CC8 (96% G2. *repB* genes showed the highest nucleotide identity and coverage to pLmN12-0935 (CP038643.1). Two of the three plasmid contigs with unknown replication systems were most similar to plasmid pLIS55 (MZ151539.1), while the *repB* gene in isolate CDL77 had no match to any known sequence in NCBI. The majority of the isolates which carried simultaneously two *repA* genes came from FPEs and belonged to CC3, CC5, CC8 and CC89.

Further, blastn screening using a custom database of *L. monocytogenes* virulence and stress resistance genes against the contigs carrying replicon genes (n = 735) identified co-location of *qacH* and *repA* genes in six of the G4 plasmid group isolates. In these isolates, *qacH* had a different genetic context from the previously described Tn*6188* transposon structure, consisting of *tetR* and *qacH* genes, and an upstream transcriptional regulator *mutR* (Fig. S3).

All *bcrABC* and *emrC* genes were detected on plasmid contigs as determined by their co-location with other plasmid-associated genes. While the *bcrABC* gene was located on contigs of varying sizes (3.8 – 91.2 kbps), the *emrC* was found on small contigs (4.4 - 4.7 kbps) with higher nucleotide coverage

compared to the coverage of the other contigs in the assemblies. The *qacH* genes (n=237) were located on contigs with no identified replicon gene and with lengths greater than the largest identified *Listeria* spp. plasmid (>152 kbp, CP022021.1), suggesting chromosomal origin. *emrE* as part of the LGI-1 was located on the chromosome in all isolates as previously described (18).

The LIPI-1 genes, *plcA*, *hly*, *mpl*, *plcB* and a truncated *actA* gene, associated with increased *L. monocytogenes* virulence, were also found co-located with a *repA* gene in two isolates (ERR1432982 and ERR1432994). Genes encoding for resistance to arsenic (*ACR3*, *lmo2230*, *arsA*, *arsD*, *arsR*) (n = 8), cadmium resistance genes *cadA1C1* (n= 522) and *cadA2C2* (n = 45), heat resistance gene *clpL* (n = 348), genes for copper resistance *copB* (n = 198) and *copL* (n = 54), mobile genetic element consisting of *gbuC* (n = 173), NiCo riboswitch (n = 168) and *npr* (n = 173), *tmr* (n=20), *mco* (n = 197) and the mercury resistance operon (n = 2) were also co-carried with a *repA* gene (Table S6).

Genetic organisation of the *bcrABC*-harbouring plasmid contigs.

The *bcrABC* gene was detected in 96 isolates representing 11 CCs and four isolation sources (Fig. S4). Using multiple sequence alignment, the *bcrABC*-carrying contigs were grouped into eight structures (Fig. S5). Structures 1 and 1a, detected only in CC9 isolates, were identical except for the presence of the genetic module NiCo riboswitch-*gbuC*-*npr* that was seen in structure 1 but absent from 1a. Structures 2, 3 and 4 had a common backbone consisting of *ltrC*, *bcrABC*, recombinase, glyoxalase, *tmr*, *qorB* and *ravA*. While structure 2 was detected only in the CC5 231a1 isolate, structures 3 and 4 were distributed among various CCs. A shared feature of the structures 5, 5a and 6, found in three non-clonal *L. monocytogenes* isolates from CC1 (LI),) and CC31 and CC204 (LII), was the presence of the mercury resistance (*mer*) operon. Blastn search in NCBI of the *bcrABC* contigs of structures 5 and 5a identified similarity to pLMR479a (HG813248.1) from a CC8 isolate from smoked salmon, with 84% coverage, lacking only the *bcrABC* cassette and the *mer* operon, suggesting that these genetic elements could horizontally be transferred together. The isolate with structure 6, N21-0102 isolated from poultry in Switzerland, had 99.89-100% identity and over 100% coverage to the *Listeria innocua* plasmids CP095722.1 and CP095729.1, found in isolates from meat in Jamaica. The sequences from the three CC31 isolates from milk in Switzerland with structure 7 had 100% coverage and 100% identity to a *L.*

innocua pLIS35 plasmid (MZ127844.1), isolated from a food contact surface swab in Poland. Structure 8 was detected only in CC321 isolates from North America, in which the *bcrABC* cassette was carried on a composite transposon associated with IS*1216* (Fig. S5). This shows that there is a wide diversity of genetic contexts surrounding the *bcrABC* operon, although it seems that the cassette is always located on a plasmid.

Identification of virulence and stress resistance genes among QAC tolerant and sensitive isolates

The carriage of virulence factors and stress resistance genes have previously been linked to specific CCs, environments and QAC phenotypes in small- and large-scale screening studies. To validate and extent these observations with a larger isolate collection a variety of virulence and stress resistance genes were identified in the 1,671 *L. monocytogenes* isolates.

Ten different mutations (including four newly detected in this study) and two internal deletions in the *inlA* gene, leading to premature stop codons (PMSCs) and proteins with various lengths, were found in 21% of the isolates (Table S5). The majority of these isolates belonged to lineage II (97%) and specifically CC121 (61%) and CC9 (30%). The most common mutation (C -> T at nucleotide 1474 leading to a protein with 491 aa), accounting for 61% of the total mutations, was only seen in CC121 isolates (34). All but five CC121 isolates (98%) carried an *inlA* PMSC mutation. Truncated *inlB* genes, with deletion of amino acids from 356/357 to 630 (leading to a protein of 356/357 aa), were detected in 19 CC5 isolates originating from FPEs and food samples. PMSCs in *inlB* gene resulting in shorter protein sequences were identified in five food isolates belonging to CC1 and CC8 (data not shown). All isolates with truncated *inlB* genes harboured intact *inlA* genes. The *actA* gene, located on LIPI-1, was truncated in all isolates from 20 CCs, including CC1, CC121, CC31, CC4, CC379 and CC88 (Fig. 6A).

The cadmium resistance genes *cadA1C1* and *cadA2C2*, located on plasmids with various sizes, were the predominant cadmium resistance genes, while *cadA3C3*, *cadA4C4*, *cadA5C5*, *cadA6aC6a*, *cadA6bC6b* were less frequently detected. Copper (*mco*, *copB*, *copY*, *zosA*, *ctpA*), arsenic (*arsA1D1R1D2R2A2B1B2* and *arsCBADR*) and mercury (*merR1ETPAR2B1*) resistance genes were unevenly distributed (Fig. 6A). The oxidative stress, heavy metal, salt and acid resistance cassette NiCo

284 riboswitch-*npr-gbuC*-like, were transferred together by the same mobile genetic element and found in
285 CCs that carry plasmid genes. The environmental stress islet SSI-1 was present in 45% of the CCs,
286 while SSI-2 was less prevalent (9%). The gene *comK*, a common hotspot for prophages in *L.*
287 *monocytogenes*, was truncated in 82% of the isolates, and a truncation was found in all but three CCs.
288 Two insertion sites were identified, TAA-TAAAA (at nucleotides 212 and 214) seen among 25% of the
289 isolates, and GGA (at nucleotide 190) seen in 57% of the isolates. Additionally, 98.6% of the isolates
290 contained full or partial dsDNA phage sequences and 52.3% had full phage sequences as identified by
291 VirSorter with minimum length of 1500 bp. The *bapL* gene, associated with biofilm formation, was
292 present in CCs from LII only. The LIPI-1 pathogenicity island, that carries the master regulator for
293 virulence transcription (*prfA*), and cluster of virulence genes used to escape from vacuolar
294 compartments (*hly*, *plcA*; *actA*; *mpl*, *plcB*, and *orfX*), was present in all isolates. The pathogenicity
295 islands LIPI-3 and LIPI-4 were present in the majority of the isolates in CCs from LI and were absent
296 in isolates from LII. CC218, CC217, CC506, CC1258 and CC639 harboured all LIPI-1, -3, and -4 genes,
297 indicative of potential to exhibit higher virulence level. The isolates from LIII and LIV did not carry
298 any of the screened stress resistance genes (Fig. 6A, Table S7). When isolated were grouped as sensitive
299 and QAC tolerant there were differences in the prevalence of several stress resistance genes between
300 the groups. When adjusting for the higher number of sensitive isolates in the isolate collection there
301 were a significantly (P<0.05) higher prevalence in QAC tolerant isolates of genes conferring resistance
302 to cadmium, and heat as well as SSI-2 (Fig. 6B).
303
304 **Global distribution of QAC tolerance genes in *L. monocytogenes***
305 The raw sequencing data from 39,196 *L. monocytogenes* isolates deposited in the European Nucleotide
306 Archive (ENA) were screened for the presence of QAC-tolerance genes *bcrABC*, *emrE*, *emrC* and *qacH*
307 (Table S8). Of the analyzed sequencing data, 10,953 (28%) carried one or more QAC tolerance genes.
308 *bcrABC* was the most abundant gene, present in 72% of the QAC tolerance gene positive isolates,
309 followed by *qacH* (19%), *emrC* (7%) and *emrE* (2%) (Fig. S2B). *qacH* had the highest nucleotide
310 variability among the four QAC tolerance genes, and nine *qacH* variants with nucleotide identity above
311 90% were detected in the global dataset based on analysis of raw reads using KMA (Fig. S6). Notably,

23 isolates carried simultaneously two QAC tolerance genes. Since several large-scale studies using in-house sequencing data (35,36), including this study, detected no more than one QAC tolerance gene in a single *L. monocytogenes* genome, the raw sequencing data of the isolates carrying two QAC tolerance genes were assembled, re-screened by blastn and sub-typed by MLST. Subsequently, 7/23 isolates carried only one QAC gene. Of the 16 isolates that still carried two QAC genes, two alleles of an MLST gene were detected in three isolates, suggesting contamination of the sequencing reads. For the remaining 13 genomes, in which *bcrABC* and *qacH* (seven isolates), *bcrABC* and *emrC* (one isolate), and *qacH* and *emrC* (four isolates) were simultaneously detected, re-isolation and re-sequencing could reveal if two QAC tolerance genes could in fact be carried by the same *L. monocytogenes* isolate (Table S9). All QAC tolerance genes were present in *L. monocytogenes* isolates from LI and LII, with the exception of one LIII (ST1142) isolate (SRR8223359) that harbored *bcrABC* and recovered from a FPE isolate in the United States.

There were differences in the dissemination of the QAC tolerance genes among countries/continents and CCs. The *bcrABC* cassette was widely distributed in the United States, while *qacH* and *emrC* were mainly associated with Europe, and *emrE* had highest occurrence in Australia/Oceania (Fig. 7, Fig. S7, Fig. S8). The *emrC, emrE* and *qacH* were associated with CC8, CC6 and CC121, respectively, while the *bcrABC* cassette was distributed among several CCs, having the highest occurrence in CC5 followed by CC321, CC155, CC9 and CC7 (Fig. S9). All genes were associated with FPE or food and feed sources, except *emrC*, which was also associated with clinical sources (Fig. S10,).

DISCUSSION

The adaptation and persistence of *L. monocytogenes* in FPEs is a complex issue, often dependent on a combination of factors, such as biofilm formation capacity of different strains, improper equipment design, the presence of organic matter, insufficient removal of disinfectants during cleaning and disinfection and presence of genes conferring tolerance to various environmental stressors, among others (8,37). The QAC genes conferring tolerance to low concentrations of QACs can in combination

with the above events contribute to persistence of *L. monocytogenes* as isolates harboring such genes have potentially higher survivability in the presence of low disinfectant concentrations in niches and often carry additional stress genes as seen in present study (Fig. 6B).

L. monocytogenes susceptibility to food industry disinfectants, specifically QACs, which are applied in the food industry, has been extensively examined (17,22,27,28,31,38). Most studies have included either a small number of isolates or isolates restricted to certain geographical areas, or focused on isolates recovered from specific environments, e.g., meat (24,39) or pork processing environments (40). The 1,671 *L. monocytogenes* isolates in this study were collected to represent diverse origin sources and geographical locations. Of all the isolates, 74% were selected solely based on the available metadata, therefore minimizing the bias of choosing isolates with certain genetic characteristics. Yet, *L. monocytogenes* isolates originating from FPEs and food prevailed in our collection, which could be based on sampling bias, as FPEs and foods make up the important niches for studying *L. monocytogenes* transmission routes (35,36,41,42).

Unlike for antibiotics, the lack of a harmonized assay for disinfectant susceptibility testing affects the comparability of the phenotypic results produced by different studies. The previously reported QAC MIC values vary significantly between studies due to differences in the biocide and/or inoculum concentrations, variations in media and assay conditions (27,28,29,54,55). For instance, *L. monocytogenes* exhibits different physiological traits at lower temperatures (e.g., 15ºC) as compared to 37ºC, such as the different types of motility and having different membrane fatty acid composition (43). An important aspect of our large-scale study is the consistency and reliability of the phenotypic data obtained using the same antimicrobial susceptibility method and conditions, which allowed us to establish a cut-off for tolerance based on the presence/absence of known genetic determinants. The observed distribution of the QAC MIC values allowed for a clear division of the sensitive and tolerant phenotypes. The high concordance between a tolerant phenotype and the presence of a QAC tolerance gene (99%) means that the phenotype of QAC tolerance can be predicted *in-silico* if the genotype is known without the need for conducting wet lab phenotypic testing.

The few tolerant *L. monocytogenes* isolates with no known QAC genetic determinants belonged to various CCs. This suggests a spread of mobilizable genetic elements distributed among diverse CCs, or that these isolates may be tolerant due to altered expression regulation of efflux pump genes, such as *mdrL* or *lde*. Other recent studies also reported QAC tolerant *L. monocytogenes* isolates with no known QAC tolerance genes, e.g., a food environment ST155 isolate from the United States (44) and two isolates from German food production facilities (22). Genome-wide association studies (GWAS) have the power to identify novel genetic determinants associated with disinfectant tolerant phenotypes using large collections of isolates (17). Among the pangenome orthologous groups (POGs), in the study of Palma et al. (17), Tn*6188_qacH* and a cluster of 14 POGs, a part of a prophage region, were strongly associated with BC tolerance (17). However, such findings have not yet been confirmed phenotypically. It is also worth mentioning that the general efflux pumps genes previously associated with an increased QAC tolerance (e.g., *mdrL* and *lde*) were detected in both QAC tolerant and sensitive isolates. The role of these genes may be explained by gene expression regulation differences between resistant and sensitive strains. It is possible that genetic variation in these genes cannot explain different QAC tolerance phenotypes, but rather that the differential expression regulation may be responsible for these effects.

Gene variants or specific genetic organizations located on plasmids, or a chromosome have been reported to have an effect on the disinfectant inhibitory concentrations. For instance, Dutta et al. (57) demonstrated that *L. monocytogenes* that carried *bcrABC* in different genetic locations, such as in categories VI (presumed chromosomal origin) and VII (composite transposon), exhibited lower MICs (30 mg/L) compared to categories I-V (MIC = 40 mg/L), which have been associated with plasmid origin. In the present study all *brcABC* carrying isolates showed the same tolerance to BC. Møretrø et al. (27) reported a variation among the tolerant phenotypes due to a non-synonymous mutation in the *qacH* gene. QacH harboring ^{42}Ser exhibited higher MIC values than those carrying ^{42}Cys. In this study, ^{42}Ser was detected in QacH variants 4, 6 and 7, but no difference in the MIC values was observed. If further in-depth phenotypic differences could be investigated with a more sensitive method such as growth curve analysis (44).

Association between *qacH* location and its sequence has not previously been discussed in the literature, except for the report of a *qacH* homolog in a *L. monocytogenes* isolate from FPE in a German production facility, which lacked the other Tn*6188* genes (NG_076646.1 of isolate 16-LI00532-0) (22) and a recent report in a clinical isolate from Norway (35). In our study, we detected a *qacH* homolog in RepA G4 plasmid, with 100% nucleotide similarity to NG_076646.1 and 91-92% nucleotide identity to the chromosomal *qacH* (HG329628.1). The plasmid-borne *qacH* homolog was associated with a *tetR* gene, with 72% nucleotide identity to the chromosomal *tetR* located immediately upstream of *qacH* . Similarly, Chmielowska et al. (45) and Schmitz-Esser et al. (12) observed that the RepA G4 plasmids, although rarely found in their studies, encoded a putative novel QAC transporter. However, when we used the broth microdilution method to test the isolates, differences in MIC values to BC there were not difference in the tolerance of isolates with plasmid-borne *qacH* and the *qacH* on Tn*6188*. Additionally, it was seen that all RepA G4 plasmids had a complete transfer module, which suggests that they can be mobilized by a conjugative transfer. However, despite potential mobility, the occurrence of the RepA G4 plasmids tends to be low. In our study only seven isolates carried these plasmids, with similar low numbers observed by Chmielowska et al. (45) (n = 4), Schmitz-Esser et al. (12) (n = 8), one isolate in a study by Fagerlund et al. (35) (n = 1). Notably, five of the strains carrying the novel *qacH* homolog gene in the present study are from FPEs in three European countries (Italy, Spain and UK), isolated between 2015 and 2020.

Regarding the *bcrABC* cassette structures identified in this study, the *bcrABC* genetic environment consisted of genes encoding recombinase, glyoxalase and oxidoreductase, with exception of CC321 isolates which carried *bcrABC* on a composite transposon flanked by two IS*1216* elements. It is important to note that structure 8 in this study (Fig. S5) corresponds to category 7 in Dutta et al. (46). Contrary to Dutta et al., who observed both plasmid- and chromosome-located *bcrABC* cassettes, all structures in this study carried either the *cadC1* gene conferring resistance to cadmium, the mercury resistance operon, the stress response module NiCo riboswitch-*gbuC-npr* and/or the *clpL* gene, all suggesting plasmid origin of the contigs. None of the *bcrABC* genes in this study were associated with chromosomal origin. As all *bcrABC* structures were found to co-occur with other genetic elements

conferring stress resistance, their impact on the persistence potential of *L. monocytogenes* in FPEs should be further investigated. Castro et al. (32) found that mobile genetic elements harbouring resistance genes for arsenic and cadmium were significantly more prevalent among persistent *L. monocytogenes* isolates from farm environments than in non-persistent isolates. Conversely, persistent isolates have been found to carry mutations associated with attenuated virulence, including a truncated *inlA* gene (47), which is in line with our findings of 98% of the FPE CC121 isolates possessing a PMSC in the *inlA* gene. Additionally, co-occurrence of QAC tolerance genes with other stress resistance genes, such as metal and antibiotic resistance genes, can contribute to their adaptation and persistence in FPEs. If exerted on disinfectant selective pressure, a response together with tolerance to other stressors may occur. It is common for BC tolerant isolates from food and food processing facilities to be cadmium resistant (48). However, our results also showed that in addition to cadmium resistance genes, the heat resistance gene *clpL*, plasmid replicon genes and the SSI-2 genes are strongly associated with QAC tolerant phenotype (Fig. 5A, Fig 6B). The SSI-2, involved in alkaline and oxidative stress responses, has been described to predominate in the hypovirulent CC121 isolates, mostly associated with FPE, and also carrying the *qacH* gene (49).

While not highly prevalent in our isolate collection, the high global prevalence of the *bcrABC* cassette and its predominance in phylogenetically distant CCs, may be due to its location on plasmids with various structures that allow for transfer between many different CCs (50). Evolutionary, the high prevalence could be due to an earlier acquisition of the cassette in the *Listeria* genome compared to the other QAC *L. monocytogenes* genes. Another possible explanation for the *bcrABC* predominance in food production facilities is their co-occurrence with other stress resistance genes, which can aid the survival of the isolates that carry them. It is also possible that they have better genomic compatibility with CCs that are likely to get established as persisters in FPEs, such as CC121, CC9, or in hypovirulent genomes that carry PMSCs, or other signature genetic characteristics associated with FPE.

On the other hand, the global differences in the gene prevalence as clearly seen for *emrE* (Fig. S7, S8) with much higher prevalence rates in some countries could be related to trading patterns between different countries/continents (e.g., US and Oceania/Australia) as contamination of raw materials with

specific strains in one country can affect the *L. monocytogenes* prevalence rates in another country, as well as practices for application and use of disinfectants. The higher sequencing rates and public availability of bacterial genomes in US could partially explain the increased finding of *brcABC* among the global dataset in this study (Fig. S8).

The majority of the QAC tolerance genes found in the global dataset were associated with food production environment and food and feed samples, as previously reported by other studies (50,51), with the exception of *emrC* whose predominance in the production environment was merely 1.6%, compared to 44.2% seen in clinical samples. The latter is in agreement with the studies of Daeschel et al. (51) and He et al. (50), who examined the prevalence of QAC tolerance genes among *L. monocytogenes* isolates from US processing facilities and produce processing environments, respectively, and did not detect *emrC*.

Despite being associated with one of the most prevalent CCs within the FPEs, CC121 and CC9, the negligible detection of the *qacH* gene compared to the *bcrABC* cassette in the North American *L. monocytogenes* isolates, could possibly be due to a lower CC121 distribution in this geographic region or differences in the use of disinfectants in different countries/continents. It appears that the QAC-tolerant *L. monocytogenes*, which persist in the FPE in US, are non-CC121 *bcrABC*-harbouring isolates, such as CC5, CC321, CC155, CC7, CC9 and CC199 (35,41,46). Other large-scale studies comprising of *L. monocytogenes* isolates from Europe, report *qacH* as the predominant QAC tolerance gene, 18.8% in a French study (52) and 18.9% in an EU-study (53). This also leads to the idea that the distribution of QAC tolerance genes in different geographic regions is related to the predominance of CCs (51).

Despite its widespread distribution among geographic regions and CCs, *bcrABC* has never been associated with CC121, probably due to the high prevalence of *qacH* in this CC or possible incompatibility of two QAC tolerance genes in one *L. monocytogenes* genome. We also found 17 isolates that carried two different QAC tolerance genes, specifically *bcrABC* and *qacH*, *bcrABC* and *emrC*, and *qacH* and *emrC*. Kropac et al. (54) showed that *emrC* could be transformed in a *L. monocytogenes* CC2 genome carrying chromosomal *qacH*. The authors even reported slight increase in

the BC MIC value of the transformant compared to the MIC value of the recipient cell and hypothesized that acquiring additional BC efflux pump gene can have an additive effect on the tolerance level. It is unknown, however, if in the nature it is advantageous for *L. monocytogenes* to harbour two QAC tolerance genes, especially the *emrC* gene which is located on a high copy plasmid. Similarly, Daeschel et al. (51) reported that 50% of the isolates screened in their study contained one or more of *bcrABC*, *emrE* or *qacH* genes, however, harbourage of more than one QAC tolerance gene in a single *L. monocytogenes* genome similarly to our study could be explained by their dataset, consisting of sequencing data downloaded from NCBI Sequence Read Archive that could contain contaminating reads.

In conclusion, this study showed that QAC tolerance can be *in-silico* predicted as genotype-phenotype concordance was 99% for QAC-based disinfectants for the four assessed QAC tolerance genes which all conferred the same level of tolerance to QACs as seen in the MICs. The propagation of QAC tolerance genes is however not evenly distributed globally with differences between continents and countries which showcases that understanding the development of *L. monocytogenes* disinfectant sensitivity and tolerance can be monitored using publicly available WGS data as different QAC gene variants and genetic contexts exist.

MATERIALS AND METHODS

Listeria monocytogenes isolates.

A total of 1671 *L. monocytogenes* isolates recovered from human (n = 83; 5%), animal (n = 122; 7%), food (n = 839; 50%), FPE (n = 488; 29%), feed (n = 4; <1%), natural environment (n = 32; 2%), farm environment (n = 66; 4%) and unknown sources (n = 22; 1%) were included in this study. Additionally, 14 *L. monocytogenes* isolates (<1%) with unspecified clinical origin and one *L. monocytogenes* isolate (<1%) with unspecified environmental origin were also included. The isolates originated from 19 countries and were collected within a time span of 98 years, from 1924 to 2021 (Fig. 1C, Table S1).

Determination of the *Listeria monocytogenes* minimum inhibitory concentrations to QACs.

The minimum inhibitory concentrations (MIC) of *L. monocytogenes* to two pure biocide substances, benzalkonium chloride (BC) (500 g/L, Thermo Fisher, Kandel, Germany), didecyldimethylammonium chloride (DDAC) (500 g/L, Sigma-Aldrich, Denmark), and a commercial disinfectant - Mida SAN 360 OM (10-15% 2-methoxymethylethoxy propanol, 3-5% didecyl dimethyl ammonium chloride, < 3% 1,2-ethanediol; Christeyns, Denmark) were tested by an in-house optimized broth microdilution assay according to Wiegand et al. (55). The *L. monocytogenes* isolates were streaked from -80°C on trypticase soy agar plates (TSA; 40 g/L) and incubated at 37°C overnight. Three single colonies per isolate were transferred to 0.1x trypticase soy broth (TSB; 3 g/L) and cultured at 15°C for 48 h under stationary conditions until the optical density at 620 nm (OD_{620}) measured ~0.1, corresponding to 10^8 CFU/mL. The final inoculum concentration was 10^5 CFU/mL per well. The range of the disinfectant concentrations was selected according to previously reported MIC values (17) using two-fold dilutions. A positive control consisting of inoculated 0.1x TSB and a negative control consisting of sterile 0.1x TSB were included in each 96-well plate. The plates were sealed with adhesive film (ThermoFisher, Denmark) and incubated at 15°C for 48 h. The MIC, defined as the lowest biocide concentration at which the *L. monocytogenes* growth was inhibited, was determined by measuring the *L. monocytogenes* OD_{620} by a Multiscan FC Microplate Reader (Thermo Scientific, Denmark). The threshold for growth was set at $OD_{620} \geq 0.08$, which is 60-65% of OD_{620} of the positive control and 200% of the negative control. The experiment was performed in two independent biological replicates with three technical replicates each. The result was considered valid if two out of the three technical replicates had identical MIC values. One two-fold MIC variation between biological replicates was considered acceptable and the higher MIC value was reported as the final result, unless the two-fold dilution difference was at the cut-off for tolerance (MIC ≥ 1.25 mg/L). In this case, the test was repeated a third time.

DNA extraction and whole-genome sequencing.

For whole-genome sequencing (WGS), as part of the current study, the *L. monocytogenes* isolates (n = 1,244) were grown on TSA at 37°C, overnight. A single colony per strain was transferred to 1.8 ml TSB

and grown at 37°C overnight. Genomic DNA was extracted by the DNeasy Blood and Tissue Kit (Qiagen, Denmark) following the manufacturer's recommendations except that the DNA was eluted in 10 mM Tris-HCl (pH = 8.5) (BioNordika, Denmark). The DNA concentration was measured using Quant-iT dsDNA high sensitivity kit (Invitrogen, Denmark) by VICTOR X2 Multilabel Microplate Reader (Spectralab Scientific Inc.). Sequencing libraries were constructed using the Nextera XT Library Prep Kit (Illumina, San Diego, CA, USA), normalized and denatured for loading in a NextSeq 500/550 Mid Output v2.5 Kit (300 cycles) (Illumina), and pair-end sequenced on a NextSeq 500 platform (Illumina). The remaining 427 *L. monocytogenes* isolates were previously sequenced (Table S1).

Assembly, species identification and *in silico* sub-typing.

The raw sequencing data of the 1,671 isolates were processed with the FoodQCpipeline v1.6 (https://bitbucket.org/genomicepidemiology/foodqcpipeline/src/master/), which uses bbduk2 from bbtools (Bushnell B. sourceforge.net/projects/bbmap/) for trimming and SPAdes (56) for genome assembly. Quality control of the sequencing reads was performed before and after trimming by FastQC v0.11.5 (57). Quality of the assemblies was assessed by Quast v4.5 (58) and thresholds for number of contigs (≤ 300 contigs) and genome size (3 Mb ± 0.5 Mb) were established. *In silico* species identification was performed using KmerFinder v2.0 (59). Assemblies were submitted to the BIGSdb-*L. monocytogenes* Pasteur MLST database (https://bigsdb.pasteur.fr/listeria/) for sub-typing (Table S1).

Core-genome MLST and phylogenetic analysis.

The assemblies were annotated with Prokka v1.14.6 referencing the genus *Listeria* and the species *monocytogenes* (60). Core-genome alignment was generated by Roary v3.13.0 using the .gff files from Prokka as input, with 95% blastp identity threshold and paralog splitting (-s) disabled (61) to prevent presumed paralogous genes from being split into different gene groups. The core-gene alignment was trimmed with trimAI (62) and option -gappyout to decide optimal thresholds based on the gap percentage count over the whole alignment. A maximum-likelihood (ML) phylogenetic tree was inferred using IQ-TREE v2.1.3 with the GTR+G nucleotide substitution model and 1000 bootstrap

replicates (--ufboot 1000) and mid-point rooted. The trees were visualized and annotated in iTOL (63) or ggtreeExtra R package (64).

Identification of plasmids.

The plasmid replicon gene screening was carried out by BLAST+ v2.2.31 (65) with a database consisting of representative *repA* genes from each of the eleven phylogenetic groups identified by Chmielowska et al. (2021) (66), the *repA* G12 by Fagerlund et al. (2022) (35), the *repB* genes from the small *Listeria* spp. plasmids, and the complete sequences of the plasmids with an unknown replication system (Table S3). The blastn was performed with the fasta files as queries against the *repA* database with minimum identification (--minid) and coverage (--mincov) thresholds of 80 and 95, respectively, in ABRicate v1.0.1 (Seemann T, Abricate, Github https://github.com/tseemann/abricate). The *repA*-harboring contigs were extracted from the assemblies using *awk* and the *repA* gene sequences subsequently extracted by getfasta in bedtools v2.30.0 (67). Translated *repA* gene sequences were aligned with MAFFT v1.5.0 and used to infer maximum likelihood tree in IQ-TREE using the Le and Gascuel (LG) amino-acid substitution model and ultrafast bootstrapping (--ufboot 1000). Reference RepA sequences from each of the twelve RepA groups (G1-G12) (35,45) were included in the tree. The tree was visualized and annotated in iTOL.

Genetic organization of the *bcrABC*-harbouring plasmids.

Plasmid contigs carrying *bcrABC* were annotated by Prokka with --compliant option. The contigs were aligned in Geneious Prime 2023.0.4 by MAUVE and MAFFT and grouped according to their genetic organization. From each group of identical contigs (indels or SNPs were ignored), one contig was chosen as a representative and the .gbf from Prokka uploaded to Clinker (68) for pairwise alignment and visualization. Plasmid contigs from each *bcrABC* category were aligned to a custom database consisting of complete *Listeria* spp. plasmids carrying the *bcrABC* cassette (45,69). Additionally, the plasmid contigs from each category were compared with publicly available sequences in NCBI.

Virulence and resistance genes profiles.

A set of 97 virulence and resistance genes identified based on the literature (Table S4) were searched in the *L. monocytogenes* genomes using a custom database in ABRicate with minimum identification (--minid) threshold of 90%. Genes detected on contigs with very low coverage (<2×) compared to the rest of the contigs in each assembly were excluded as possible contamination. Except for truncations/interruptions in *actA*, *inlA*, *inlB* and *comK* genes, a gene was considered present when coverage was >90%. Additionally, the assemblies were screened against the ResFinder v4.0 database for the presence of antimicrobial resistance (AMR) genes. Mutations and internal truncations in *inlA* and *inlB* genes were identified by extracting the genes as explained for plasmid replicon genes, translated into protein sequences and aligned to the EGD-e *inlA* (NC_003210.1:454534..456936) and *inlB* (NC_003210.1:457021..458913) as reference genes.

In silico screening for disinfectant resistance genes in global _L. monocytogenes_ isolates.

To study the global prevalence of the genes associated with increased tolerance to QACs (*emrC*, *emrE*, *bcrABC* and *qacH*) in *L. monocytogenes*, publicly available raw sequencing data deposited in ENA as of 26 November 2018 were screened using COBS (COmpact Bit-sliced Signature index) v0.1.2 with default settings for *emrE* (NC_013766.2:c1850670-1850347), *emrC* (MT912503.1:2384-2770) *bcrABC* (JX023284.1) sequences and homology reduced to maximum 90% nucleotide identity for *qacH* variants (Table S1). COBS consisted of 661,405 assembled and indexed bacterial genomes and 26,006 of them were annotated as *L. monocytogenes* (71). The STs of the *L. monocytogenes* isolates positive for QAC tolerance genes were obtained from the metadata associated with COBS. Furthermore, the pair-end sequencing runs deposited in ENA between 27 November 2018 and 29 April 2021 were downloaded and screened for QAC tolerance genes using KMA with minimum template identity of 90% (72) . To index the database of biocide genes, a k-mer = 16 was used as a default. The STs of the *L. monocytogenes* isolates harboring the disinfectant resistance genes were determined by stringmlst v0.6.3 (73) with k-mer = 35 using *L. monocytogenes* MLST database and converted to CCs using the Pasteur's BIGSdb-Lm MLST database (https://bigsdb.pasteur.fr/). The world map for QAC tolerance gene distribution was produced using the R packages *mapPies* (https://search.r-

606 project.org/CRAN/refmans/rworldmap/html/mapPies.html) and *rworldmap* version 1.3-6
607 (https://cran.r-project.org/web/packages/rworldmap/rworldmap.pdf).
608
609 **Statistical analysis.**
610 The heterogeneity in proportion of clonal complexes (CCs), geographical locations and isolation
611 sources that were positive within and between genes was estimated using a random-effects model
612 proposed by and implemented in the R package meta to produce forest plots (74). Statistical
613 heterogeneity within and between groups was estimated using the Cochran chi-square test and the
614 Cochrane I^2 index. The Pearson's chi-squared association test was performed using Microsoft Excel to
615 determine statistically significant ($p < 0.05$) association between presence or absence of plasmids, stress
616 survival and virulence genes in the phenotypes defined as tolerant or sensitive to BC.
617
618
619 **Data availability.**
620 The raw sequencing data have been deposited in the European Nucleotide Archive with metadata
621 overview in Table S1.
622

623 **ACKNOWLEDGEMENTS**

624 The authors gratefully acknowledge Pia Engelsmann, Resadije Idrizi, Rannvá Høgnadóttir Houmann
625 (Research Group for Food Microbiology and Hygiene, DTU) and Jacob Dyring Jensen, Gunhild Larsen
626 and Christina Aaby Svendsen (Research Group for Genomic Epidemiology, DTU) for excellent
627 technical assistance throughout the project. Stephanie Brown (Food Innovation Center, Oregon State
628 University, Portland, OR, USA) is thanked for her assistance with some of the US *L. monocytogenes*
629 isolates. Lone Gram (Department of Biotechnology and Biomedicine, DTU) is kindly acknowledged
630 for sharing *L. monocytogenes* isolates for this project. The Institut Pasteur teams are acknowledged for
631 the curation and maintenance of BIGSdb-Pasteur databases at https://bigsdb.pasteur.fr/. This work was

supported by the Danish Dairy Research Foundation, the Milk Levy Fund and Karl Pedersen og Hustrus Industrifond (DI-2019-07020) grants.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Radoshevich L, Cossart P. 2018. *Listeria monocytogenes*: towards a complete picture of its physiology and pathogenesis. Nat Rev Microbiol 16:32-46.

2. Huang C, Lu TL, Yang Y. 2023. Mortality risk factors related to listeriosis — A meta-analysis. J Infect Public Health 16:771-83.

3. Thomas J, Govender N, McCarthy KM, Erasmus LK, Doyle TJ, Allam M, Ismail A, Ramalwa N, Sekwadi P, Ntshoe G, Shonhiwa A, Essel V, Tau N, Smouse S, Ngomane HM, Disenyeng B, Page NA, Govender NP, Duse AG, Stewart R, Thomas T, Mahoney D, Tourdjman M, Disson O, Thouvenot P, Maury MM, Leclercq A, Lecuit M, Smith AM, Blumberg LH. 2020. Outbreak of listeriosis in South Africa associated with processed Meat. N Engl J Med 13:632-43.

4. Gilmour MW, Graham M, Domselaar GV, Tyler S, Kent H, Trout-Yakel KM, Larios O, Allen V, Lee B, Nadon C. 2010. High-throughput genome sequencing of two *Listeria monocytogenes* clinical isolates during a large foodborne outbreak. BMC Genomics 11:120.

5. Kvistholm Jensen A, Nielsen EM, Björkman JT, Jensen T, Müller L, Persson S, Bjerager G, Perge A, Krause TG, Kiil K, Sørensen G, Andersen JK, Mølbak K, Ethelberg S. 2016. Whole-genome sequencing used to investigate a nationwide

outbreak of listeriosis caused by ready-to-eat delicatessen meat, Denmark, 2014. Clin Infect Dis 63:64-70.

6. Linnan MJ, Mascola L, Lou XD, Goulet V, May S, Salminen C, Hird DW, Yonekura ML, Hayes P, Weaver R, Audurier A, Plikaytis BD, Fannin SL, Kleks A, Broome CV. 1988. Epidemic listeriosis associated with Mexican-style cheese. N Engl J Med 319:823-8.

7. Orsi RH, Borowsky ML, Lauer P, Young SK, Nusbaum C, Galagan JE, Birren BW, Ivy RA, Sun Q, Graves LM, Swaminathan B, Wiedmann M. 2008. Short-term genome evolution of *Listeria monocytogenes* in a non-controlled environment. BMC Genomics 9:539.

8. Belias A, Sullivan G, Wiedmann M, Ivanek R. 2022. Factors that contribute to persistent Listeria in food processing facilities and relevant interventions: A rapid review. Food Control 133:108579.

9. Maury MM, Tsai YH, Charlier C, Touchon M, Chenal-Francisque V, Leclercq A, Criscuolo A, Gaultier C, Roussel S, Brisabois A, Disson O, Rocha EPC, Brisse S, Lecuit M. 2016. Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. Nat Genet 48:308–13.

10. Moura A, Lefrancq N, Leclercq A, Wirth T, Borges V, Gilpin B, Dallman TJ, Frey J, Franz E, Nielsen EM, Thomas J, Pightling A, Howden BP, Tarr CL, Gerner-Smidt P, Cauchemez S, Salje H, Brisse S, Lecuit M. 2020. Emergence and global spread of *Listeria monocytogenes* main clinical clonal complex. Sci Adv 49:eabj9805.

11. Fagerlund A, Langsrud S, Schirmer BCT, Møretrø T, Heir E. 2016. Genome analysis of *Listeria monocytogenes* sequence type 8 strains persisting in salmon and poultry processing environments and comparison with related strains. PLOS ONE. 8:11:e0151117.

12. Schmitz-Esser S, Anast JM, Cortes BW. 2021. A large-scale sequencing-based survey of plasmids in *Listeria monocytogenes* reveals global dissemination of plasmids. Front Microbiol 12.

13. Bland R, Brown SRB, Waite-Cusic J, Kovacevic J. 2022. Probing antimicrobial resistance and sanitizer tolerance themes and their implications for the food industry through the *Listeria monocytogenes* lens. Compr Rev Food Sci Food Saf 21:1777-802.

14. McDonnell G, Russell AD. 1999. Antiseptics and disinfectants: activity, action, and resistance. Clin Microbiol Rev 12:147–79.

15. Fagerlund A, Heir E, Møretrø T, Langsrud S. 2020. *Listeria monocytogenes* biofilm removal using different commercial cleaning agents. Molecules 12;25(4):792.

16. Manso B, Melero B, Stessl B, Jaime I, Wagner M, Rovira J, Rodríguez-Lázaro D. 2020. The response to oxidative stress in *Listeria monocytogenes* is temperature dependent. Microorganisms 8:521.

17. Palma F, Radomski N, Guérin A, Sévellec Y, Félix B, Bridier A, Soumet C, Roussel S, Guillier L. 2022. Genomic elements located in the accessory repertoire drive the adaptation to biocides in *Listeria monocytogenes* strains from different ecological niches. Food Microbiol 106:103757.

18. Kovacevic J, Ziegler J, Wałecka-Zacharska E, Reimer A, Kitts DD, Gilmour MW. 2016. Tolerance of *Listeria monocytogenes* to quaternary ammonium sanitizers is mediated by a novel efflux pump encoded by *emrE*. Appl Environ Microbiol 82:939–53.

19. Müller A, Rychli K, Muhterem-Uyar M, Zaiser A, Stessl B, Guinane CM, Cotter PD, Wagner M, Schmitz-Esser S. 2013. Tn6188 - A novel transposon in *Listeria monocytogenes* responsible for tolerance to benzalkonium chloride. PLOS ONE 2;8:e76835.

20. Elhanafi D, Dutta V, Kathariou S. 2010. Genetic characterization of plasmid-associated benzalkonium chloride resistance determinants in a *Listeria monocytogenes* strain from the 1998-1999 outbreak. Appl Environ Microbiol 76:8231–8.

21. Kremer PHC, Lees JA, Koopmans MM, Ferwerda B, Arends AWM, Feller MM, Schipper K, Seron MV, Ende A van der, Brouwer MC, Beek D van de, Bentley SD. 2017. Benzalkonium tolerance genes and outcome in *Listeria monocytogenes* meningitis. Clin Microbiol Infect 23:265

22. Roedel A, Dieckmann R, Brendebach H, Hammerl JA, Kleta S, Noll M, S AD, Vincze S. 2019. Biocide-tolerant *Listeria monocytogenes* isolates from German food production plants do not show cross-resistance to clinically relevant antibiotics. Appl Environ Microbiol 85:e01253-19.

23. Moretro T, Schirmer BCT, Heir E, Fagerlund A, Hjemli P, Langsrud S. 2017. Tolerance to quaternary ammonium compound disinfectants may enhance growth of *Listeria monocytogenes* in the food industry. Int J Food Microbiol 241:215–24.

24. Heir E, Lindstedt BA, Røtterud OJ, Vardund T, Kapperud G, Nesbakken T. 2004. Molecular epidemiology and disinfectant susceptibility of *Listeria monocytogenes* from meat processing plants and human infections. Int J Food Microbiol 96:85–96.

25. Assisi C, Forauer E, Oliver HF, Etter AJ. 2020. Genomic and transcriptomic analysis of biofilm formation in persistent and transient *Listeria monocytogenes* isolates from the retail deli environment does not yield insight into persistence mechanisms. Foodborne Pathog Dis 18:179–88.

26. Bland R, Brown SRB, Waite-Cusic J, Kovacevic J. 2022. Probing antimicrobial resistance and sanitizer tolerance themes and their implications for the food industry through the *Listeria monocytogenes* lens. Compr Rev Food Sci Food Saf 21:1777–802.

759 27. Møretrø T, Schirmer BCT, Heir E, Fagerlund A, Hjemli P, Langsrud S. 2017.
760 Tolerance to quaternary ammonium compound disinfectants may enhance growth of
761 *Listeria monocytogenes* in the food industry. Int J Food Microbiol 241:215–24.
762
763 28. Luque-Sastre L, Fox EM, Jordan K, Fanning S. 2018. A comparative study of the
764 susceptibility of *Listeria* species to sanitizer treatments when grown under planktonic
765 and biofilm conditions. J Food Prot 81:1481–90.
766
767 29. Conficoni D, Losasso C, Cortini E, Di Cesare A, Cibin V, Giaccone V, Corno G,
768 Ricci A. 2016. Resistance to biocides in *Listeria monocytogenes* collected in meat-
769 processing environments. Front Microbiol 2016 7.
770
771 30. Rakic-Martinez M, Drevets DA, Dutta V, Katic V, Kathariou S. 2011. *Listeria*
772 *monocytogenes* Strains selected on ciprofloxacin or the disinfectant benzalkonium
773 chloride exhibit reduced susceptibility to ciprofloxacin, gentamicin, benzalkonium
774 chloride, and other toxic compounds. Appl Environ Microbiol 77:8714–21.
775
776 31. Gelbicova T, Florianova M, Hluchanova L, Kalova A, Korena K, Strakova N,
777 Karpiskova R. 2021. Comparative analysis of genetic determinants encoding
778 cadmium, arsenic, and benzalkonium chloride resistance in *Listeria monocytogenes* of
779 human, food, and environmental origin. Front Microbiol 11.
780
781 32. Castro H, Douillard FP, Korkeala H, Lindström M. 2021. Mobile elements harboring
782 heavy metal and bacitracin resistance genes are common among *Listeria*
783 *monocytogenes* strains persisting on dairy farms. mSphere 6:4
784
785 33. Hurley D, Luque-Sastre L, Parker CT, Huynh S, Eshwar AK, Nguyen SV, Andrews
786 N, Moura A, Fox EM, Jordan K, Lehner A, Stephan R, Fanning S. 2019. Whole-
787 genome sequencing-based characterization of 100 *Listeria monocytogenes* isolates
788 collected from food processing environments over a four-year period. mSphere 4:4.
789
790 34. Olier M, Pierre F, Rousseaux S, Lemaître JP, Rousset A, Piveteau P, Guzzo J. 2003.
791 Expression of truncated internalin a is involved in impaired internalization of some

Listeria monocytogenes isolates carried asymptomatically by humans. Infect Immun 71:1217–24.

35. Fagerlund A, Wagner E, Møretrø T, Heir E, Moen B, Rychli K, Langsrud S. 2022. Pervasive *Listeria monocytogenes* is common in the Norwegian food system and is associated with increased prevalence of stress survival and resistance determinants. Appl Environ Microbiol 88:e00861-22.

36. Fagerlund A, Møretrø T, Aspholm M, Lindbäck T, Langsrud S. 2022. WGS analysis of *Listeria monocytogenes* from rural, urban, and farm environments in Norway: Genetic diversity, persistence, and relation to clinical and food isolates. Appl Environ Microbiol 88:e02136-21.

37. Martínez-Suárez JV, Ortiz S, López-Alonso V. 2016. Potential impact of the resistance to quaternary ammonium disinfectants on the persistence of *Listeria monocytogenes* in food processing environments. Front Microbiol 7.

38. Stoller A, Stevens M, Stephan R, Guldimann C. 2019. Characteristics of *Listeria monocytogenes* strains persisting in a meat processing facility over a 4-year period. Pathogens 8:32.

39. Romanova N, Favrin S, Griffiths MW. 2002. Sensitivity of *Listeria monocytogenes* to sanitizers used in the meat processing industry. Appl Environ Microbiol 68:6405–9.

40. Ortiz S, López-Alonso V, Rodríguez P, Martínez-Suárez JV. 2016. 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 82:308–17.

41. 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 Cont 57:321–6.

42. Meier AB, 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.

43. Touche C, Hamchaoui S, Quilleré A, Darsonval M, Dubois-Brissonnet F. 2023. Growth of *Listeria monocytogenes* is promoted at low temperature when exogenous unsaturated fatty acids are incorporated in its membrane. Food Microbiol 110.

44. Bland RN, Johnson JD, Waite-Cusic JG, Weisberg AJ, Riutta ER, Chang JH, Kovacevic J. 2021. Application of whole genome sequencing to understand diversity and presence of genes associated with sanitizer tolerance in *Listeria monocytogenes* from produce handling sources. Foods 10.

45. Chmielowska C, Korsak D, Chapkauskaitse E, Decewicz P, Lasek R, Szuplewska M, Bartosik D. 2021. Plasmidome of *Listeria spp*. the repA-family business. Int J Mol Sci 22:19.

46. Dutta V, Elhanafi D, Kathariou S. 2013. Conservation and distribution of the benzalkonium chloride resistance cassette *bcrABC* in *Listeria monocytogenes*. Appl Environ Microbiol 79:6067–74.

47. Nightingale KK, Ivy RA, Ho AJ, Fortes ED, Njaa BL, Peters RM, Wiedmann M. 2008. *inlA* premature stop codons are common among *Listeria monocytogenes* isolates from foods and yield virulence-attenuated strains that confer protection against fully virulent strains. Appl Environ Microbiol 74:6570–83.

48. Ratani SS, Siletzky RM, Dutta V, Yildirim S, Osborne JA, Lin W, Hitchins AD, Ward TJ, Kathariou S. 2012. Heavy metal and disinfectant resistance of *Listeria monocytogenes* from foods and food processing plants. Appl Environ Microbiol 78:6938–45.

49. Harter E, Wagner EM, 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:e00827-17.

50. He Y, Xu T, Li S, Mann DA, Britton B, Oliver HF, Den Bakker HC, Deng X. 2022. Integrative assessment of reduced *Listeria monocytogenes* susceptibility to benzalkonium chloride in produce processing environments. Appl Environ Microbiol 88:e01269-22.

51. Daeschel D, Pettengill JB, Wang Y, Chen Y, Allard M, Snyder AB. 2022. Genomic analysis of *Listeria monocytogenes* from US food processing environments reveals a high prevalence of QAC efflux genes but limited evidence of their contribution to environmental persistence. BMC Genomics 23:11.

52. Maury MM, Bracq-Dieye H, Huang L, Vales G, Lavina M, Thouvenot P, Disson O, Leclercq A, Brisse S, Lecuit M. 2019. Hypervirulent *Listeria monocytogenes* clones' adaption to mammalian gut accounts for their association with dairy products. Nat Commun 10.

53. Painset A, Björkman JT, Kiil K, Guillier L, Mariet JF, Félix B, Amar C, Rotariu O, Roussel S, Perez-Reche F, Brisse S, Moura A, Lecuit M, Forbes K, Strachan N, Grant K, Møller-Nielsen E, Dallman TJ. 2019. LiSEQ – whole-genome sequencing of a cross-sectional survey of *Listeria monocytogenes* in ready-to-eat foods and human clinical cases in Europe. Microb Genomics 5:2.

54. Kropac AC, Eshwar AK, Stephan R, Tasara T. 2019. New insights on the role of the *pLMST6* plasmid in *Listeria monocytogenes* biocide tolerance and virulence. Front Microbiol 10.

55. Wiegand I, Hilpert K, Hancock REW. 2008. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. Nat Protoc. 3:163–75.

56. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. Curr Protoc Bioinforma 70:1.

57. Andrews S. 2010. FastQC: A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

58. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: Quality assessment tool for genome assemblies. Bioinformatics 29:1072-5

59. Larsen MV, Cosentino S, Lukjancenko O, Saputra D, Rasmussen S, Hasman H, Sicheritz-Pontén T, Aarestrup FM, Ussery DW, Lund O. 2014. Benchmarking of methods for genomic taxonomy. J Clin Microbiol 52:1529–39.

60. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068-9

61. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics. 31:3691–3.

62. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–3.

63. Letunic I, Bork P. 2019. Interactive tree of life (iTOL) v4: Recent updates and new developments. Nucleic Acids Res 47.

64. Xu S, Dai Z, Guo P, Fu X, Liu S, Zhou L, Tang W, Feng T, Chen M, Zhan L, Wu T, Hu E, Jiang Y, Bo X, Yu G. 2021. ggtreeExtra: Compact visualization of richly annotated phylogenetic data. Mol Biol Evol 38:4039–42.

65. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10:421.

66. Chmielowska C, Korsak D, Szuplewska M, Grzelecka M, Mackiw E, Stasiak M, Macion A, Skowron K, Bartosik D. 2021. Benzalkonium chloride and heavy metal

resistance profiles of *Listeria monocytogenes* strains isolated from fish, fish products and food-producing factories in Poland. Food Microbiol 98.

67. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–2.

68. Gilchrist CLM, Chooi YH. 2021. Clinker & clustermap.js: Automatic generation of gene cluster comparison figures. Bioinformatics 37:2473–5.

69. Galata V, Fehlmann T, Backes C, Keller A. 2019. PLSDB: a resource of complete bacterial plasmids. Nucleic Acids Res 47:195–202.

70. Roux S, Enault F, Hurwitz BL, Sullivan MB. 2015. VirSorter: mining viral signal from microbial genomic data. PeerJ 28:3.

71. Blackwell G, Hunt MR, Malone K, Leandro GL, Horesh, Alako BTF, Thomson N, Iqbal Z. 2021. Exploring bacterial diversity via a curated and searchable snapshot of archived DNA sequences. PLOS Biol 19:11.

72. Clausen PTLC, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. 2018. BMC Bioinformatics 19:307.

73. Gupta A, Jordan IK, Rishishwar L. 2017. stringMLST: a fast k-mer based tool for multilocus sequence typing. Bioinformatics 33:119–21.

74. DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7:177–88.

Figure 1 A. Percentage of *L. monocytogenes* isolates by source of isolation. Sources below 0.1% are not shown. **B.** Division of isolates by source of isolation for each country of collection. Countries with < 10 isolates are not shown on the figure: Turkey (n=3, food), Russia (n=3, food), Belgium (n=1, food), China (n=1, food), Finland (n=1, food) and France (n=1, production environment). **C.** Temporal distribution of the isolates in the collection covering a time span of 98 years.

Figure 2. Distribution of the main isolation sources within dominating CCs (CCs with > 50 isolates are shown in the figure).

Figure 3 Distribution of the MIC values of sensitive (blue) and tolerant (red) *L. monocytogenes* isolates to Benzalkonium chloride (n=1671), Mida SAN 360 OM (cQAC, n=155) and Didecyldimethylammonium chloride (DDAC, n=247).

Figure 4 Genomic characterisation of the *L. monocytogenes* tolerance to QACs. Distribution of CCs **(A)** and isolation sources **(B)** among the isolates harbouring QAC tolerance genes and those isolates with unknown mechanism of tolerance. Graphs were generated by Circos Viewer v0.63-10 (http://mkweb.bcgsc.ca/tableviewer/).

Figure 5 A. Mid-point rooted ML phylogenetic tree of the RepA sequences identified in 728 of the *L. monocytogenes* isolates. The 13 plasmid replicon groups detected among the isolates in this study and the reference groups are indicated. **B.** Prevalence of plasmid carrying isolates determined as sensitive or tolerant to QAC and their isolation source. Asterisks indicate significant differences determined with the Pearson's chi-squared association test with $p < 0.05$.

Figure 6 A. Stress resistance, virulence and plasmid replicon gene prevalence among the 95 CCs and singleton STs in this study presented as percentage within each CC. A gene is considered present if the gene coverage is > 90%. The phylogeny is a core-genome ML tree built with a representative isolate (the most recent ancestor) from each CC with serogrouping based on the serogrouping scheme from https://bigsdb.pasteur.fr/_nuxt/img/serogroups.9daaa98.png **B.** Prevalence of stress resistance genes among the QAC tolerant (n=388) and sensitive (n=1283) isolates. Asterisks indicate significant differences in prevalence determined with the Pearson's chi-squared association test with $P < 0.05$.

Figure 7 Distribution of *bcrABC, qacH, emrE, emrC* genes globally **(A)** and in Europe **(B)** in *L. monocytogenes* sequencing runs deposited in ENA as of 21 April 2021. The pie charts indicate the proportion (rate) of the genes in each of the country in which at least one of the genes was present and their size reflects the rate of each gene per 1000 genomes (overview in Table S11).

FIGURES

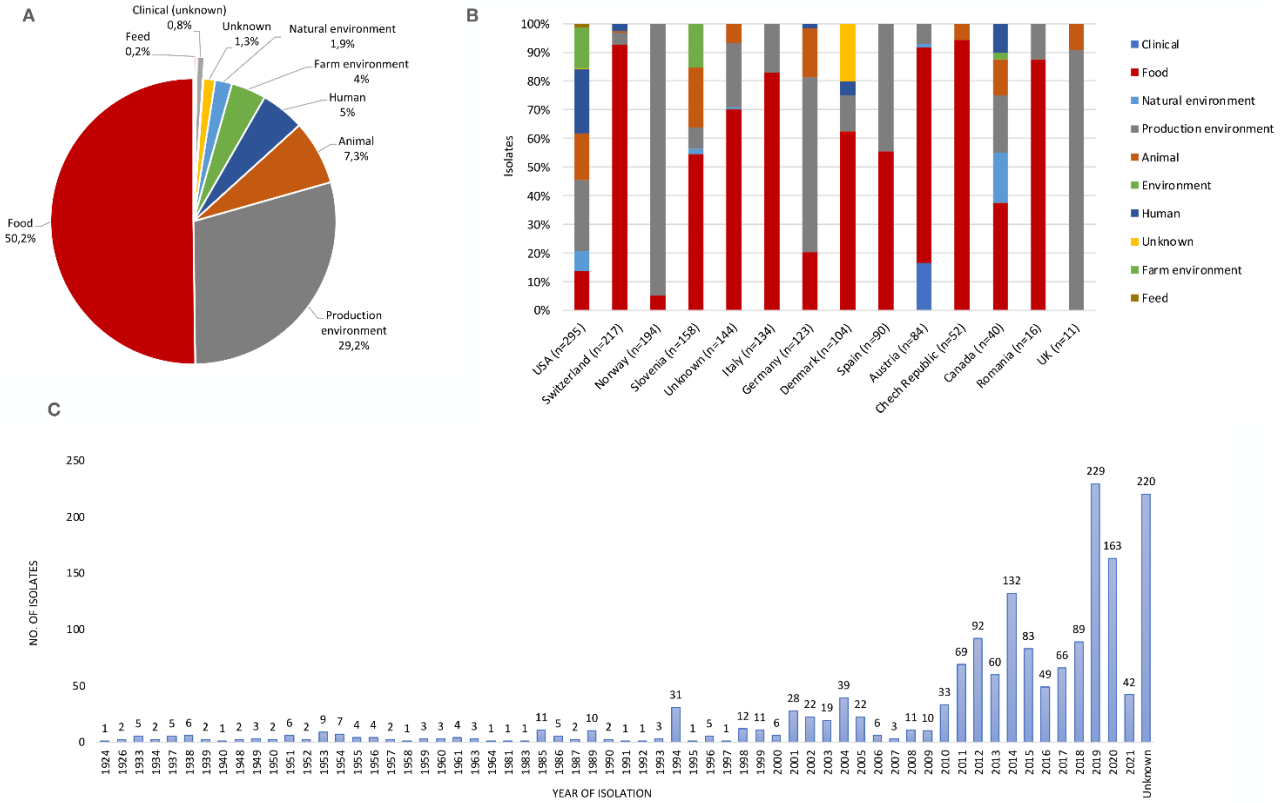


Figure 1 A. Percentage of *L. monocytogenes* isolates by source of isolation. Sources below 0.1% are not shown. **B.** Division of isolates by source of isolation for each country of collection. Countries with < 10 isolates are not shown on the figure: Turkey (n=3, food), Russia (n=3, food), Belgium (n=1, food), China (n=1, food), Finland (n=1, food) and France (n=1, production environment). **C.** Temporal distribution of the isolates in the collection covering a time span of 98 years.

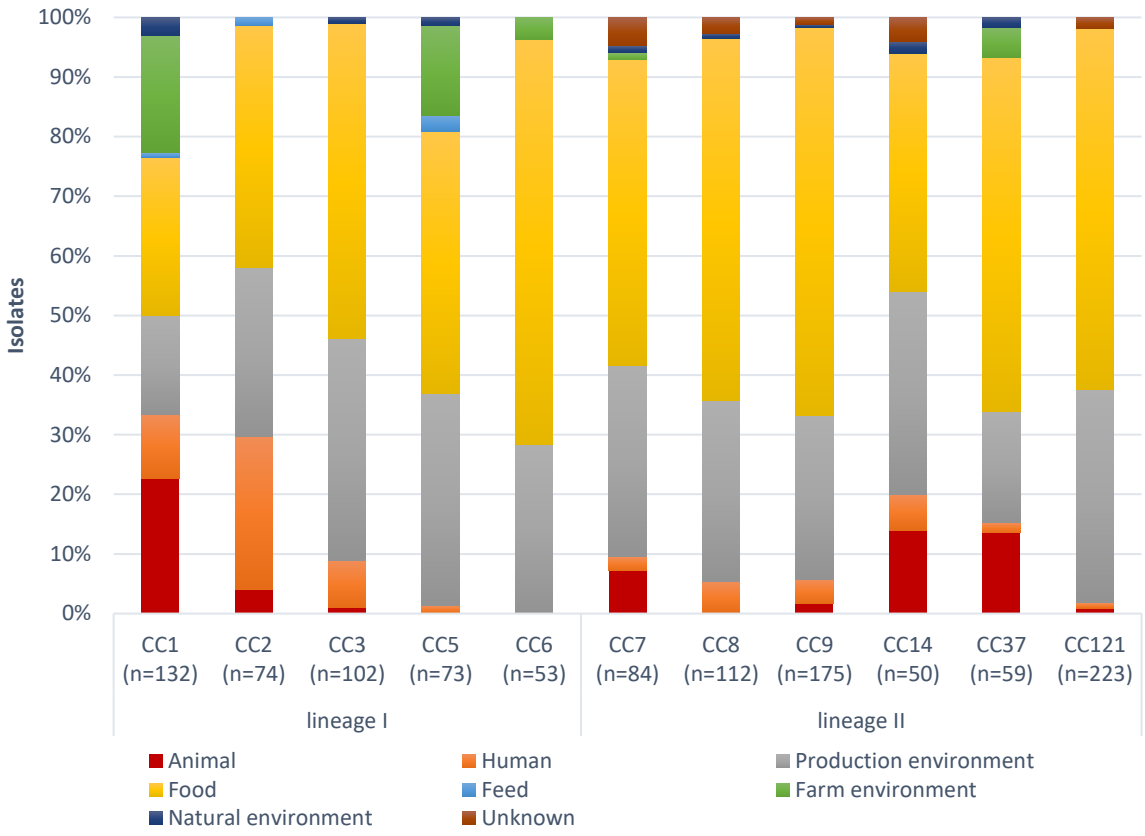


Figure 2. Distribution of the main isolation sources within dominating CCs (CCs with > 50 isolates are shown in the figure).

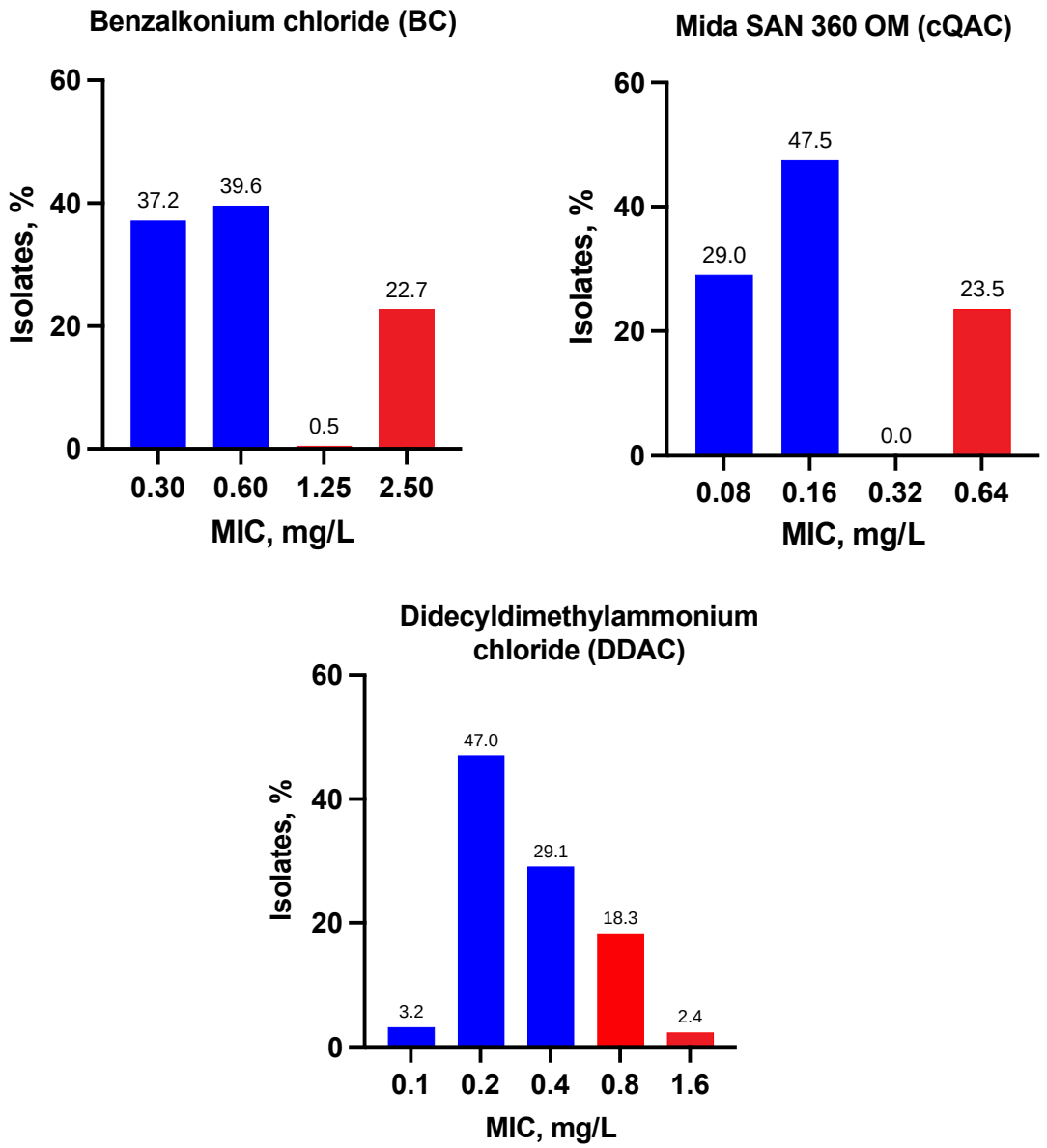
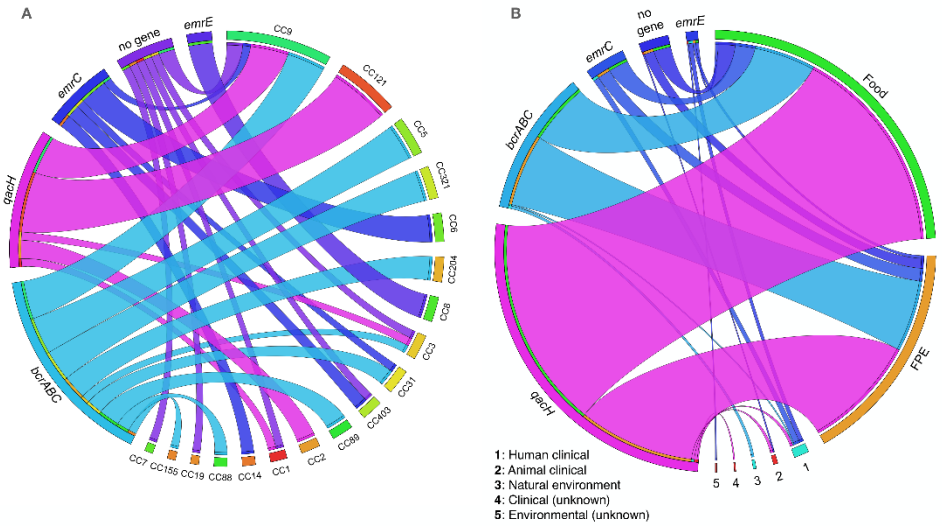


Figure 3 Distribution of the MIC values of sensitive (blue) and tolerant (red) *L. monocytogenes* isolates to Benzalkonium chloride (n=1671), Mida SAN 360 OM (cQAC, n=155) and Didecyldimethylammonium chloride (DDAC, n=247).

1019



1020

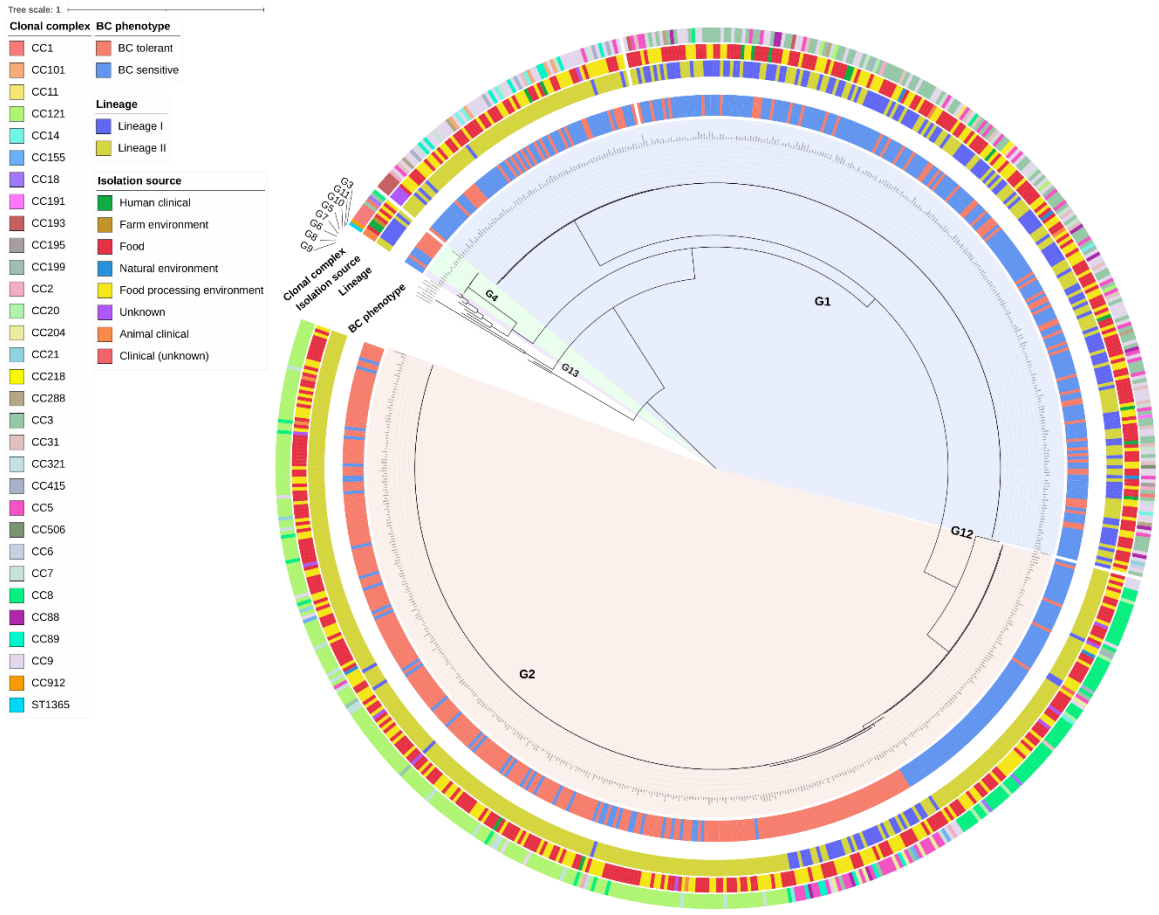
1021 **Figure 4** Genomic characterisation of the *L. monocytogenes* tolerance to QACs. Distribution of CCs

1022 **(A)** and isolation sources **(B)** among the isolates harbouring QAC tolerance genes and those isolates

1023 with unknown mechanism of tolerance. Graphs were generated by Circos Viewer v0.63-10

1024 (http://mkweb.bcgsc.ca/tableviewer/).

A.



B.

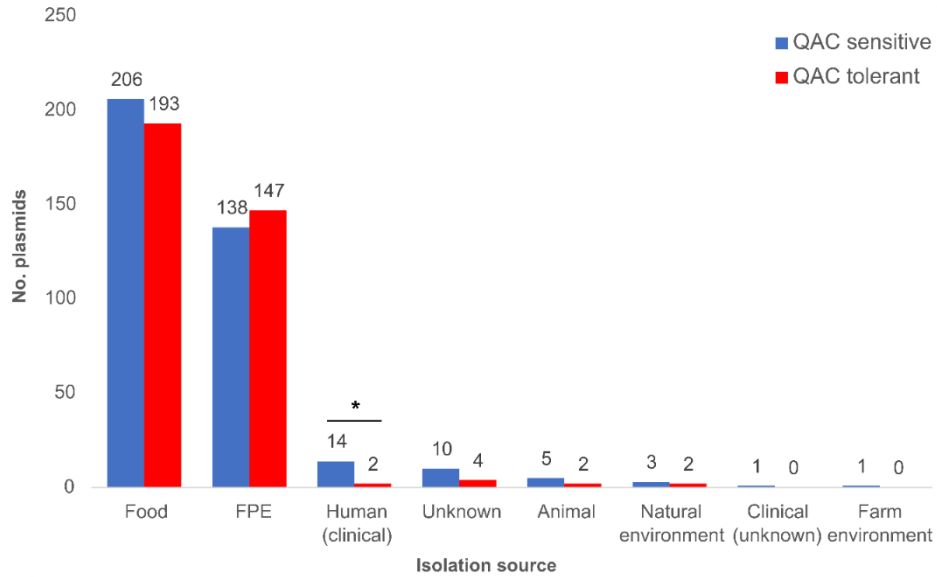
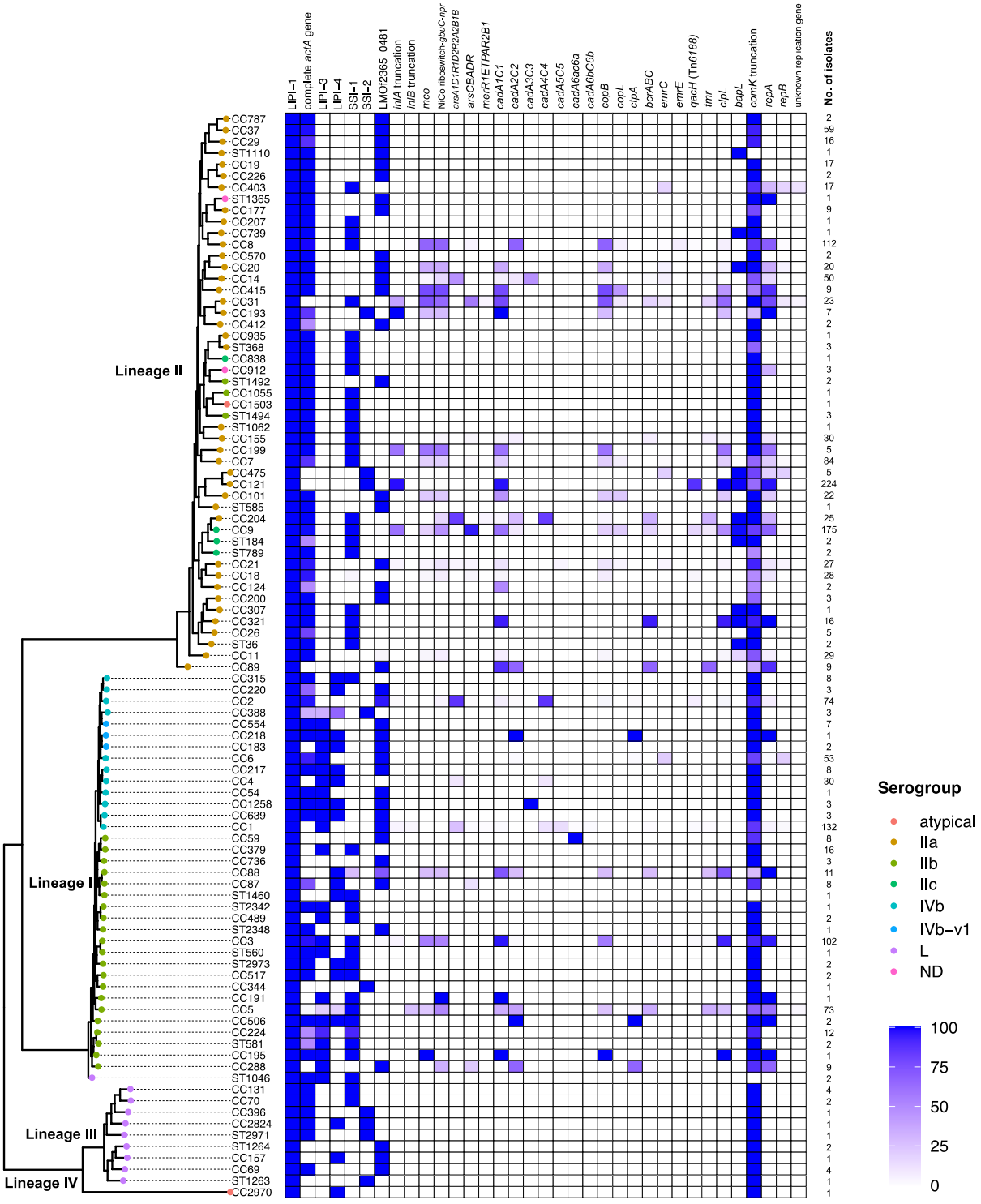


Figure 5 A. Mid-point rooted ML phylogenetic tree of the RepA sequences identified in 728 of the *L. monocytogenes* isolates. The 13 plasmid replicon groups detected among the isolates

in this study and the reference groups are indicated. **B.** Prevalence of plasmid carrying isolates determined as sensitive or tolerant to QAC and their isolation source. Asterisks indicate significant differences determined with the Pearson's chi-squared association test with $p < 0.05$.



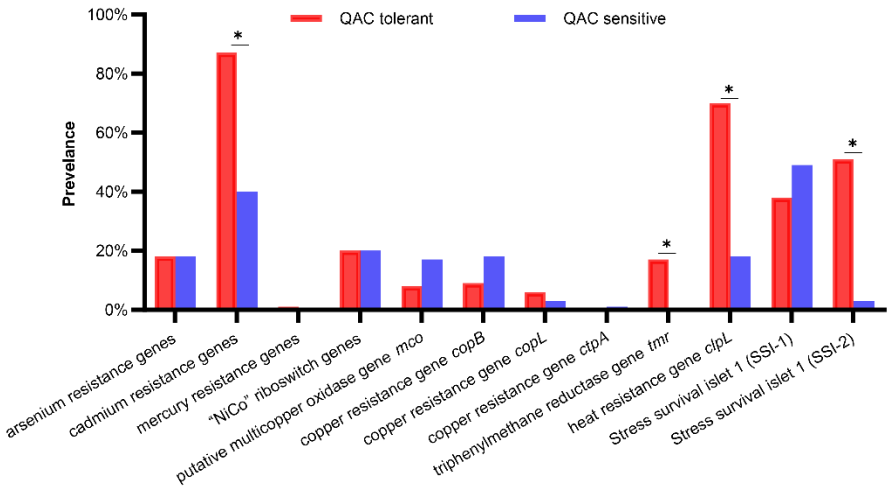


Figure 6 A. Stress resistance, virulence and plasmid replicon gene prevalence among the 95 CCs and singleton STs in this study presented as percentage within each CC. A gene is considered present if the gene coverage is > 90%. The phylogeny is a core-genome ML tree built with a representative isolate (the most recent ancestor) from each CC with serogrouping based on the serogrouping scheme from https://bigsdb.pasteur.fr/_nuxt/img/serogroups.9daaa98.png **B.** Prevalence of stress resistance genes among the QAC tolerant (n=388) and sensitive (n=1283) isolates. Asterisks indicate significant differences in prevalence determined with the Pearson's chi-squared association test with $P < 0.05$.

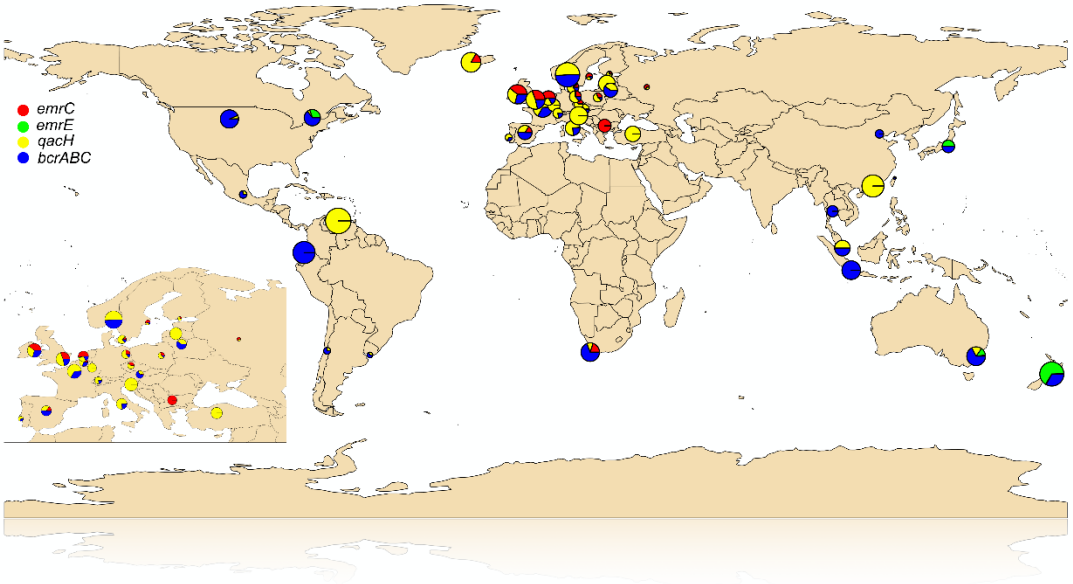
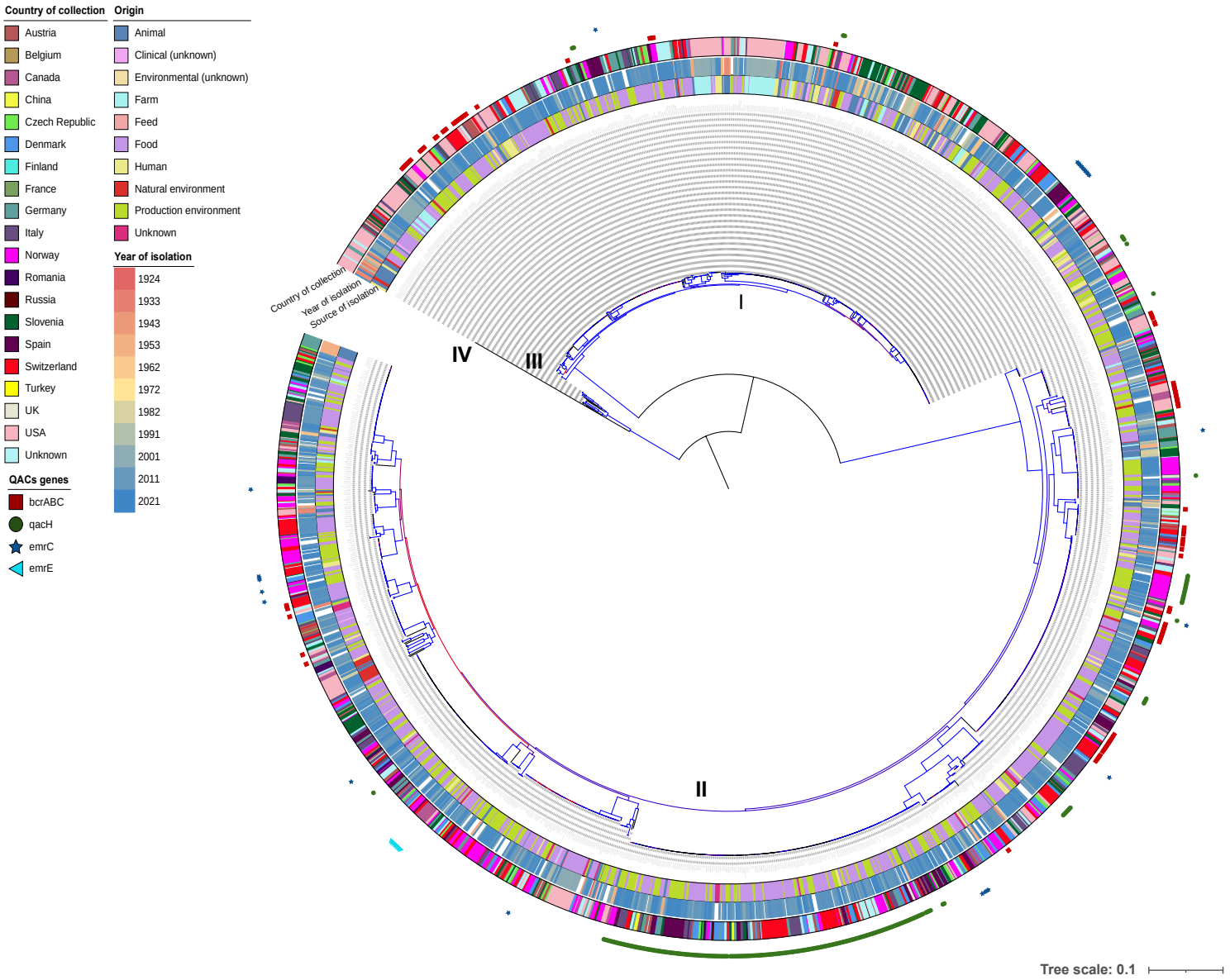


Figure 7 Distribution of *bcrABC*, *qacH*, *emrE*, *emrC* genes globally **(A)** and in Europe **(B)** in *L. monocytogenes* sequencing runs deposited in ENA as of 21 April 2021. The pie charts indicate the proportion (rate) of the genes in each of the country in which at least one of the genes was present and their size reflects the rate of each gene per 1000 genomes (overview in Table S11).

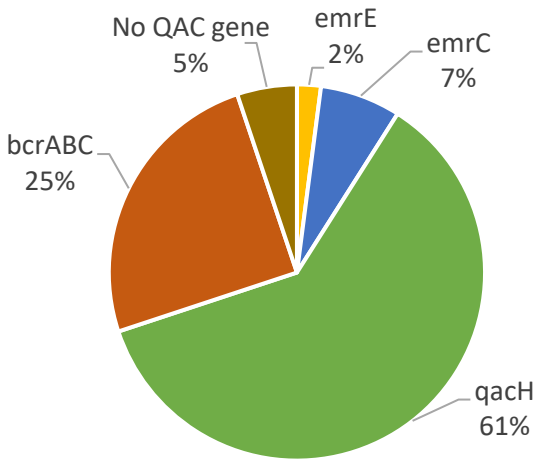
SUPPLEMENTARY FIGURES


Country of collection
Austria
Belgium
Canada
China
Czech Republic
Denmark
Finland
France
Germany
Italy
Norway
Romania
Russia
Slovenia
Spain
Switzerland
Turkey
UK
USA
Unknown
QACs genes
bcrABC
qacH
emrC
emrE
Origin
Animal
Clinical (unknown)
Environmental (unknown)
Farm
Feed
Food
Human
Natural environment
Production environment
Unknown
Year of isolation
1924
1933
1943
1953
1962
1972
1982
1991
2001
2011
2021
Country of collection
Year of isolation
Source of isolation
I
II
III
IV
Tree scale: 0.1

1051

Figure S1 Mid-rooted maximum likelihood (ML) tree constructed from 2.19 Mb core-genome alignment using IQ-TREE with 1000 ultrafast bootstraps and GTR+G nucleotide substitution model. The color of the branches represents bootstrap support values, from 50 (red) to 100 (blue). Roman numbers represent the *L. monocytogenes* phylogenetic lineages. The rings from the inner to outer direction represent the source of isolation (origin), year of isolation (if known) and country of collection. Presence of QAC tolerance genes is presented by symbols outside the rings.

A.



B.

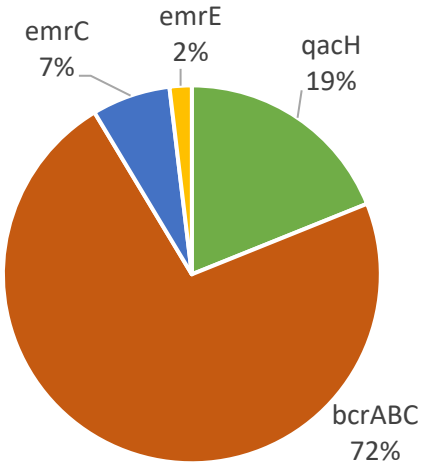
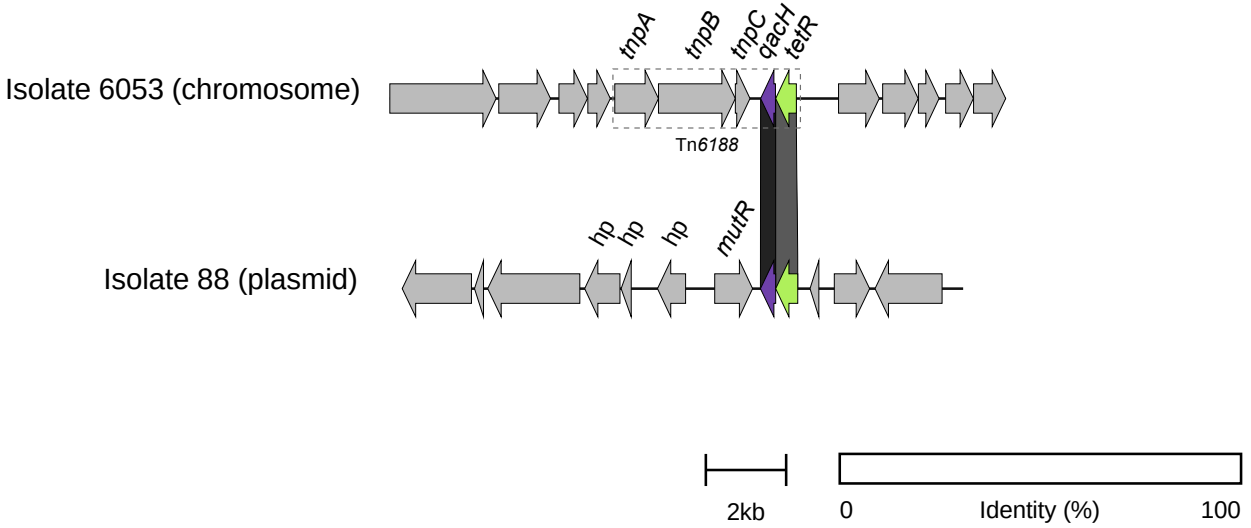


Figure S2 Distribution of the QAC tolerance genes in the 388 QAC-tolerant *L. monocytogenes* isolates in this study (A) and in the ENA *L. monocytogenes* isolates as of April 2021 (B).



1062

1063

1064 **Figure S3** Genetic organisation of the *qacH* gene located on the chromosome and plasmids. The chromosome-located *qacH* is part of the previously

1065 described Tn*6188* (19), while plasmid-located *qacH* has different genetic environment, consisting of *tetR* (72% nucleotide identity to *tetR* on the

1066 chromosome), *mutR* transcriptional regulator and genes encoding hypothetical proteins. The nucleotide identity between plasmid- and

1067 chromosome-located *qacH* was 91%.

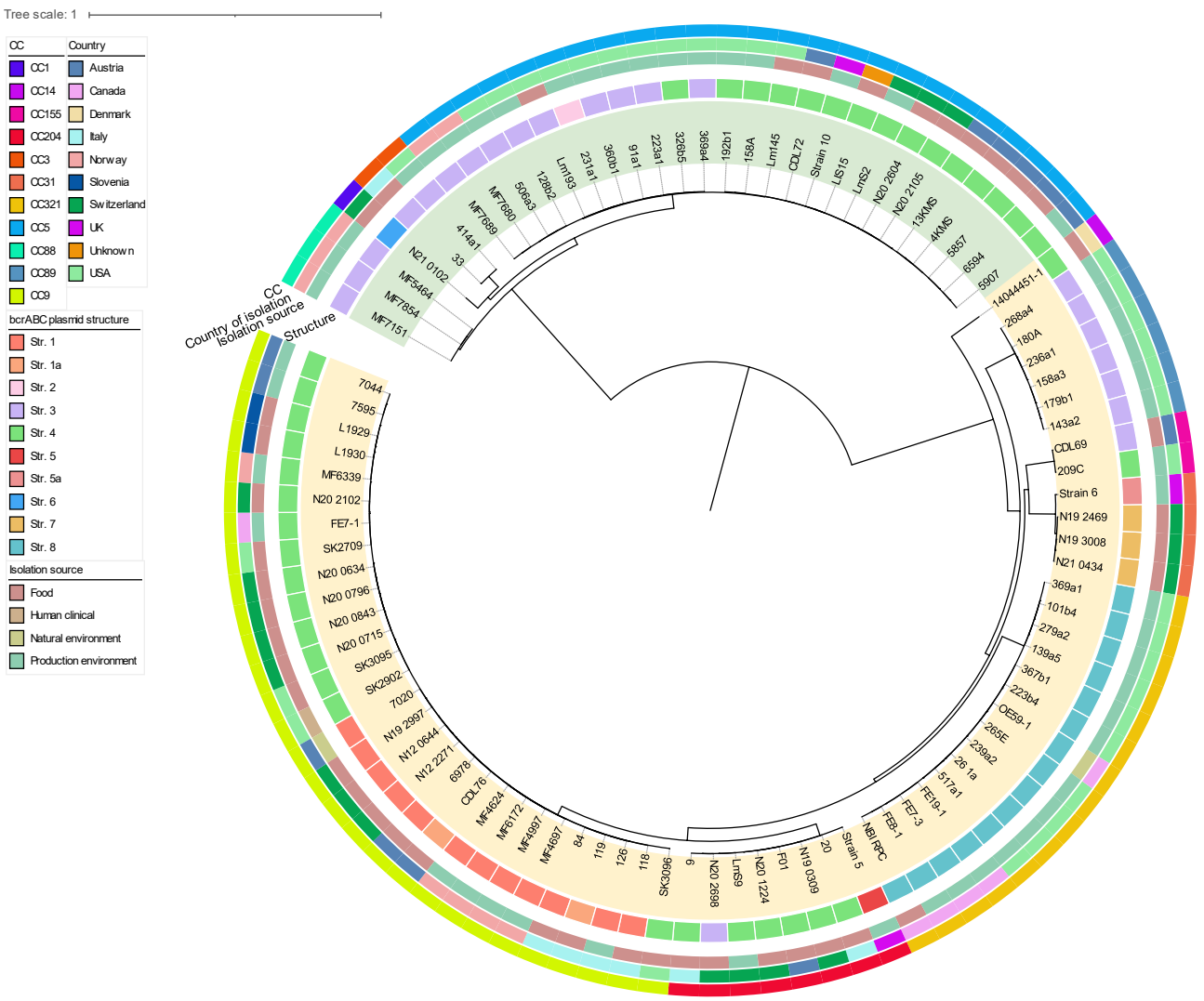


Figure S4 Mid-point rooted SNP-based phylogenetic tree of the 96 *bcrABC*-containing *L. monocytogenes* isolates.

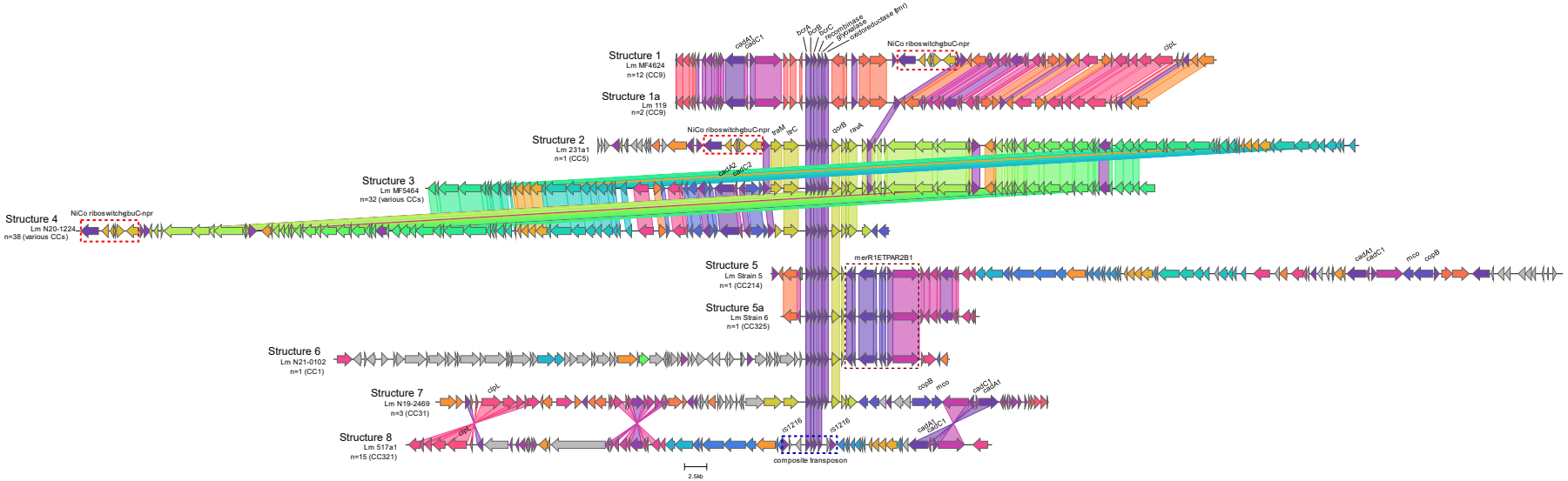


Figure S5 Organisation of the *bcrABC*-harbouring contigs grouped into structures according to their genetic contexts. Links between homologous genes are shown in different colours (the threshold for percentage identity is 90%). Genes in grey have no homologous genes within the analysed plasmid contigs. The red dashed rectangles represent a mobile genetic element carrying Nico riboswitch, *gbuC* and *npr* genes, and the blue dashed rectangle represents the composite transposon in plasmid structure 8. All stress resistance genes identified by blastn are annotated in the contigs; "n" indicates the number of isolates in each structure. The genes are automatically coloured by clinker according to their protein functions

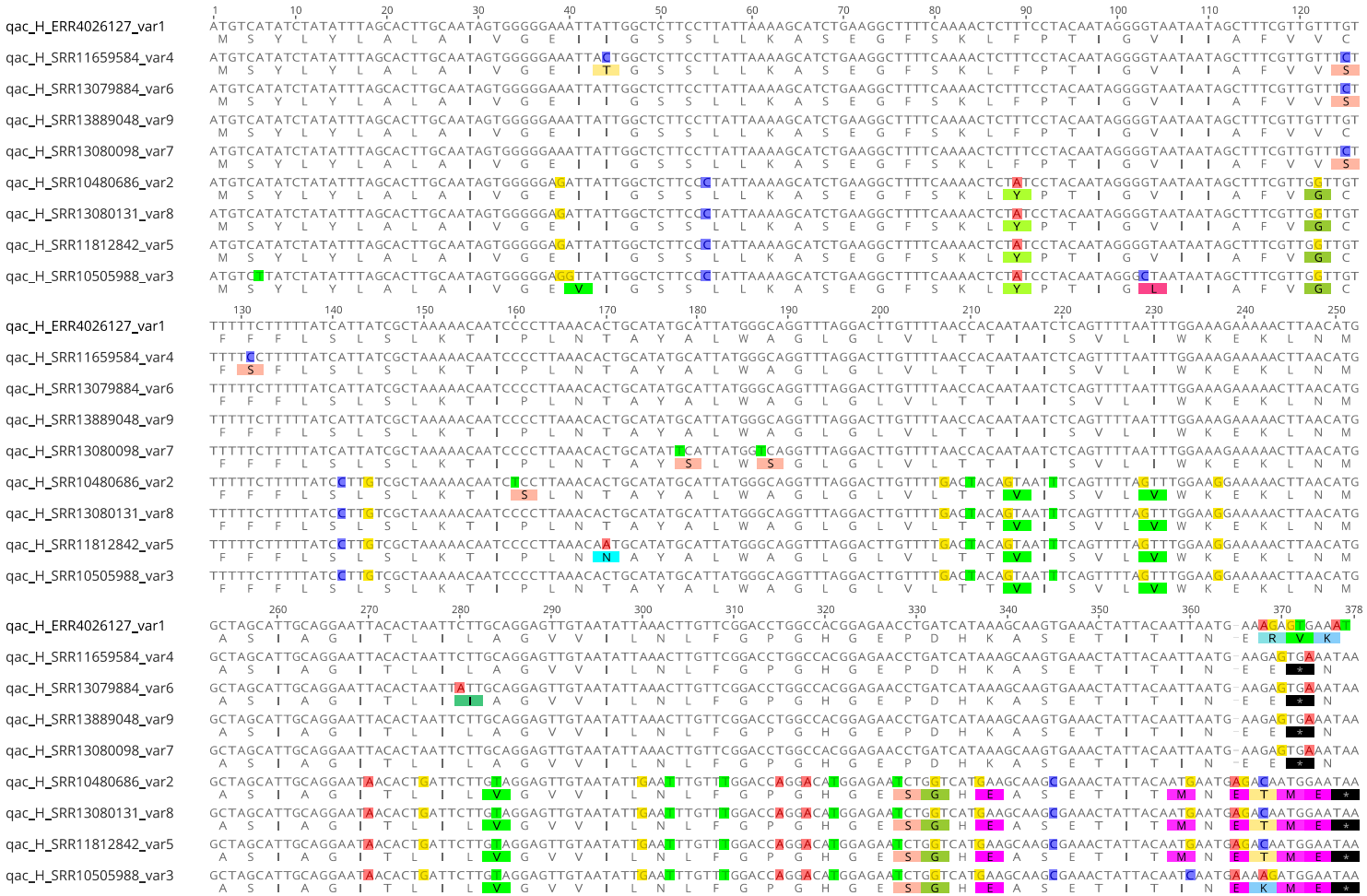


Figure S6 Nucleotide alignment and translation of the nine *qacH* variants identified among the global *L. monocytogenes* dataset with nucleotide identity above 90%. The ENA run accession numbers are given for each isolate harbouring the respective variant.

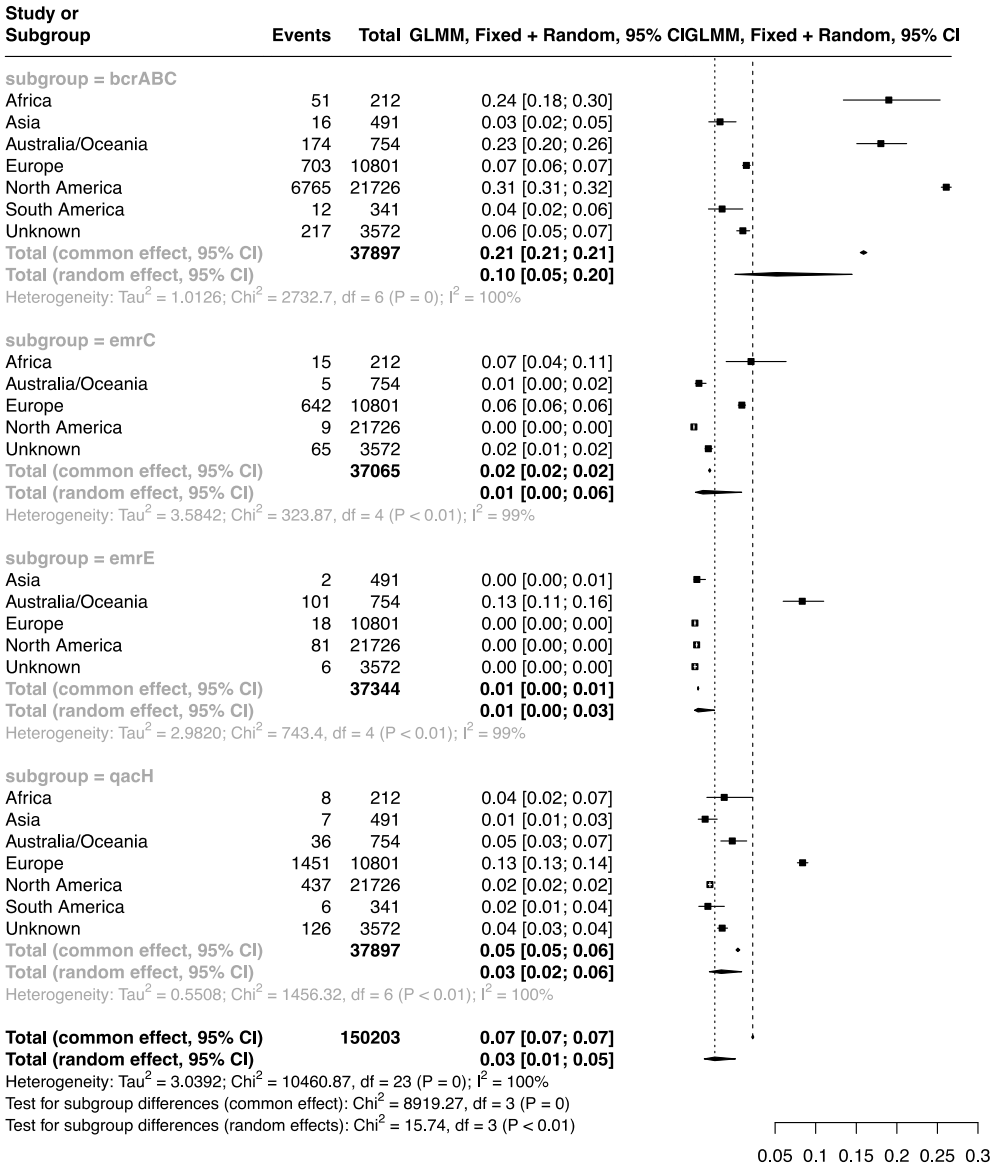


Figure S7 Forest plot showing the association between QAC tolerance genes and continent in the global *L. monocytogenes* dataset.

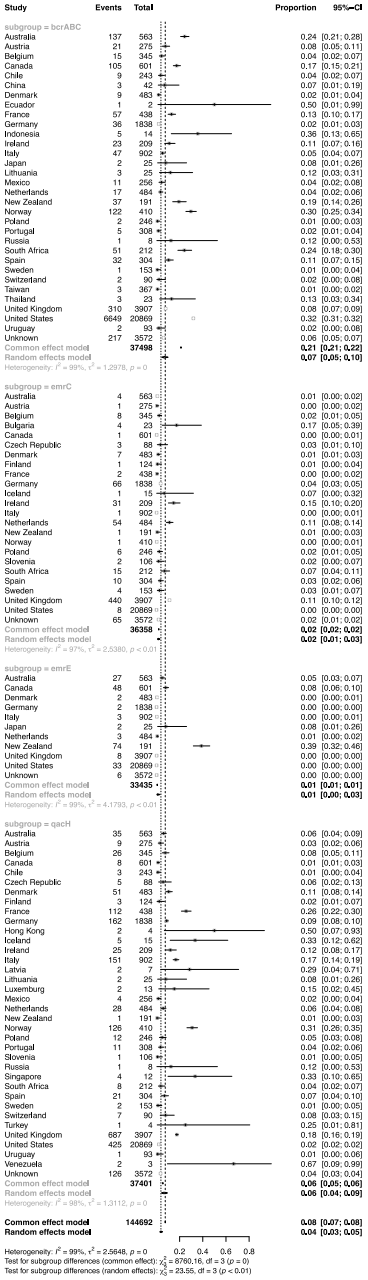


Figure S8 Forest plot showing the association between QAC tolerance genes and country of isolation in the global *L. monocytogenes* dataset.

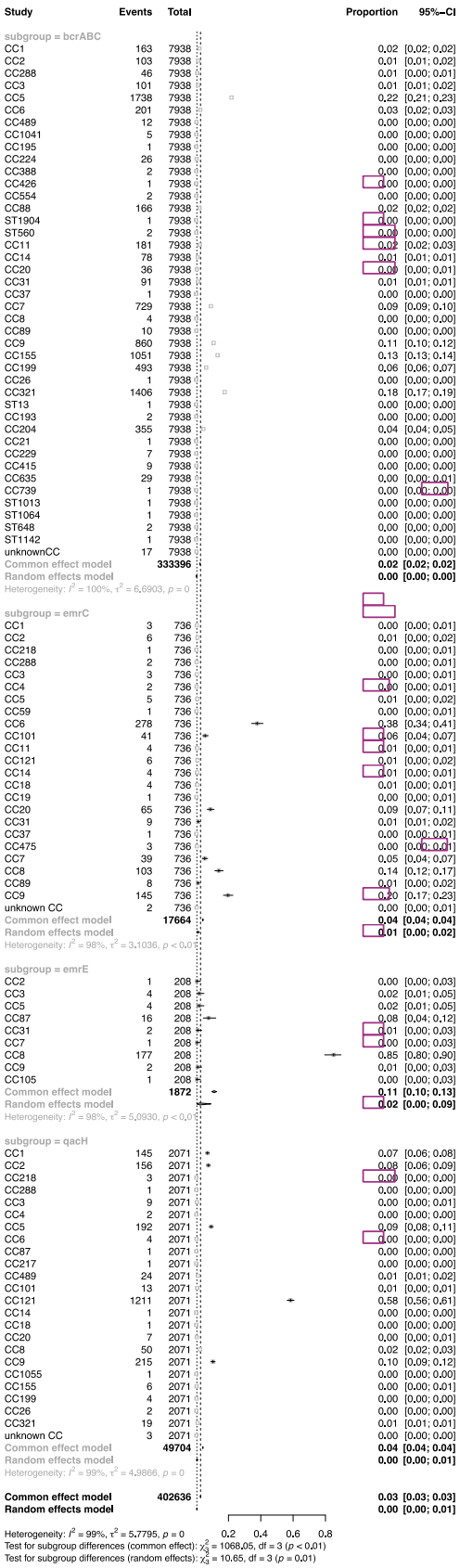


Figure S9 Forest plot showing the association between QAC tolerance genes and CC in the global *L. monocytogenes* dataset.

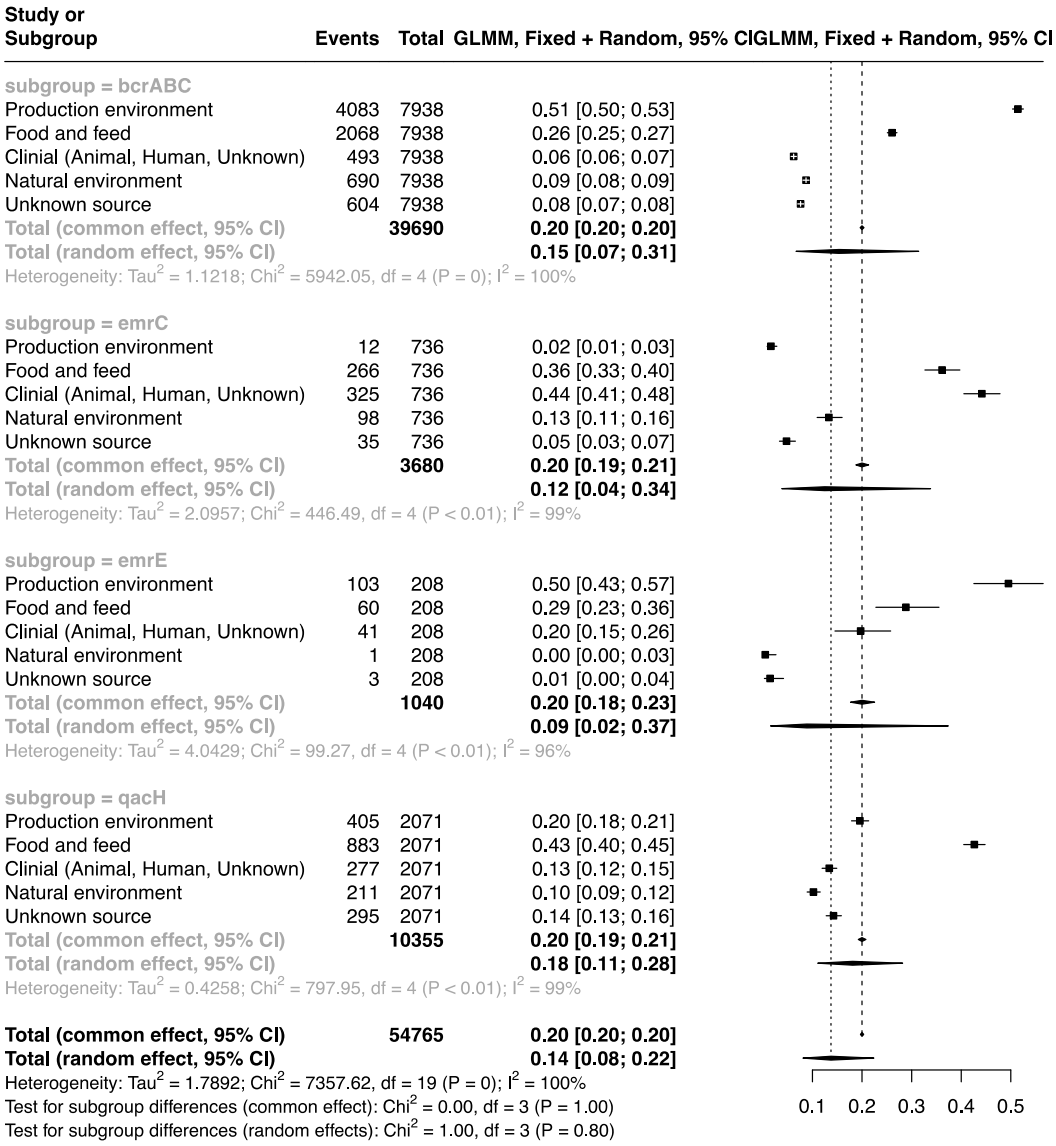


Figure S10 Forest plot showing the association between QAC tolerance genes and source of isolation in the global *L. monocytogenes* dataset.