bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 # Pervasive *Listeria monocytogenes* are common
2 in Norwegian food chains and associated with
3 increased prevalence of stress survival and
4 resistance determinants

5 Annette Fagerlund[a]#, Eva Wagner[a*], Trond Møretrø[a], Even Heir[a], Birgitte Moen[a], Kathrin Rychli[b],
6 Solveig Langsrud[a]

7 [a] Nofima - Norwegian Institute of Food, Fisheries and Aquaculture Research, Ås, Norway

8 [b] Unit for Food Microbiology, Institute for Food Safety, Food Technology and Veterinary Public
9 Health, University of Veterinary Medicine, Vienna, Austria

10 *current address: Institute of Food Science, University of Natural Resources and Life Science,
11 Vienna, Austria

12 # Corresponding author: Annette Fagerlund, annette.fagerlund@nofima.no

13 Keywords: *Listeria*; *L. monocytogenes*; *Listeria monocytogenes*; whole genome sequencing;
14 WGS; food processing environment; food safety; meat industry; salmon industry; meat
15 processing; salmon processing; plasmids; stress resistance; stress survival; antibiotic resistance;
16 persistence; pervasive; pervasion; *inlA*

17 Running title: Pervasive *L. monocytogenes* in Norwegian food industry

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

18 ## Abstract

19 To investigate the diversity, distribution, persistence, and stress resistome of *Listeria*
20 *monocytogenes* clones dominating in food processing environments in Norway, genome
21 sequences from 769 *L. monocytogenes* isolates from food industry environments, foods, and
22 raw materials (of which 512 were sequenced in the present study) were subjected to wgMLST,
23 SNP, and comparative genomic analyses. The dataset comprised isolates from nine meat and six
24 salmon processing facilities in Norway collected over a period of three decades. The most
25 prevalent clonal complex (CC) was CC121, found in ten factories, followed by CC7, CC8, and CC9,
26 found in seven factories each. Overall, 72% of the isolates were classified as persistent, showing
27 20 or fewer wgMLST allelic differences towards an isolate found in the same factory in a
28 different calendar year. Moreover, over half of the isolates (56%) showed this level of genetic
29 similarity towards an isolate collected from a different food processing facility. These were
30 designated as pervasive strains, defined as clusters with the same level of genetic similarity as
31 persistent strains but isolated from different factories. The prevalence of genetic determinants
32 associated with increased survival in food processing environments, including heavy metal and
33 biocide resistance determinants, stress response genes and *inlA* truncation mutations, showed a
34 highly significant increase among pervasive isolates, but not among persistent isolates.
35 Furthermore, these genes were significantly more prevalent among the isolates from food
36 processing environments compared to in isolates from natural and rural environments (n=218)
37 and clinical isolates (n=111) from Norway.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

38 ## Importance

39 *Listeria monocytogenes* can persist in food processing environments for months to decades and
40 spread through food chains by e.g., contaminated raw materials. Knowledge about the
41 distribution and diversity of *L. monocytogenes* is of importance in outbreak investigations and
42 essential to effectively track and control this pathogen in the food chain. The current study
43 presents a comprehensive overview of the prevalence of persistent clones and of the diversity
44 of *L. monocytogenes* in Norwegian food processing facilities. The results demonstrate extensive
45 spread of highly similar strains throughout Norwegian food chains, in that 56% of the 769
46 collected isolates from food processing factories belonged to clusters of *L. monocytogenes*
47 identified in more than one facility. These strains were associated with an overall increase in the
48 prevalence of plasmids and determinants of heavy metal and biocide resistance as well as other
49 genetic elements associated with stress survival mechanisms and persistence.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

50 ## Introduction

51 *Listeria monocytogenes* is a foodborne pathogen responsible for the deadly disease listeriosis.
52 Cross-contamination of food products with *L. monocytogenes* during processing is a major
53 concern, especially with regard to ready-to-eat (RTE) products that support growth of the
54 pathogen prior to consumption. As the pathogen is widespread in natural and urban
55 environments (1, 2) and able to form biofilms and withstand various stresses such as
56 disinfection agents, high and low pH, and low temperatures (3, 4), it is very difficult to eliminate
57 *L. monocytogenes* from food processing environments. Clonal populations of *L. monocytogenes*
58 that survive in the processing environment over an extended time-period (months or years) are
59 referred to as persistent *L. monocytogenes*. In contrast, transient *L. monocytogenes* enter the
60 processing environment but are eliminated through cleaning and disinfection (5, 6). Some
61 authors also define a category of "persistent transient" *L. monocytogenes* contamination which
62 is a consequence of continual introduction of one or more subtypes into the processing
63 environment from outside reservoirs combined with a failure to apply sufficient *Listeria* control
64 measures (7). The concept of pervasive bacterial strains is sometimes used to describe
65 subpopulations of bacteria with enhanced ability to spread or migrate to new geographical
66 locations or ecological habitats (8, 9). This term has not previously been used to describe
67 subpopulations of *L. monocytogenes*, although the dissemination of persistent strains to more
68 than one food processing facility is a well-documented phenomenon (10-18).

69 In a phylogenetic context, *L. monocytogenes* comprises four separate deep-branching lineages,
70 which are further subdivided into sequence types (STs) and clonal complexes (CCs or clones) by
71 multilocus sequence typing (MLST) (19). Certain clones such as CC1 and CC4 belonging to

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

72 lineage I are commonly associated with clinical disease while others, often belonging to lineage
73 II (e.g., CC9, CC121), are frequently found in food processing environments and food, but rarely
74 among clinical cases (19-21). The underlying causes behind these differences are not fully
75 understood, but thought to be linked to differences in virulence potential and the ability to
76 survive and multiply in food processing environments (11, 22-29). An increased capacity for
77 biofilm formation can contribute to the survival and persistence of *L. monocytogenes* in both
78 natural and food processing environments (3, 30-32). Resistance to stressors encountered in
79 food processing environments, e.g., biocides and alkaline pH, may also contribute to survival.
80 Associated stress resistance determinants can be spread through mobile genetic elements such
81 as plasmids, prophages and transposons (21, 33-37). For example, it has been shown that the
82 presence of *bcrABC* or *qacH* (located on plasmids and transposon Tn*6188*, respectively) results
83 in tolerance to low concentrations of quaternary ammonium compounds (QAC), biocides
84 commonly used in the food industry (26, 36, 38-40). Another genetic determinant associated
85 with CCs commonly found in food processing plants is premature stop codon (PMSC) truncation
86 mutations in *inlA* encoding the virulence factor internalin A (41, 42). Although the ecological
87 significance of *inlA* PMSC mutations is not fully understood, some studies indicate that they
88 mediate increased adhesion and biofilm formation (43-45) and increased tolerance to
89 desiccation (27).

90 There is a consensus that certain *L. monocytogenes* strains are more frequently isolated from
91 food processing factories because of their increased ability to survive and multiply in niches that
92 are difficult to keep clean (2, 7, 28, 46). However, there is no consensus on the operational
93 definition of a persistent strain in terms of the number of independent isolation events or the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

94 time-frame (7, 28). Furthermore, the level of genetic relatedness required to delineate a
95 persistent clone is defined by the resolution of the employed molecular subtyping technique,
96 and with increased sensitivity of subtyping methods, the criteria for defining persistent clones
97 have to be reconsidered. Persistent strains are indistinguishable when characterized with
98 traditional subtyping techniques such as multi locus variable-number tandem repeat analysis
99 (MLVA) and pulsed-field gel electrophoresis (PFGE). These methods have limited resolution as
100 they capture genetic diversity in a small portion of the microbial genome. In contrast, whole
101 genome sequencing (WGS)-based typing strategies determine the diversity across the entire
102 genome and can accurately define genetic distances and differentiate between closely related
103 strains (47, 48). Therefore, WGS-based analysis usually implies setting a threshold of genetic
104 relatedness for identification of clusters or "strains" from the same contamination source. This
105 threshold commonly constitutes 7 to 10 core genome MLST (cgMLST) differences (19, 49, 50),
106 20 single nucleotide polymorphisms (SNPs)(51, 52), or 20 whole genome MLST (wgMLST)
107 differences, as SNP and wgMLST analyses have similar resolution (10, 53, 54). However, these
108 thresholds must be used with caution as bacteria continuously diversify through evolutionary
109 processes. Different outbreak strains and persistent strains will therefore show varying levels of
110 genetic relatedness (47, 53, 55).

111 Interpretation of WGS-based typing results should also consider that highly similar isolates may
112 be found across several food processing facilities (10-18). This may occur due to contamination
113 from a common source of raw materials or pre-processed product (e.g., slaughtered salmon) or
114 through transfer of used processing equipment between food processing plants. In addition,
115 evidence suggests that highly similar *L. monocytogenes* strains may be present in apparently

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

116 unassociated locations, at least in natural environments (2). However, there is limited
117 knowledge regarding the extent to which highly similar genetic clones disseminate or pervade
118 and establish in multiple separate locations. In case of public investigations of listeriosis
119 outbreaks, it is of importance both for public health authorities and for food industry
120 representatives to know more about the prevalence of pervasive strains.

121 The present study aimed to investigate the diversity, distribution, persistence, and stress
122 resistome of *L. monocytogenes* clones dominating in food processing environments in Norway.
123 The analysis comprises 769 *L. monocytogenes* isolates collected during 1990-2020, including 257
124 *L. monocytogenes* belonging to ST8 and ST9 subjected to WGS analysis as part of earlier studies
125 (10, 11). The aims of the study were to (i) assess the genomic diversity of *L. monocytogenes* in
126 Norwegian food chains, (ii) identify persistence, contamination routes, and cases where the
127 same strain is present in more than one factory, (iii) evaluate these aspects in light of the
128 presence of genetic determinants associated with stress survival, antimicrobial resistance and
129 persistence, and (iv) compare the prevalence of genetic determinants of stress survival in the
130 isolates from food processing with that found in Norwegian clinical isolates and environmental
131 isolates from urban and natural locations (2, 56).

132 Results

133 Diversity of isolates from the Norwegian food chain

134 The basis of the current study was a collection of 769 *L. monocytogenes* isolates from the
135 Norwegian food industry from 1990 to 2020 (see Table S1 in the supplemental material). Of
136 these, 441 were from meat processing factories, 19 were from zones with lower hygienic

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

137 conditions associated with meat processing (i.e., animal transport vehicles, animal holding pens,
138 and slaughter departments), and 306 isolates were from salmon processing factories. Samples
139 were mainly from the processing environment (floors, drains, and or food processing
140 equipment) and to a minor degree (n=12) from raw materials and products. All were collected
141 from nine different meat production plants and six different salmon processing plants (Table
142 S2), except for eight isolates from other factories and three isolates from other food associated
143 sources (salmon and cheese product, domestic kitchen). A subset of the isolates belonging to *L.*
144 *monocytogenes* ST8 and ST9 (5 and 252 isolates, respectively) were previously subjected to
145 WGS analysis as part of earlier studies (10, 11). The additional 512 isolates were subjected to
146 WGS, *in silico* MLST, and wgMLST profiling. MLST showed that 13% of the 769 studied isolates
147 belonged to lineage I and the remaining isolates to lineage II, while lineage III or IV isolates were
148 not detected. The isolates were assigned to 33 different STs and 28 different CCs (Fig. 1).

149 Genetic distances obtained from wgMLST analysis were compared with results obtained from a
150 SNP analysis performed separately for each CC using the CFSAN SNP pipeline (57), with
151 reference genomes selected from each CC. The average number of SNPs and wgMLST loci
152 detected within each CC was 126 and 140, respectively. The results show that with default
153 filtering settings, wgMLST analysis was somewhat more sensitive than SNP analysis (Table 1).

154 Presence of plasmids and genetic determinants of stress response and resistance

155 A BLAST analysis was carried out to detect plasmids and genetic determinants associated with
156 stress survival and antimicrobial resistance. Overall, plasmids were identified in 58% of the
157 *L. monocytogenes* isolates. All were *repA*-family theta-replicating plasmids belonging to either
158 group 1 (G1) or group 2 (G2) (58), with one exception. Isolate MF6196 belonging to CC7

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

159 harbored a novel RepA group protein, hereby named RepA group 12 (G12), which was 76%
160 identical to RepA G2 (Table 2). In total, 39% of the isolates harbored *repA* G1 plasmids while
161 18% harbored *repA* G2 plasmids (Table S3). In addition, both G1 and G2 *repA* genes were
162 identified in ten CC5 isolates collected in factory M6 between 2016 and 2019, indicating that
163 they harbored two plasmids. These isolates were related (13-45 wgMLST differences) to three
164 CC5 isolates containing *repA* G1 but not *repA* G2, collected in the same factory during 2010 and
165 2012, suggesting that this strain had acquired a second plasmid harboring a *repA* G2 gene
166 during the intervening years.

167 The search for genetic determinants associated with stress survival, resistance, and persistence
168 identified 76 core genes that were present in all genomes, and 20 accessory genes, gene loci or
169 gene variants that were present in a subset of genomes (Fig. 2A and Table S4). The only
170 identified antibiotic resistance gene, detected by searching the ResFinder database (59), was
171 the fosfomycin resistance gene (*lmo1702*, *fosX_2*) (60), present in all genomes. The accessory
172 genetic determinants comprised cadmium resistance (*cadA1C1*, *cadA2C2*, *cadA4C4*, *cadA5C5*)
173 and arsenic resistance operons (*arsA1D1R1D2R2A2B1B2* on *Listeria* Genomic Island 2 (LGI2) and
174 *arsCBADR* on a Tn*554*-like transposon) (61), QAC resistance loci (chromosomally encoded qacH
175 and plasmid encoded bcrABC (40, 62)), various additional plasmid-associated stress response
176 genes (*clpL*, *mco*, *npr*, a *gbuC*-like gene, a NiCo riboswitch, and *tmr* (63)), genes located on the
177 stress survival islets (SSIs) SSI-1 and SSI-2 (64, 65), biofilm associated genes (*bapL* and *inlL* (66,
178 67)), and PMSC and internal deletion mutations in *inlA* and *inlB* (45). Further details regarding
179 the identified accessory loci found in food environmental isolates are presented in Text S1 in the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

180 supplemental material and Table 2. The unique combinations of accessory genes are presented
181 in Fig. 2A and the complete phylogeny is shown in Fig. S1.

182 For statistical analyses, the accessory stress survival genes were grouped into categories of
183 cadmium resistance, arsenic resistance, QAC resistance, stress survival islets, biofilm-associated
184 genes, and *inlA* PMSC mutations (the identified *inlA* 3-codon deletion (3CD) *inlA* mutation
185 associated with an increased Caco-2 cell invasion phenotype (68, 69) was not included in this
186 category) (Fig. 2B). The prevalence of plasmids and all tested categories of stress survival genes
187 was significantly higher in isolates from hygienic zones in meat processing factories than in
188 salmon processing environments and in low-hygienic zones associated with meat production
189 ($p<0.007$; Table S5). However, it should be emphasized that the high prevalence among isolates
190 from meat processing was associated with their high prevalence within CC9 isolates, which
191 constituted the majority of isolates in this category (n=286; 65%). When CC9 isolates were
192 excluded from the analysis, the occurrence of cadmium and arsenic resistance loci was
193 significantly higher in isolates from salmon processing environments than from hygienic zones in
194 meat processing factories ($p\leq0.001$). In contrast, the occurrence of biofilm-associated genes was
195 significantly higher in isolates from hygienic zones in meat processing factories than from
196 salmon processing environments ($p=0.001$) and low-hygienic zones associated with meat
197 production ($p=0.02$) (Fig. 2B and Table S5).

198 Proportion of food processing plants harbouring the different CCs

199 As the numbers of sequenced isolates from each plant varied widely, ranging from 4 to 192 for
200 the meat processing plants and from 2 to 188 for the salmon processing plants (Table S2), the
201 number of processing plants in which each CC was detected was summarized to assess the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

202 prevalence of different CCs (Fig. 3). Prevalence was not counted as raw number of isolates to
203 reduce bias due to differing number of isolates included from each plant. The clone detected in
204 the greatest number of processing plants overall was CC121 (found in 6/9 meat factories and
205 4/6 salmon factories), followed by CC7, CC8, and CC9. CC9 was detected in the largest number
206 of meat processing plants (7/9).

207 Closely related isolates were present over time within individual factories

208 We next investigated whether persistence of specific strains of *L. monocytogenes* occurred in
209 processing plants, and whether this was associated with certain CCs or the presence of plasmids
210 or stress survival genes. In total, 551 isolates (72%), belonging to 15 different CCs, were linked
211 to persistence. The proportion of persistent isolates was significantly higher in lineage II (74%)
212 than in lineage I (53%) ($p < 0.001$). A persistent isolate was here defined as an isolate that
213 showed 20 or fewer wgMLST allelic differences towards an isolate collected from the same
214 factory in a different calendar year. A persistent strain was defined as a clonal population
215 showing this level of similarity towards at least one other isolate in a cluster found across more
216 than one calendar year in the same factory. The definition of persistence is irrespective of
217 whether the isolates originated from e.g., an established house strain, reintroduction from raw
218 materials or external environment, or from house strains present at a supplier's factory. Clonal
219 clusters of isolates collected within the same calendar year and thus not designated as
220 persistent included 23 isolates belonging to CC3 (0-8 wgMLST differences) collected from
221 factory S6 between January and April of 2020, and 55 CC9 isolates collected during an eight-
222 week period in 2014 at factory M4 in connection with the previously described event related to
223 installation of a contaminated second-hand slicer line (10).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

224 Analysis of the prevalence of the examined categories of genes associated with stress survival
225 showed that QAC resistance genes and biofilm associated genes were more prevalent among
226 persistent than non-persistent strains (p=0.02 and p=0.04, respectively) (Fig. 4A and Table S6).
227 No significant differences were identified between persistent or non-persistent isolates with
228 respect to the presence of plasmids, cadmium and arsenic resistance gene, stress survival islets,
229 or *inlA* PMCS mutations.

230 Previous studies show that the most common sites contaminated with persistent *Listeria* were
231 floors, drains, conveyor belts, slicers, and tables (7). Factories and typical niches where
232 persistent *L. monocytogenes* strains were found are summarized in Table 3, showing that these
233 were largely consistent with previous studies, although isolation of persistent strains on
234 conveyors was only observed in salmon processing plants. In addition, salmon gutting machines
235 were identified as common sites for isolation. The results furthermore show that many different
236 CCs can be associated with persistence. However, some appear to have a greater tendency than
237 others for becoming persistent, e.g., CC7 and CC8, each identified as persistent in four factories,
238 and CC9 and CC121, each in three factories.

239 From the salmon slaughterhouses (S1-S6), persistent strains were identified in ten CCs (11 STs;
240 Table 3). Six of these CCs were present among isolates repeatedly found in S6, where the
241 greatest number of isolates of the same persistent CC121 strain (n=59) was obtained through
242 sampling in the period from February 2019 to April 2020. The emergence of CC121 was followed
243 by extensive sampling in S6 in this period and the CC121 strain was detected in the
244 slaughterhouse processing equipment, machines and environment. Interestingly, the CC121
245 strain was also found onboard a salmon slaughter ship supplying S6 with fresh salmon for

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

246 further processing and on samples from fresh salmon. This indicated that fresh slaughtered
247 salmon contaminated with *L. monocytogenes* was a likely source for the introduction and
248 subsequent persistence of this particular CC121 strain in the plant. Salmon slaughtered in other
249 slaughterhouses and further processed in factory S2 was likely the source of a persistent CC1
250 strain repeatedly and extensively isolated in samples from production environment and
251 equipment in S2 over a two-year period: In a cluster of 30 isolates differing by 0-16 wgMLST
252 alleles, eight isolates were obtained from samples of supplied slaughtered salmon. For salmon
253 processing plant S1, we previously reported isolation of the same CC8 strain ten years apart
254 (11). This factory also harbored a persistent strain of CC7 that was repeatedly isolated from the
255 processing environment and equipment throughout a three-year sampling period. A similar
256 situation was observed in factory S5, with repeated isolation of a CC177 strain over a two-year
257 sampling period. In factory S2, clonal ST732 isolates (CC7) sampled eight years apart (2011-
258 2019) were isolated from various surfaces of equipment and processing environments. These
259 observations indicate that *L. monocytogenes* strains had persisted in the respective
260 slaughterhouses or were repeatedly reintroduced between sampling events during the study
261 period.

262 For the meat processing plants, repeated isolation of the same strain over at least two different
263 years was observed for nine CCs (Table 3). The dominance and persistence of CC9/ST9 over
264 several years in meat processing plants M1 and M4 was previously described (10). In the
265 present study, persistence of a CC9 strain over a two-year period was also confirmed in factory
266 M8. For factory M1, the CC9 strains were repeatedly isolated from the department producing
267 heat-treated products (10), while in raw meat departments, persistent strains were identified

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

268 for CC8, CC19, CC91 and CC415. Closely related isolates of CC19 and CC415 were also repeatedly
269 isolated in factory M4, but in contrast to M1, only in the heat-treated department. Persistent
270 CC7 strains were isolated from floors in poultry processing plants M2 and M6. CC7 was the
271 dominant clonal group in M2, and the only CC from which the same strain was repeatedly
272 isolated in this factory. In factory M6, a persistent strain of CC5 was repeatedly isolated in 2012
273 and during 2016-2019 from floors in a room used for processing of heat-treated products.
274 Closely related CC199 was isolated three times over an eight-month period from factory M8.
275 Two different clusters of CC121 isolates were identified in drains in factory M4, one consisting
276 of two isolates collected three years apart and differing by six wgMLST alleles, and another
277 comprising six isolates differing by 0-6 wgMLST alleles, collected over a period of four years. A
278 cluster of CC121 was also detected in different sites in factory M6, comprising six isolates
279 differing by 3-23 wgMLST alleles over a two-year period. Thus, for several CCs, including CC7,
280 CC9, CC19 and CC121, persistent strains were repeatedly isolated from more than one meat
281 processing factory, while some CCs (e.g., CC5, CC199) were repeatedly isolated only at single
282 plants.

283 ## Pervasive strains: The same strain found in more than one food processing plant

284 Phylogenetic analysis revealed that many isolates were closely related despite being isolated
285 from different processing plants (Fig. 1). We hereby designate the observation of clonal
286 populations of *L. monocytogenes* found in more than one factory as pervasion. The definition
287 does not differentiate between the mode of dissemination between factories and includes
288 isolates with a common source and ancestor. In total, 433 of the isolates (56%) were pervasive,
289 showing 20 or fewer wgMLST allelic differences towards an isolate found in a different factory.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

290 A pervasive strain was defined as the isolates belonging to such a cluster. The proportion of
291 pervasive isolates differed between food source, with 70% of isolates from meat processing
292 environments designated as pervasive, in contrast to only 40 % of isolates from salmon
293 processing industry. The proportion of pervasive isolates was significantly higher in lineage II
294 (59%) than in lineage I (34%) (p<0.001).

295 With two exceptions, pervasive strains were either shared between meat processing factories or
296 between salmon processing factories (Fig. 5A and 5B). Both exceptions involved salmon factory
297 S1. The first case involved CC7 and a cluster of 14 isolates collected at S1 between 2011 and
298 2014 and a cluster of 14 isolates from meat factory M6 collected in 2004 and 2011, differing by
299 15-47 allelic differences (median 22). In the second case, a CC8 isolate from a floor sample in
300 the raw meat zone in factory M1 collected in 2019 differed by 11-17 wgMLST alleles to a cluster
301 of five isolates from S1, one from 2001 and four from 2011.

302 Factories S2 and S6 were the two most heavily sampled salmon slaughterhouses. The two
303 factories are located in different geographic regions in Norway and belong to different
304 companies. WGS was performed for 68 isolates collected during 2011-2012 and 2018-2019 from
305 S2, and for 188 isolates collected between June 2017 and April 2020 from S6. We identified
306 three different CCs (CC1, CC121, and CC88) in which closely related isolates (≤17 wgMLST allelic
307 differences) were found in both factories. In the case of CC1, one isolate from a product sample
308 collected in October 2019 from factory S6 (MF7739) differed by 7-15 wgMLST alleles from the
309 previously mentioned persistent strain comprising 30 environmental isolates from factory S2,
310 for which the likely source was slaughtered salmon. These were obtained between July 2018
311 and May 2019. Also, one CC121 isolate collected from a drain at factory S2 in 2012 (MF4804)

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

312 showed 5-11 wgMLST allelic differences towards the previously mentioned cluster of 59
313 persistent CC121 isolates collected in factory S6 during 2019-2020, which included one isolate
314 collected onboard the slaughter ship. However, the slaughter ship had not been in operation in
315 2012 when the S2 isolate was detected. The third case linking these two factories were two
316 isolates separated by 17 wgMLST alleles belonging to CC88, obtained from the boots of a factory
317 worker at S2 in November 2018 and from salmon from an external supplier at factory S6 one
318 year later, in November 2019. Genetically similar isolates were also found at factories S2 and S3,
319 where two CC14 isolates collected in S2 in 2011-2012 showed 12-28 wgMLST allelic differences
320 towards a cluster of 13 isolates from S3 collected during 2018-2019.

321 In the meat industry, the majority of isolates from pervasive strains belonged to CC9, for which
322 we previously described a close genetic relationship between *L. monocytogenes* from four meat
323 processing plants (M1, M4, M5 and M7) collected during 2009-2017 (10). In the current study,
324 38 additional CC9 isolates, including representatives from three additional processing plants
325 (M3, M8, and M9) were sequenced (Fig. 5C). Of particular interest was a group of CC9 isolates
326 collected in 1998 and 2001 from the raw side of a meat factory (M3) that is no longer in
327 operation (70). One of the isolates from 1998 differed by only 5-8 wgMLST allelic differences to
328 isolates from five other processing plants (M1, M4, M5, M7, and M8), suggesting that this CC9
329 strain has circulated in Norwegian meat chains for at least two decades.

330 Pervasion was significantly associated with presence of stress survival genes

331 In contrast to persistent strains (Fig. 4A), isolates identified as pervaders showed significantly
332 increased prevalence of plasmids and all examined categories of stress and persistence
333 associated genes ($p<0.001$) (Fig. 4B and Table S6). The prevalence of individual gene variants

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

334 (*clpL*, *cadA1C1*, *cadA2C2*, *bcrABC*, *qacH,* the *ars* operon on LGI2, *arsCBADR* on Tn*554*, SSI-1, SSI-
335 2, *bapL*, and *inlL*) was also significantly higher in isolates classified as pervaders compared to
336 non-pervaders ($p \leq 0.004$) (Table S6).

337 Of the pervasive isolates, 81% (351/433) were also persistent. The 82 pervaders classified as
338 non-persistent (11% of the total number of isolates) included the 55 CC9 isolates associated
339 with the previously described contaminated second-hand slicer line collected from factory M4
340 during an eight-week period in 2014 (10). Only 63% (351/551) of the persistent isolates were
341 also pervaders. If the CC9 isolates were excluded from the analysis, 92% of pervaders were also
342 persistent, only 51% of persistent isolates were also pervasive, and only 3% of pervasive isolates
343 were classified as non-persistent.

344 ## Stress survival determinants in clinical and environmental isolates

345 To compare the prevalence of plasmids and stress survival genes in isolates from food
346 processing environments to human clinical isolates and isolates from natural environments in
347 Norway, the BLAST analysis was also performed for the genomes of 111 Norwegian clinical
348 isolates and 218 isolates from rural, urban, and farm environments in Norway, which were
349 examined in a recent study (2). The prevalence of plasmids was significantly different between
350 isolates from the three sources ($p<0.001$; Table S7), with 26% and 8% of clinical and
351 environmental isolates harbouring plasmids, respectively, compared to 41% for the isolates
352 from food processing environments (Table S3). The lowest prevalence of plasmids was found in
353 the 87 isolates from dairy farms (5%). In addition to *repA* G1 and *repA* G2 plasmids, one *repA*
354 group 4 (G4) plasmid and two small non-*repA* plasmids were identified among the clinical
355 isolates (Table 2 and Text S1).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

356 The clinical and natural environment isolates harboured the same core and accessory stress
357 genes as the isolates from food processing environments with three exceptions: i) one clinical
358 isolate harboured a chromosomally encoded *tet*(M) tetracycline resistance gene, identified
359 through search against the ResFinder database (59), ii) one clinical isolate belonging to CC7
360 harboured the *emrC* gene conferring QAC resistance (71), and iii) none of the isolates harboured
361 internal deletion mutations in *inlB*. All unique combinations of the accessory gene subset for
362 clinical and environmental isolates are presented in Fig. 6A and 6B and further discussed in Text
363 S1. The prevalence of all examined categories of stress survival and resistance genes was
364 highest in the isolates from food processing environments (p<0.003; Table S7) (Fig. 6C).
365 Comparison of the clinical and environmental isolates showed a higher prevalence among
366 clinical isolates in all categories of stress survival determinants (p<0.006), with the exception of
367 biofilm associated (p=0.688) and arsenic resistance genes (p=0.131). Of note, none of the 218
368 isolates collected from rural, urban, and farm environments harbored the *bcrABC* or *qacH* QAC
369 resistance genes or *inlA* PMSC mutations.

370 Within the 1098 examined *L. monocytogenes* genomes, the prevalence of *inlA* PMSC truncation
371 mutations was 9% and 41% in lineages I and II, respectively. In contrast, the 3CD *inlA* genotype
372 associated with increased invasion (68, 69) showed an occurrence of 14% in lineage I, but was
373 detected only once among the lineage II isolates (in CC89) (Fig. 2A and Fig. 6AB).

374 Discussion

375 To gain a better understanding of the population structure and genomic diversity of *L.*
376 *monocytogenes* in Norwegian food chains, genome sequences from 769 *L. monocytogenes*

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

377 isolates from the Norwegian food industry collected over three decades from 15 food
378 processing factories were characterized using WGS-based comparative genomic analyses. The
379 study showed that 56% of isolates were closely related (2-20 wgMLST allelic differences) to an
380 isolate collected from a different factory. These isolates were designated as *pervasive*, a term
381 which has previously been used to describe subpopulations of bacteria with enhanced ability to
382 spread or migrate to new geographical locations (8, 9). Several other studies have similarly
383 reported that *L. monocytogenes* isolates from geographically and temporally unrelated sources
384 were separated by equally short genetic distances (2, 12-18). WGS analyses are increasingly
385 used in epidemiology and have in recent years been essential for solving several European
386 foodborne listeriosis outbreaks (72-75). However, as we know from the use of DNA as forensic
387 evidence in criminal trials, the apparent certainty of DNA evidence can be deceptive, and there
388 is a danger that the statistical significance of a DNA match can be overstated (76, 77). This is
389 particularly likely in the case of bacteria which reproduce using binary fission and especially for
390 *L. monocytogenes* which has an extremely low evolutionary rate (16, 19). Indeed, in one case
391 described by Lüth *et al*. (18), the same *L. monocytogenes* strain identified in two different
392 processing plants matched the same CC5 outbreak cluster. There is a risk that authorities could
393 mistakenly consider a WGS match between two bacterial isolates as proof of identification of a
394 contamination source, also in cases lacking other epidemiological evidence. It is therefore of
395 crucial importance to consider the possibility of highly similar isolates being found in multiple
396 factories when WGS analyses of *L. monocytogenes* are performed, both for outbreak
397 investigations and for food safety risk-based decisions and risk assessment in food industry. The
398 high prevalence of *L. monocytogenes* isolates identified in more than one processing plant in the

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

399 current study suggests that the occurrence of pervasive strains in the Norwegian food industry –
400 particularly in the meat distribution chain where 70% of isolates were pervasive – may be
401 significantly higher than in other countries. This possibly reflects a particularly complex and
402 interconnected Norwegian meat supply chain. Regardless, mistaken identification of an
403 outbreak source can have enormous economic impacts and cause significant food waste due to
404 unnecessary recalls. Examples include the German 2011 outbreak of enterohaemorrhagic
405 *Escherichia coli* (EHEC) where initially cucumbers imported from Spain were erroneously
406 implicated as the source of the infections (78, 79), and the Norwegian 2006 EHEC outbreak
407 caused by contaminated traditional cured sausage ("morrpølse"), in which minced meat was
408 initially indicated as the source (80).

409 The prevalence of genetic determinants associated with stress survival, metal and biocide
410 resistance and biofilm formation, as well as PMSC mutations in *inlA*, was higher in isolates
411 collected from food processing environments than among clinical isolates and isolates from
412 natural environments, thus supporting previous studies suggesting that these factors are
413 involved in the adaptation to food processing environments (21, 35, 39, 64, 69, 81, 82). Notably,
414 none of the 218 isolates from natural environments contained biocide resistance genes *qacH*,
415 *bcrABC*, or *inlA* PMSC mutations. This concurs with a recent study of *L. monocytogenes* collected
416 from surface waters in California (83), in which *qacH* and *bcrABC* were detected in zero and 18
417 isolates respectively, while *inlA* PMSC mutations were present in only four of 1248 examined
418 isolates. In contrast, we identified *qacH* or *bcrABC* genes in 45% and *inlA* PMSC mutations in
419 51% of isolates from food processing environments. These values coincide with data from
420 numerous previous studies (26, 28, 68, 81, 82). Antibiotic resistance determinants present on

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

421 mobile genetic elements appear to have low prevalence in *L. monocytogenes* isolated in
422 Norway, as only one such antibiotic resistance gene was detected among the 1098 examined
423 genomes (a *tet*(M) tetracycline resistance gene in a clinical isolate). This is lower than that
424 recently reported among isolates from other countries (84-89).

425 It is acknowledged that the distribution and abundance of *L. monocytogenes* CCs varies between
426 different environments, such as humans, animals, soil, water, plants, and various types of food,
427 and that this is likely driven by selective adaptation (2, 22, 24, 29, 90-93). However, the same is
428 not always true when comparing persistent and transient isolates from food processing
429 environments, although it is highly likely that inheritable genetic traits are responsible also for
430 the ability to survive long-term in food processing facilities. It is frequently reported that certain
431 CCs have been identified as persistent in food processing environments, notably CC9 and CC121,
432 however, persistent strains have been detected in many CCs including CC5, CC7, CC8, CC31,
433 CC155, and CC321 (14, 16, 18, 21, 26, 94-99). This largely corresponds with observations from
434 the current study, in which CC7, CC8, CC9 and CC121 were identified as persistent in the
435 greatest number of processing plants. Nevertheless, persistent strains were identified within 15
436 of the 25 CCs represented by more than one isolate (60%), including in three lineage I CCs and
437 12 lineage II CCs.

438 It is likely that the genetic basis behind increased prevalence of certain *L. monocytogenes* in
439 food processing environments aligns with the genetic basis behind persistence in the ecological
440 sense of the word, i.e., an increased ability to survive long-term in food processing
441 environments (5, 6). There is some evidence of association of persistence with the presence of
442 the *bcrABC* cassette conferring resistance to QAC disinfection agents (26, 39, 100) and perhaps

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

443 also biofilm formation capacity (30, 99). However, several studies have failed to identify any
444 phenotypic or genotypic differences between persistent and transient *L. monocytogenes* (6, 95,
445 101, 102) or failed to associate persistent strains with differences in stress response, sanitizer
446 resistance, or adhesion properties (16, 25, 32, 70, 102-104). In line with these reports, the
447 current study showed a relatively weak statistically significant association ($p \leq 0.04$) between
448 persistence and QAC resistance determinants and biofilm-associated genes, and none between
449 persistence and plasmids, heavy metal resistance genes, SSIs, or *inlA* PMSC mutations. In
450 contrast, we found that pervasive isolates, belonging to strains present in more than one
451 factory, showed strong statistically significant association ($p<0.001$) with all examined
452 categories of stress survival and persistence genes, as well as towards plasmids and several
453 individual genes or gene loci. Pervasive strains were identified in eight CCs: one belonging to
454 lineage I and seven belonging to lineage II. Notably, only a small proportion of pervaders were
455 classified as non-persistent (i.e., transient). It thus appeared that strains that occurred at several
456 factories and were repeatedly isolated over time in one or more of these facilities were more
457 likely to show problematic properties, i.e., carry genetic determinants that enable them to
458 establish as house strains and disperse to new environments. Our results furthermore support
459 the hypothesis that there is not one single genetic determinant responsible for survival in food
460 processing environments, but rather an accumulation of stress resistance genes, biofilm
461 associated genes, and *inlA* PMSC mutations.

462 The identity of the isolates classified as non-persistent pervaders in the current study exemplify
463 one of the difficulties in separating between "true" persistent and "true" transient isolates using
464 operational definitions. These included 82 isolates of which 55 were from the previously

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

465 described contaminated second-hand slicer line isolated from factory M4 (10). These were not
466 defined as persistent, as the contamination was only detected during an eight-week period in
467 2014 in which the slicer line was installed in the factory. However, from an ecological viewpoint,
468 they obviously belonged to a house strain established in a difficult to clean niche (5, 6). Another
469 challenge with operational definitions of persistence is that strains without any specific adaptive
470 features responsible for increased survival in processing environments may survive there under
471 permissive conditions, e.g., in a period with higher temperatures or inadequate cleaning and
472 sanitation, or alternatively, reoccurrence can be due to repeated introduction from an outside
473 reservoir (7, 46). Therefore, even when using a sampling method targeted towards house
474 strains, including sampling after cleaning and disinfection and detection of reoccurrence over a
475 longer time-period, the obtained isolates may not carry specific genetic determinants for
476 survival in the factory environment. The results from the current study suggest that the
477 operational definition of pervasion is superior to those used to define persistence in identifying
478 strains that carry adaptations responsible for increased ability to survive and multiply in food
479 processing environments.

480 Material and methods

481 Source of isolates

482 The isolates from food processing environment, raw materials and processed foods included in
483 the study are listed in Table S1 and were from the *L. monocytogenes* strain collection at Nofima,
484 Norway. A total of 305 isolates from both meat and salmon industry were collected during
485 2011-2015 as part of a previous study (38). The majority of these were isolated after sanitation,

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

486 and before start of production. Most of the isolates obtained after 2016 were from the
487 factories' own sampling programs, and for these, sampling was mainly performed during
488 production. A subset of 252 isolates belonging to CC8 and CC9 has been described in previous
489 studies (10, 11).

490 WGS and genome assembly

491 Bacteria were grown on BHI agar overnight at 37 °C before a loopful of cells was suspended in
492 500 µl 2x Tris-EDTA buffer with 1.2 % Triton X-100. Cells were lysed using lysing matrix B and a
493 FastPrep instrument (both MP Biomedicals), and genomic DNA was isolated using the DNeasy
494 blood and tissue kit (Qiagen). Libraries were prepared using the Nextera XT DNA sample
495 preparation kit or the Nextera DNA Flex Library prep kit (both Illumina) and sequenced on a
496 MiSeq platform with 300-bp paired-end reads. Raw reads were filtered on q15 and trimmed of
497 adaptors before *de novo* genome assembly was performed using SPAdes v3.10.0 or v.3.13.0
498 (105) with the careful option and six k-mer sizes (21, 33, 55, 77, 99, and 127). Contigs with sizes
499 of <500 bp and *k*-mer coverage of <5 were removed from the assemblies (a coverage cutoff 15
500 was used for MF7896). The average coverage for the genome assemblies was calculated using
501 BBmap v36.92 (106). The quality of all assemblies was evaluated using QUAST v5.0.2 (107).

502 MLST analyses

503 Classical MLST analysis followed the MLST scheme described by Ragon *et al*. (42) and the
504 database maintained at the Institute Pasteur's *L. monocytogenes* online MLST repository
505 (http://bigsdb.web.pasteur.fr/listeria/listeria.html). CC14 and CC91 were defined as previously
506 described (2). The wgMLST analysis was performed using a whole-genome scheme containing

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

507 4797 coding loci from the *L. monocytogenes* pan-genome and the assembly-based BLAST
508 approach, implemented in BioNumerics 7.6 (http://www.applied-maths.com/news/listeria-
509 monocytogenes-whole-genome-sequence-typing).

510 Minimum spanning trees were constructed using BioNumerics based on the categorical
511 differences in the allelic wgMLST profiles for each isolate. Loci with no allele calls were not
512 considered in the pairwise comparison between two genomes. The number of allelic differences
513 between isolates was read from genetic distance matrices computed from the absolute number
514 of categorical differences between genomes.

515 Calculation of pairwise wgMLST distances for Neighbor-Joining (NJ) trees was performed using
516 the daisy function (108) from the cluster package v2.1.1 (109) in R and selection of the gower
517 metric. NJ trees were generated using an improved version of the NJ algorithm (BION) (110)
518 implemented in the ape package v5.4-1 (111) in R v4.0.4 (112) as function bionjs. Interactive
519 Tree Of Life (iTOL) v6.5.2 (113) was used for visualization.

520 SNP analysis

521 Read mapping based SNP analysis was performed separately for each CC. An internal reference
522 genome was selected from each CC (listed in Table S1) using the following criteria: Centrally
523 positioned in CC clusters by wgMLST analysis, from larger sub-clusters if more than one cluster
524 was present in a CC, older isolates preferred to more recent isolates, higher quality assemblies
525 (high coverage, few contigs) preferred. The reference based SNP analysis was performed using
526 the CFSAN SNP pipeline v2.1.1 (57) and default filtering settings, except that regions of high-
527 density SNPs were defined for each sample individually instead of filtering a dense region found

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

528 in any genome from all genomes. The default filtering removes SNPs located closer than 500 bp
529 to the end of a contig on the reference genome, SNPs where there are more than 3 SNPs in a
530 1000 base window, more than 2 SNPs in a 125 base window, and more than 1 SNP in a 15 base
531 window, reads with map quality below than 30, reads with a base quality of at least 15 at a
532 given position, SNPs where less than 90% of the calls agree, and SNPs with a read coverage of
533 less than 5. Results were read from the output matrix of pairwise SNP distances.

534 ## BLAST analyses

535 The selection of plasmids included in the current analysis was based on a study by Chmielowska
536 *et al*. (58), which included 113 unique completely sequenced plasmid replicons from *Listeria*
537 spp. strains, of which 63 were assigned to 19 groups (different only by point mutations or small
538 indels of size <1 kb). In the current analysis, one plasmid each from the 19 groups (selecting the
539 largest plasmid in each group) and the 50 plasmids that did not belong to a group were
540 included, in total 69 different plasmids (Table S8). All contigs >2000 bp from the *L.*
541 *monocytogenes* genomes were used as queries in a BLAST search (blastn v2.10.0+) against a
542 local nucleotide BLAST database created for the 69 *Listeria* spp. plasmids. A contig was viewed
543 as a plasmid contig if the query coverage (qcov) was ≥85%, and the percentage identity of the
544 best alignment was >90%. Hits were inspected manually.

545 A BLAST search for plasmid replication genes was carried out using one *repA* gene selected from
546 each of the eleven RepA groups (G1-G11) identified by Chmielowska *et al*. (58) as queries (Table
547 S9), against a local nucleotide BLAST database created for the *L. monocytogenes* genomes. The
548 same analysis was also performed using as queries the nine plasmid-encoded/associated
549 genetic elements involved in stress response identified by Schmitz-Esser *et al*. (63), as well as

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

550 additional genetic elements associated with stress response and resistance (Table S4) and the
551 entire content of the ResFinder database (59) downloaded on November 11, 2021. Only the
552 best hit for each query sequence in each genome was kept. When the minimum nucleotide
553 identity was <99%, and/or the length ratio of the query sequence relative to the match in the
554 genome (length/qlen) was ≠ 1, the alignments and contigs were manually inspected before a
555 presence/absence gene call was made. In addition, the protein sequences of the BLAST hits
556 were aligned to the query protein sequences. A call was made for the presence of the gene if
557 the protein identity was >92%. Differences in the nucleotide sequences that lead to alterations
558 of the protein sequences, such as insertions, deletions, premature stop codons or elongations,
559 were recorded.

560 The Pearson's chi-squared association test was performed using Minitab v.19.2020.1 to
561 determine whether there was any statistically significant association between the presence or
562 absence of stress survival genes (or groups of genes) or lineage and the source of the isolates
563 (e.g., meat vs. salmon, clinical vs. food processing) or classification as persistent or pervasive. All
564 tests were performed separately for pairs of categories.

565 Calculation of weighted similarity scores

566 Genetic associations between pervasive isolates from different factories, weighted by their
567 similarity, and illustrated in the chord diagram in Fig. 5B were calculated as follows: All pairs of
568 isolates from different factories separated by 20 or fewer wgMLST allelic differences were
569 counted towards the total strength of links between factories. For these pairs, the genetic
570 distance D was converted to a similarity score $S = 21 - D$. Thus, a distance of 20 wgMLST
571 alleles corresponds to a similarity score of 1, and a distance of 0 wgMLST alleles corresponds to

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

572 a similarity score of 21. Of note, the lowest genetic distance separating two isolates from
573 different factories was 2. Then, all similarity scores were grouped by CC and factory and
574 summed to generate the final scores, which are represented by the thickness of each arc in Fig.
575 5B. The image was created using the R package circlize (114).

576 ## Data availability

577 The raw data and assembled genomes for the 512 genomes sequenced in the current study has
578 been submitted to NCBI as BioProject accession PRJNA689484. For GenBank and Sequence Read
579 Archive (SRA) accession numbers, see Table S1. The assemblies were annotated using the NCBI
580 Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) server.

581 # Acknowledgments

582 The authors wish to thank Merete Rusås Jensen and Anette Wold Åsli at Nofima for excellent
583 technical assistance. This work was supported by the Research Council of Norway through grant
584 no. 294910 and the Norwegian Fund for Research Fees for Agricultural Products (FFL) through
585 grant no. 314743.

586

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

587 ## References

588 1. **Vivant AL, Garmyn D, Piveteau P.** 2013. *Listeria monocytogenes*, a down-to-earth
589 pathogen. Front Cell Infect Microbiol **3:**87.

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

594 3. **Colagiorgi A, Bruini I, Di Ciccio PA, Zanardi E, Ghidini S, Ianieri A.** 2017. *Listeria*
595 *monocytogenes* biofilms in the wonderland of food industry. Pathogens **6:**41.

596 4. **NicAogáin K, O'Byrne CP.** 2016. The role of stress and stress adaptations in determining
597 the fate of the bacterial pathogen *Listeria monocytogenes* in the food chain. Front
598 Microbiol **7:**1865.

599 5. **Ferreira V, Wiedmann M, Teixeira P, Stasiewicz MJ.** 2014. *Listeria monocytogenes*
600 persistence in food-associated environments: Epidemiology, strain characteristics, and
601 implications for public health. J Food Prot **77:**150-170.

602 6. **Carpentier B, Cerf O.** 2011. Review - Persistence of *Listeria monocytogenes* in food
603 industry equipment and premises. International Journal of Food Microbiology **145:**1-8.

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

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

607 8. **Numminen E, Gutmann M, Shubin M, Marttinen P, Meric G, van Schaik W, Coque TM,**
608 **Baquero F, Willems RJ, Sheppard SK, Feil EJ, Hanage WP, Corander J.** 2016. The impact of
609 host metapopulation structure on the population genetics of colonizing bacteria. J Theor
610 Biol **396:**53-62.

611 9. **Johnston ER, Rodriguez RL, Luo C, Yuan MM, Wu L, He Z, Schuur EA, Luo Y, Tiedje JM,**
612 **Zhou J, Konstantinidis KT.** 2016. Metagenomics reveals pervasive bacterial populations
613 and reduced community diversity across the Alaska tundra ecosystem. Front Microbiol
614 **7:**579.

615 10. **Fagerlund A, Langsrud S, Møretrø T.** 2020. In-depth longitudinal study of *Listeria*
616 *monocytogenes* ST9 isolates from the meat processing industry: Resolving diversity and
617 transmission patterns using whole-genome sequencing. Appl Environ Microbiol
618 **86:**e00579-20.

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

622 12. **Stasiewicz MJ, Oliver HF, Wiedmann M, den Bakker HC.** 2015. Whole-genome sequencing
623 allows for improved identification of persistent *Listeria monocytogenes* in food-associated
624 environments. Appl Environ Microbiol **81:**6024-6037.

625 13. **Morganti M, Scaltriti E, Cozzolino P, Bolzoni L, Casadei G, Pierantoni M, Foni E, Pongolini**
626 **S.** 2016. Processing-dependent and clonal contamination patterns of *Listeria*

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

627 *monocytogenes* in the cured ham food chain revealed by genetic analysis. Appl Environ
628 Microbiol **82:**822-831.

629 14. **Centorotola G, Guidi F, D'Aurizio G, Salini R, Di Domenico M, Ottaviani D, Petruzzelli A,**
630 **Fisichella S, Duranti A, Tonucci F, Acciari VA, Torresi M, Pomilio F, Blasi G.** 2021. Intensive
631 environmental surveillance plan for *Listeria monocytogenes* in food producing plants and
632 retail stores of central Italy: Prevalence and genetic diversity. Foods **10:**1944.

633 15. **Holch A, Webb K, Lukjancenko O, Ussery D, Rosenthal BM, Gram L.** 2013. Genome
634 sequencing identifies two nearly unchanged strains of persistent *Listeria monocytogenes*
635 isolated at two different fish processing plants sampled 6 years apart. Appl Environ
636 Microbiol **79:**2944-2951.

637 16. **Knudsen GM, Nielsen JB, Marvig RL, Ng Y, Worning P, Westh H, Gram L.** 2017. Genome-
638 wide-analyses of *Listeria monocytogenes* from food-processing plants reveal clonal
639 diversity and date the emergence of persisting sequence types. Environ Microbiol Rep
640 **9:**428-440.

641 17. **Lundén JM, Autio TJ, Korkeala HJ.** 2002. Transfer of persistent *Listeria monocytogenes*
642 contamination between food-processing plants associated with a dicing machine. J Food
643 Prot **65:**1129-1133.

644 18. **Lüth S, Halbedel S, Rosner B, Wilking H, Holzer A, Roedel A, Dieckmann R, Vincze S,**
645 **Prager R, Flieger A, Al Dahouk S, Kleta S.** 2020. Backtracking and forward checking of
646 human listeriosis clusters identified a multiclonal outbreak linked to *Listeria*

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

647 *monocytogenes* in meat products of a single producer. Emerg Microbes Infect **9:**1600-
648 1608.

649 19. **Moura A, Criscuolo A, Pouseele H, Maury MM, Leclercq A, Tarr C, Björkman JT, Dallman**
650 **T, Reimer A, Enouf V, Larsonneur E, Carleton H, Bracq-Dieye H, Katz LS, Jones L, Touchon**
651 **M, Tourdjman M, Walker M, Stroika S, Cantinelli T, Chenal-Francisque V, Kucerova Z,**
652 **Rocha EPC, Nadon C, Grant K, Nielsen EM, Pot B, Gerner-Smidt P, Lecuit M, Brisse S.**
653 2016. Whole genome-based population biology and epidemiological surveillance of
654 *Listeria monocytogenes*. Nat Microbiol **2:**16185.

655 20. **Gray MJ, Zadoks RN, Fortes ED, Dogan B, Cai S, Chen Y, Scott VN, Gombas DE, Boor KJ,**
656 **Wiedmann M.** 2004. *Listeria monocytogenes* isolates from foods and humans form
657 distinct but overlapping populations. Appl Environ Microbiol **70:**5833-5841.

658 21. **Palma F, Brauge T, Radomski N, Mallet L, Felten A, Mistou MY, Brisabois A, Guillier L,**
659 **Midelet-Bourdin G.** 2020. Dynamics of mobile genetic elements of *Listeria monocytogenes*
660 persisting in ready-to-eat seafood processing plants in France. BMC Genom **21:**130.

661 22. **Maury MM, Bracq-Dieye H, Huang L, Vales G, Lavina M, Thouvenot P, Disson O, Leclercq**
662 **A, Brisse S, Lecuit M.** 2019. Hypervirulent *Listeria monocytogenes* clones' adaptation to
663 mammalian gut accounts for their association with dairy products. Nat Commun **10:**2488.

664 23. **Bergholz TM, Shah MK, Burall LS, Rakic-Martinez M, Datta AR.** 2018. Genomic and
665 phenotypic diversity of *Listeria monocytogenes* clonal complexes associated with human
666 listeriosis. Appl Microbiol Biotechnol **102:**3475-3485.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

667 24. **Félix B, Feurer C, Maillet A, Guillier L, Boscher E, Kerouanton A, Denis M, Roussel S.**
668 2018. Population genetic structure of *Listeria monocytogenes* strains isolated from the pig
669 and pork production chain in France. Front Microbiol **9:**684.

670 25. **Gray J, Chandry PS, Kaur M, Kocharunchitt C, Fanning S, Bowman JP, Fox EM.** 2021.
671 Colonisation dynamics of *Listeria monocytogenes* strains isolated from food production
672 environments. Sci Rep **11:**12195.

673 26. **Cherifi T, Carrillo C, Lambert D, Miniaï I, Quessy S, Larivière-Gauthier G, Blais B, Fravalo**
674 **P.** 2018. Genomic characterization of *Listeria monocytogenes* isolates reveals that their
675 persistence in a pig slaughterhouse is linked to the presence of benzalkonium chloride
676 resistance genes. BMC Microbiol **18:**220.

677 27. **Hingston P, Chen J, Dhillon BK, Laing C, Bertelli C, Gannon V, Tasara T, Allen K, Brinkman**
678 **FS, Truelstrup Hansen L, Wang S.** 2017. Genotypes associated with *Listeria*
679 *monocytogenes* isolates displaying impaired or enhanced tolerances to cold, salt, acid, or
680 desiccation stress. Front Microbiol **8:**369.

681 28. **Unrath N, McCabe E, Macori G, Fanning S.** 2021. Application of whole genome
682 sequencing to aid in deciphering the persistence potential of *Listeria monocytogenes* in
683 food production environments. Microorganisms **9:**1856.

684 29. **Disson O, Moura A, Lecuit M.** 2021. Making sense of the biodiversity and virulence of
685 *Listeria monocytogenes*. Trends Microbiol **29:**811-822.

686 30. **Møretrø T, Langsrud S.** 2004. *Listeria monocytogenes*: Biofilm formation and persistence
687 in food processing environments. Biofilms **1:**107-121.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

688 31. **Rodríguez-López P, Rodríguez-Herrera JJ, Vázquez-Sánchez D, Lopez Cabo M.** 2018.
689 Current knowledge on *Listeria monocytogenes* biofilms in food-related environments:
690 Incidence, resistance to biocides, ecology and biocontrol. Foods **7**.

691 32. **Nowak J, Cruz CD, Tempelaars M, Abee T, van Vliet AHM, Fletcher GC, Hedderley D,**
692 **Palmer J, Flint S.** 2017. Persistent *Listeria monocytogenes* strains isolated from mussel
693 production facilities form more biofilm but are not linked to specific genetic markers. Int J
694 Food Microbiol **256:**45-53.

695 33. **Naditz AL, Dzieciol M, Wagner M, Schmitz-Esser S.** 2019. Plasmids contribute to food
696 processing environment-associated stress survival in three *Listeria monocytogenes* ST121,
697 ST8, and ST5 strains. Int J Food Microbiol **299:**39-46.

698 34. **Minarovičová J, Véghová A, Mikulášová M, Chovanová R, Šoltýs K, Drahovská H,**
699 **Kaclíková E.** 2018. Benzalkonium chloride tolerance of *Listeria monocytogenes* strains
700 isolated from a meat processing facility is related to presence of plasmid-borne *bcrABC*
701 cassette. Antonie Van Leeuwenhoek **111:**1913-1923.

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

705 36. **Müller A, Rychli K, Zaiser A, Wieser C, Wagner M, Schmitz-Esser S.** 2014. The *Listeria*
706 *monocytogenes* transposon Tn*6188* provides increased tolerance to various quaternary
707 ammonium compounds and ethidium bromide. FEMS Microbiol Lett **361:**166-173.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

708 37. **Verghese B, Lok M, Wen J, Alessandria V, Chen Y, Kathariou S, Knabel S.** 2011. *comK*
709 prophage junction fragments as markers for *Listeria monocytogenes* genotypes unique to
710 individual meat and poultry processing plants and a model for rapid niche-specific
711 adaptation, biofilm formation, and persistence. Appl Environ Microbiol **77:**3279-3292.

712 38. **Møretrø T, Schirmer BCT, Heir E, Fagerlund A, Hjemli P, Langsrud S.** 2017. Tolerance to
713 quaternary ammonium compound disinfectants may enhance growth of *Listeria*
714 *monocytogenes* in the food industry. Int J Food Microbiol **241:**215-224.

715 39. **Martínez-Suárez JV, Ortiz S, López-Alonso V.** 2016. Potential impact of the resistance to
716 quaternary ammonium disinfectants on the persistence of *Listeria monocytogenes* in food
717 processing environments. Front Microbiol **7:**638.

718 40. **Dutta V, Elhanafi D, Kathariou S.** 2013. Conservation and distribution of the benzalkonium
719 chloride resistance cassette *bcrABC* in *Listeria monocytogenes*. Appl Environ Microbiol
720 **79:**6067-6074.

721 41. **Nightingale KK, Windham K, Martin KE, Yeung M, Wiedmann M.** 2005. Select *Listeria*
722 *monocytogenes* subtypes commonly found in foods carry distinct nonsense mutations in
723 *inlA*, leading to expression of truncated and secreted internalin A, and are associated with
724 a reduced invasion phenotype for human intestinal epithelial cells. Appl Environ Microbiol
725 **71:**8764-8772.

726 42. **Ragon M, Wirth T, Hollandt F, Lavenir R, Lecuit M, Le Monnier A, Brisse S.** 2008. A new
727 perspective on *Listeria monocytogenes* evolution. PLoS Pathog **4:**e1000146.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

728 43. **Mahoney DBJ, Falardeau J, Hingston P, Chmielowska C, Carroll LM, Wiedmann M, Jang**
729 **SS, Wang S.** 2022. Associations between *Listeria monocytogenes* genomic characteristics
730 and adhesion to polystyrene at 8 °C. Food Microbiol **102:**103915.

731 44. **Franciosa G, Maugliani A, Scalfaro C, Floridi F, Aureli P.** 2009. Expression of internalin A
732 and biofilm formation among *Listeria monocytogenes* clinical isolates. Int J Immunopathol
733 Pharmacol **22:**183-193.

734 45. **Piercey MJ, Hingston PA, Truelstrup Hansen L.** 2016. Genes involved in *Listeria*
735 *monocytogenes* biofilm formation at a simulated food processing plant temperature of 15
736 °C. Int J Food Microbiol **223:**63-74.

737 46. **Spanu C, Jordan K.** 2020. *Listeria monocytogenes* environmental sampling program in
738 ready-to-eat processing facilities: A practical approach. Compr Rev Food Sci Food Saf
739 **19:**2843-2861.

740 47. **Pightling AW, Pettengill JB, Luo Y, Baugher JD, Rand H, Strain E.** 2018. Interpreting
741 whole-genome sequence analyses of foodborne bacteria for regulatory applications and
742 outbreak investigations. Front Microbiol **9:**1482.

743 48. **Jagadeesan B, Gerner-Smidt P, Allard MW, Leuillet S, Winkler A, Xiao Y, Chaffron S, Van**
744 **Der Vossen J, Tang S, Katase M, McClure P, Kimura B, Ching Chai L, Chapman J, Grant K.**
745 2019. The use of next generation sequencing for improving food safety: Translation into
746 practice. Food Microbiol **79:**96-115.

747 49. **Zamudio R, Haigh RD, Ralph JD, De Ste Croix M, Tasara T, Zurfluh K, Kwun MJ, Millard**
748 **AD, Bentley SD, Croucher NJ, Stephan R, Oggioni MR.** 2020. Lineage-specific evolution

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

749 and gene flow in *Listeria monocytogenes* are independent of bacteriophages. Environ
750 Microbiol **22:**5058-5072.

751 50. **Ruppitsch W, Pietzka A, Prior K, Bletz S, Fernandez HL, Allerberger F, Harmsen D,**
752 **Mellmann A.** 2015. Defining and evaluating a core genome multilocus sequence typing
753 scheme for whole-genome sequence-based typing of *Listeria monocytogenes*. J Clin
754 Microbiol **53:**2869-2876.

755 51. **Wang Y, Pettengill JB, Pightling A, Timme R, Allard M, Strain E, Rand H.** 2018. Genetic
756 diversity of *Salmonella* and *Listeria* isolates from food facilities. J Food Prot **81:**2082-2089.

757 52. **Allard MW, Strain E, Rand H, Melka D, Correll WA, Hintz L, Stevens E, Timme R,**
758 **Lomonaco S, Chen Y, Musser SM, Brown EW.** 2019. Whole genome sequencing uses for
759 foodborne contamination and compliance: Discovery of an emerging contamination event
760 in an ice cream facility using whole genome sequencing. Infect Genet Evol **73:**214-220.

761 53. **Jagadeesan B, Baert L, Wiedmann M, Orsi RH.** 2019. Comparative analysis of tools and
762 approaches for source tracking *Listeria monocytogenes* in a food facility using whole-
763 genome sequence data. Front Microbiol **10:**947.

764 54. **Henri C, Leekitcharoenphon P, Carleton HA, Radomski N, Kaas RS, Mariet JF, Felten A,**
765 **Aarestrup FM, Gerner Smidt P, Roussel S, Guillier L, Mistou MY, Hendriksen RS.** 2017. An
766 assessment of different genomic approaches for inferring phylogeny of *Listeria*
767 *monocytogenes*. Front Microbiol **8:**2351.

768 55. **Chen Y, Gonzalez-Escalona N, Hammack TS, Allard MW, Strain EA, Brown EW.** 2016. Core
769 genome multilocus sequence typing for identification of globally distributed clonal groups

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

770 and differentiation of outbreak strains of *Listeria monocytogenes*. Appl Environ Microbiol
771 **82:**6258-6272.

772 56. **Van Walle I, Björkman JT, Cormican M, Dallman T, Mossong J, Moura A, Pietzka A,**
773 **Ruppitsch W, Takkinen J, European Listeria WGS Typing Group.** 2018. Retrospective
774 validation of whole genome sequencing-enhanced surveillance of listeriosis in Europe,
775 2010 to 2015. Euro Surveill **23:**1700798.

776 57. **Davis S, Pettengill JB, Luo Y, Payne J, Shpuntoff A, Rand H, Strain E.** 2015. CFSAN SNP
777 Pipeline: an automated method for constructing SNP matrices from next-generation
778 sequence data. PeerJ Comput Sci **1:**e20.

779 58. **Chmielowska C, Korsak D, Chapkauskaitse E, Decewicz P, Lasek R, Szuplewska M,**
780 **Bartosik D.** 2021. Plasmidome of *Listeria* spp. – The *repA*-family business. Int J Mol Sci
781 **22:**10320.

782 59. **Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM,**
783 **Larsen MV.** 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob
784 Chemother **67:**2640-2644.

785 60. **Fillgrove KL, Pakhomova S, Newcomer ME, Armstrong RN.** 2003. Mechanistic diversity of
786 fosfomycin resistance in pathogenic microorganisms. J Am Chem Soc **125:**15730-15731.

787 61. **Parsons C, Lee S, Kathariou S.** 2020. Dissemination and conservation of cadmium and
788 arsenic resistance determinants in *Listeria* and other Gram-positive bacteria. Mol
789 Microbiol **113:**560-569.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

790 62. **Müller A, Rychli K, Muhterem-Uyar M, Zaiser A, Stessl B, Guinane CM, Cotter PD,**
791 **Wagner M, Schmitz-Esser S.** 2013. Tn*6188* - a novel transposon in *Listeria monocytogenes*
792 responsible for tolerance to benzalkonium chloride. PLoS One **8:**e76835.

793 63. **Schmitz-Esser S, Anast JM, Cortes BW.** 2021. A large-scale sequencing-based survey of
794 plasmids in *Listeria monocytogenes* reveals global dissemination of plasmids. Front
795 Microbiol **12:**653155.

796 64. **Harter E, Wagner EM, Zaiser A, Halecker S, Wagner M, Rychli K.** 2017. Stress Survival Islet
797 2, predominantly present in *Listeria monocytogenes* strains of sequence type 121, is
798 involved in the alkaline and oxidative stress responses. Appl Environ Microbiol **83:**e00827-
799 17.

800 65. **Ryan S, Begley M, Hill C, Gahan CG.** 2010. A five-gene stress survival islet (SSI-1) that
801 contributes to the growth of *Listeria monocytogenes* in suboptimal conditions. J Appl
802 Microbiol **109:**984-995.

803 66. **Jordan SJ, Perni S, Glenn S, Fernandes I, Barbosa M, Sol M, Tenreiro RP, Chambel L,**
804 **Barata B, Zilhao I, Aldsworth TG, Adriao A, Faleiro ML, Shama G, Andrew PW.** 2008.
805 *Listeria monocytogenes* biofilm-associated protein (BapL) may contribute to surface
806 attachment of *L. monocytogenes* but is absent from many field isolates. Appl Environ
807 Microbiol **74:**5451-5456.

808 67. **Popowska M, Krawczyk-Balska A, Ostrowski R, Desvaux M.** 2017. InlL from *Listeria*
809 *monocytogenes* is involved in biofilm formation and adhesion to mucin. Front Microbiol
810 **8:**660.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

811 68. **Kovacevic J, Arguedas-Villa C, Wozniak A, Tasara T, Allen KJ.** 2013. Examination of food
812 chain-derived *Listeria monocytogenes* strains of different serotypes reveals considerable
813 diversity in *inlA* genotypes, mutability, and adaptation to cold temperatures. Appl Environ
814 Microbiol **79:**1915-1922.

815 69. **Upham J, Chen S, Boutilier E, Hodges L, Eisebraun M, Croxen MA, Fortuna A, Mallo GV,**
816 **Garduño RA.** 2019. Potential ad hoc markers of persistence and virulence in Canadian
817 *Listeria monocytogenes* food and clinical isolates. J Food Prot **82:**1909-1921.

818 70. **Heir E, Lindstedt BA, Rotterud OJ, Vardund T, Kapperud G, Nesbakken T.** 2004. Molecular
819 epidemiology and disinfectant susceptibility of *Listeria monocytogenes* from meat
820 processing plants and human infections. Int J Food Microbiol **96:**85-96.

821 71. **Baines SL, da Silva AG, Carter GP, Jennison A, Rathnayake I, Graham RM, Sintchenko V,**
822 **Wang Q, Rockett RJ, Timms VJ, Martinez E, Ballard S, Tomita T, Isles N, Horan KA,**
823 **Pitchers W, Stinear TP, Williamson DA, Howden BP, Seemann T, Communicable Diseases**
824 **Genomics Network (CDGN).** 2020. Complete microbial genomes for public health in
825 Australia and the Southwest Pacific. Microb Genom **6:**471.

826 72. **Kvistholm Jensen A, Nielsen EM, Björkman JT, Jensen T, Müller L, Persson S, Bjerager G,**
827 **Perge A, Krause TG, Kiil K, Sørensen G, Andersen JK, Mølbak K, Ethelberg S.** 2016. Whole-
828 genome sequencing used to investigate a nationwide outbreak of listeriosis caused by
829 ready-to-eat delicatessen meat, Denmark, 2014. Clin Infect Dis **63:**64-70.

830 73. **Halbedel S, Wilking H, Holzer A, Kleta S, Fischer MA, Lüth S, Pietzka A, Huhulescu S,**
831 **Lachmann R, Krings A, Ruppitsch W, Leclercq A, Kamphausen R, Meincke M, Wagner-**

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

832 **Wiening C, Contzen M, Kraemer IB, Al Dahouk S, Allerberger F, Stark K, Flieger A.** 2020.
833 Large nationwide outbreak of invasive listeriosis associated with blood sausage, Germany,
834 2018-2019. Emerg Infect Dis **26:**1456-1464.

835 74. **Nüesch-Inderbinen M, Bloemberg GV, Müller A, Stevens MJA, Cernela N, Kollöffel B,**
836 **Stephan R.** 2021. Listeriosis caused by persistence of *Listeria monocytogenes* serotype 4b
837 sequence type 6 in cheese production environment. Emerg Infect Dis **27:**284-288.

838 75. **Lachmann R, Halbedel S, Lüth S, Holzer A, Adler M, Pietzka A, Dahouk SA, Stark K, Flieger**
839 **A, Kleta S, Wilking H.** 2022. Invasive listeriosis outbreaks and salmon products: a genomic,
840 epidemiological study. Emerg Microbes Infect **11:**1308-1315.

841 76. **Findlay M, Grix J.** 2003. Challenging forensic evidence? Observations on the use of DNA in
842 certain criminal trials. Curr Issues Crim Justice **14:**269-282.

843 77. **Brown TM, Geddes L, Gill P, Jesper-Mir E, Kayser M, Phillips C, Schneider PM,**
844 **Syndercombe-Court D, Thomas J, Wienroth M, Williams RTW.** 2017. Making sense of
845 forensic genetics. What can DNA tell you about a crime. Report published by: Sense about
846 Science, https://senseaboutscience.org/wp-content/uploads/2017/01/making-sense-of-
847 forensic-genetics.pdf. Accessed 13 May 2022.

848 78. **Rogers K.** 2021**.** 8 Jun. 2021. German *E. coli* outbreak of 2011. *In* Encyclopedia Britannica.
849 https://www.britannica.com/event/German-E-coli-outbreak-of-2011. Accessed 13 May
850 2022.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

851 79. **Köckerling E, Karrasch L, Schweitzer A, Razum O, Krause G.** 2017. Public health research
852 resulting from one of the world's largest outbreaks caused by entero-hemorrhagic
853 *Escherichia coli* in Germany 2011: A Review. Front Public Health **5:**332.

854 80. **Schimmer B, Nygard K, Eriksen HM, Lassen J, Lindstedt BA, Brandal LT, Kapperud G,**
855 **Aavitsland P.** 2008. Outbreak of haemolytic uraemic syndrome in Norway caused by *stx2*-
856 positive *Escherichia coli* O103:H25 traced to cured mutton sausages. BMC Infect Dis **8:**41.

857 81. **Manuel CS, Van Stelten A, Wiedmann M, Nightingale KK, Orsi RH.** 2015. Prevalence and
858 distribution of *Listeria monocytogenes inlA* alleles prone to phase variation and *inlA* alleles
859 with premature stop codon mutations among human, food, animal, and environmental
860 isolates. Appl Environ Microbiol **81:**8339-8345.

861 82. **Van Stelten A, Simpson JM, Ward TJ, Nightingale KK.** 2010. Revelation by single-
862 nucleotide polymorphism genotyping that mutations leading to a premature stop codon in
863 *inlA* are common among *Listeria monocytogenes* isolates from ready-to-eat foods but not
864 human listeriosis cases. Appl Environ Microbiol **76:**2783-2790.

865 83. **Gorski L, Cooley MB, Oryang D, Carychao D, Nguyen K, Luo Y, Weinstein L, Brown E,**
866 **Allard M, Mandrell RE, Chen Y.** 2022. Prevalence and clonal diversity of over 1,200 *Listeria*
867 *monocytogenes* isolates collected from public access waters near produce production
868 areas on the central California coast during 2011 to 2016. Appl Environ Microbiol
869 **88:**e0035722.

870 84. **Yan S, Li M, Luque-Sastre L, Wang W, Hu Y, Peng Z, Dong Y, Gan X, Nguyen S, Anes J, Bai**
871 **Y, Xu J, Fanning S, Li F.** 2019. Susceptibility (re)-testing of a large collection of Listeria

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

872 monocytogenes from foods in China from 2012 to 2015 and WGS characterization of
873 resistant isolates. J Antimicrob Chemother **74:**1786-1794.

874 85. **Zhang X, Liu Y, Zhang P, Niu Y, Chen Q, Ma X.** 2021. Genomic characterization of clinical
875 *Listeria monocytogenes* isolates in Beijing, China. Front Microbiol **12:**751003.

876 86. **Wu L, Bao H, Yang Z, He T, Tian Y, Zhou Y, Pang M, Wang R, Zhang H.** 2021. Antimicrobial
877 susceptibility, multilocus sequence typing, and virulence of listeria isolated from a
878 slaughterhouse in Jiangsu, China. BMC Microbiol **21:**327.

879 87. **Hanes RM, Huang Z.** 2022. Investigation of antimicrobial resistance genes in *Listeria*
880 *monocytogenes* from 2010 through to 2021. Int J Environ Res Public Health **19:**5506.

881 88. **Gómez D, Azón E, Marco N, Carramiñana JJ, Rota C, Ariño A, Yangüela J.** 2014.
882 Antimicrobial resistance of *Listeria monocytogenes* and *Listeria innocua* from meat
883 products and meat-processing environment. Food Microbiol **42:**61-65.

884 89. **Rostamian M, Kooti S, Mohammadi B, Salimi Y, Akya A.** 2022. A systematic review and
885 meta-analysis of *Listeria monocytogenes* isolated from human and non-human sources:
886 the antibiotic susceptibility aspect. Diagn Microbiol Infect Dis **102:**115634.

887 90. **Orsi RH, den Bakker HC, Wiedmann M.** 2011. *Listeria monocytogenes* lineages: Genomics,
888 evolution, ecology, and phenotypic characteristics. Int J Med Microbiol **301:**79-96.

889 91. **Maury MM, Tsai YH, Charlier C, Touchon M, Chenal-Francisque V, Leclercq A, Criscuolo**
890 **A, Gaultier C, Roussel S, Brisabois A, Disson O, Rocha EPC, Brisse S, Lecuit M.** 2016.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

891 Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. Nat
892 Genet **48:**308-313.

893 92. **Bechtel TD, Gibbons JG.** 2021. Population genomic analysis of *Listeria monocytogenes*
894 from food reveals substrate-specific genome variation. Front Microbiol **12:**620033.

895 93. **Lee S, Chen Y, Gorski L, Ward TJ, Osborne J, Kathariou S.** 2018. *Listeria monocytogenes*
896 source distribution analysis indicates regional heterogeneity and ecological niche
897 preference among serotype 4b clones. mBio **9:**e00396-18.

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

901 95. **Palaiodimou L, Fanning S, Fox EM.** 2021. Genomic insights into persistence of *Listeria*
902 species in the food processing environment. J Appl Microbiol **131:**2082-2094.

903 96. **Melero B, Manso B, Stessl B, Hernández M, Wagner M, Rovira J, Rodríguez-Lázaro D.**
904 2019. Distribution and persistence of *Listeria monocytogenes* in a heavily contaminated
905 poultry processing facility. J Food Prot **82:**1524-1531.

906 97. **Demaître N, Rasschaert G, De Zutter L, Geeraerd A, De Reu K.** 2021. Genetic *Listeria*
907 *monocytogenes* types in the pork processing plant environment: from occasional
908 introduction to plausible persistence in harborage sites. Pathogens **10:**717.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

909 98. **Rückerl I, Muhterem-Uyar M, Muri-Klinger S, Wagner KH, Wagner M, Stessl B.** 2014. *L.*
910 *monocytogenes* in a cheese processing facility: Learning from contamination scenarios
911 over three years of sampling. Int J Food Microbiol **189:**98-105.

912 99. **Harrand AS, Jagadeesan B, Baert L, Wiedmann M, Orsi RH.** 2020. Evolution of *Listeria*
913 *monocytogenes* in a food processing plant involves limited single-nucleotide substitutions
914 but considerable diversification by gain and loss of prophages. Appl Environ Microbiol
915 **86:**e02493-19.

916 100. **Muhterem-Uyar M, Ciolacu L, Wagner KH, Wagner M, Schmitz-Esser S, Stessl B.** 2018.
917 New aspects on *Listeria monocytogenes* ST5-ECVI Predominance in a heavily
918 contaminated cheese processing environment. Front Microbiol **9:**64.

919 101. **Taylor AJ, Stasiewicz MJ.** 2019. Persistent and sporadic *Listeria monocytogenes* strains do
920 not differ when growing at 37 °C, in planktonic state, under different food associated
921 stresses or energy sources. BMC Microbiol **19:**257.

922 102. **Assisi C, Forauer E, Oliver HF, Etter AJ.** 2021. Genomic and transcriptomic analysis of
923 biofilm formation in persistent and transient *Listeria monocytogenes* isolates from the
924 retail deli environment does not yield insight into persistence mechanisms. Foodborne
925 Pathog Dis **18:**179-188.

926 103. **Gelbíčová T, Florianová M, Tomáštíková Z, Pospíšilová L, Koláčková I, Karpíšková R.**
927 2019. Prediction of persistence of *Listeria monocytogenes* ST451 in a rabbit meat
928 processing plant in the Czech Republic. J Food Prot **82:**1350-1356.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

929 104. **Koreňová J, Oravcová K, Véghová A, Karpíšková R, Kuchta T.** 2016. Biofilm formation in
930 various conditions is not a key factor of persistence potential of *Listeria monocytogenes* in
931 food-processing environment. J Food Nutr Res **55:**189-193.

932 105. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM,**
933 **Nikolenko SI, Son P, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev**
934 **MA, Pevzner PA.** 2012. SPAdes: A new genome assembly algorithm and its applications to
935 single-cell sequencing. J Comput Biol **19:**455-477.

936 106. **Bushnell B.** 2014. BBMap: A Fast, Accurate, Splice-Aware Aligner, 9th Annual Genomics of
937 Energy & Environment Meeting, Walnut Creek, CA, USA
938 https://www.osti.gov/biblio/1241166-bbmap-fast-accurate-splice-aware-aligner.

939 107. **Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A.** 2018. Versatile genome
940 assembly evaluation with QUAST-LG. Bioinformatics **34:**i142-i150.

941 108. **Kaufman L, Rousseeuw PJ.** 1990. Introduction, p 1-67, Finding groups in data: An
942 introduction to cluster analysis, Ch. 1. doi:10.1002/9780470316801.ch1. John Wiley &
943 Sons, Inc.: Hoboken, NJ, USA.

944 109. **Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K.** 2021. cluster: Cluster analysis
945 basics and extensions. R package version 2.1.1, https://CRAN.R-
946 project.org/package=cluster.

947 110. **Gascuel O.** 1997. BIONJ: an improved version of the NJ algorithm based on a simple model
948 of sequence data. Mol Biol Evol **14:**685-695.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

949 111. **Paradis E, Schliep K.** 2019. ape 5.0: an environment for modern phylogenetics and
950 evolutionary analyses in R. Bioinformatics **35:**526-528.

951 112. **R_Core_Team.** 2021. R: A language and environment for statistical computing, R
952 Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

953 113. **Letunic I, Bork P.** 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic
954 tree display and annotation. Nucleic Acids Res **49:**W293-W296.

955 114. **Gu Z, Gu L, Eils R, Schlesner M, Brors B.** 2014. *circlize* implements and enhances circular
956 visualization in R. Bioinformatics **30:**2811-2812.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

957 # Tables

958 **Table 1: Comparison of number of differences detected in each CC using SNP and wgMLST**
959 **analysis[a]**

	CC group	Number of genomes analysed	Number of detected SNPs (after filtering)	Number of differing wgMLST loci	Pairwise distance (wgMLST)		
					Maximum	Minimum	Median
lineage I	CC1	34	222	230	170	0	5
	CC2	7	6	11	7	2	4
	CC3	24	79	84	74	0	2.5
	CC5	13	79	88	45	0	26
	CC6	3	31	36	35	4	34
	CC88	2	9	17	17	17	17
	CC220	2	105	105	105	105	105
	CC315	10	4	11	5	0	2
lineage II	CC7	68	633	694	198	0	79
	CC8	19	203	288	181	1	37
	CC9	290	966	896	179	0	94
	CC11	9	138	215	83	3	49
	CC14	20	176	221	103	0	20
	CC18	5	106	140	88	2	63.5
	CC19	55	209	199	105	0	77
	CC20	4	57	99	77	3	56
	CC21	2	126	118	118	118	118
	CC31	3	193	198	132	112	131
	CC37	14	81	78	50	0	5
	CC91	14	468	628	512	0	102
	CC121	86	573	608	238	0	48
	CC177	22	222	223	160	0	2
	CC199	3	4	6	6	0	6
	CC403	27	19	27	9	0	2
	CC415	30	203	231	118	0	81
Median		14	126	140	103	0	37

960 [a] CC4, CC101, and CC200 are not shown as they comprised only one isolate (from M4, M4, and M3,
961 respectively).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

962 **Table 2: Novel or rare genetic elements**

Isolate	CC	Source and year of isolation	Genetic determinant	Further description
Food industry isolates				
MF6196	CC7	Factory M2, 2013	*repA* G12	Novel *repA* variant found on a 58 Kb contig, locus tag: JKS07_13825.
MF4562	CC9	Factory M1, 2012	*cadA1C1*	The only food isolate with Tn4422/*cadA1C1* located on the chromosome
MF1548	CC21	Factory M3, 1998	*cadA5C5*	The only genome identified with this locus. The genome also contains *cadA1C1* and both arsenic resistance operons (on LGI2 and Tn*554*-like transposon)
Clinical isolates				
ERR2522309	CC1	Clinical, 2013	*repA* G4	The only identified genome with this *repA* variant. Located on a 65 kb contig.
			qacH-like gene	A plasmid-encoded variant of QacH, 90% identical to QacH encoded on Tn*6188*.
ERR2522330	CC7	Clinical, 2014	plasmid pAUSMDU00000235	This genome contained a contig that could be circularized and was 100% identical to the 2776 bp pAUSMDU00000235 plasmid found in a clinical strain in Australia in 2009 (71).
ERR2522310	CC7 (ST691)	Clinical, 2013	plasmid pLMST6 (pLmN12-0935)	This genome contained a contig that aligned with 99.98% identity over 99.86% of the plasmid/contig lengths with the 4392 bp long plasmid pLMST6 (pLmN12-0935) from a clinical strain belonging to ST403/CC403 isolated in Switzerland in 2012.
			emrC	The only genome identified with this gene is present on the plasmid.
ERR2522285	CC59	Clinical, 2012	*tet*(M) allele 7	The only antibiotic resistance gene identified in this study, found by search against the ResFinder database (59).

963

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

964 **Table 3. *L. monocytogenes* clones with repeated isolation of same strain in same factory**[a]

CC	ST	Processing plant										Niches with repeated isolation of same strain				
		M1	M2	M4	M6	M8	S1	S2	S3	S5	S6	Floors[b]	Drains	Conveyors	Gutting	Other
CC1	ST1							X					S2	S2	S2	Grader, containers
CC5	ST5				X							M6				
CC7	ST7		X		X		X	X				M2, M6, S1	M2, S1	S1		M2: pallet; M6: products; S1: bone napping machine
	ST732							X				S2	S2	S2	S2	Inflow tube
CC8	ST8	X			X		X				X	M1, S1	M1, S1	S1	S6	S6: salmon; M1: all samples from raw meat department
CC9	ST9[c]	X		X		X						M1, M4	M1, M4 M8			M1: equipment framework (near floor); M4: 3 month period 55 isolates from slicer, products
CC14	ST14							X	X			S3	S2	S2, S3	S2	
CC19	ST1416	X		X								M4	M4			M1: Specific room, raw side
	ST19										X	S6			S6	Head cutter, packaging machine
CC37	ST37										X				S6	Helix, Grader
CC91	ST91	X										M1	M1			M1: All from raw meat departments
CC121	ST121			X		X					X	M4, S6	M4	S6	S6	S6: Head cutter, scaleoff
CC177	ST177									X				S5		Grader
CC199	ST199					X										M8: Three isolates; drain, gate, product
CC315	ST249										X					Filleting machine, scaleoff
CC403	ST403										X			S6	S6	Freezing tray
CC415	ST394	X		X								M1, M4	M4			M1: Raw meat department

965 [a] Included isolates comprise clones with ≤20 wgMLST allelic differences collected in at least two calendar
966 years.
967 [b] Floors includes floor-related sites such as wheels, floor mats, shoes, or floor weights.
968 [c] Includes one ST2344 isolate.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure legends

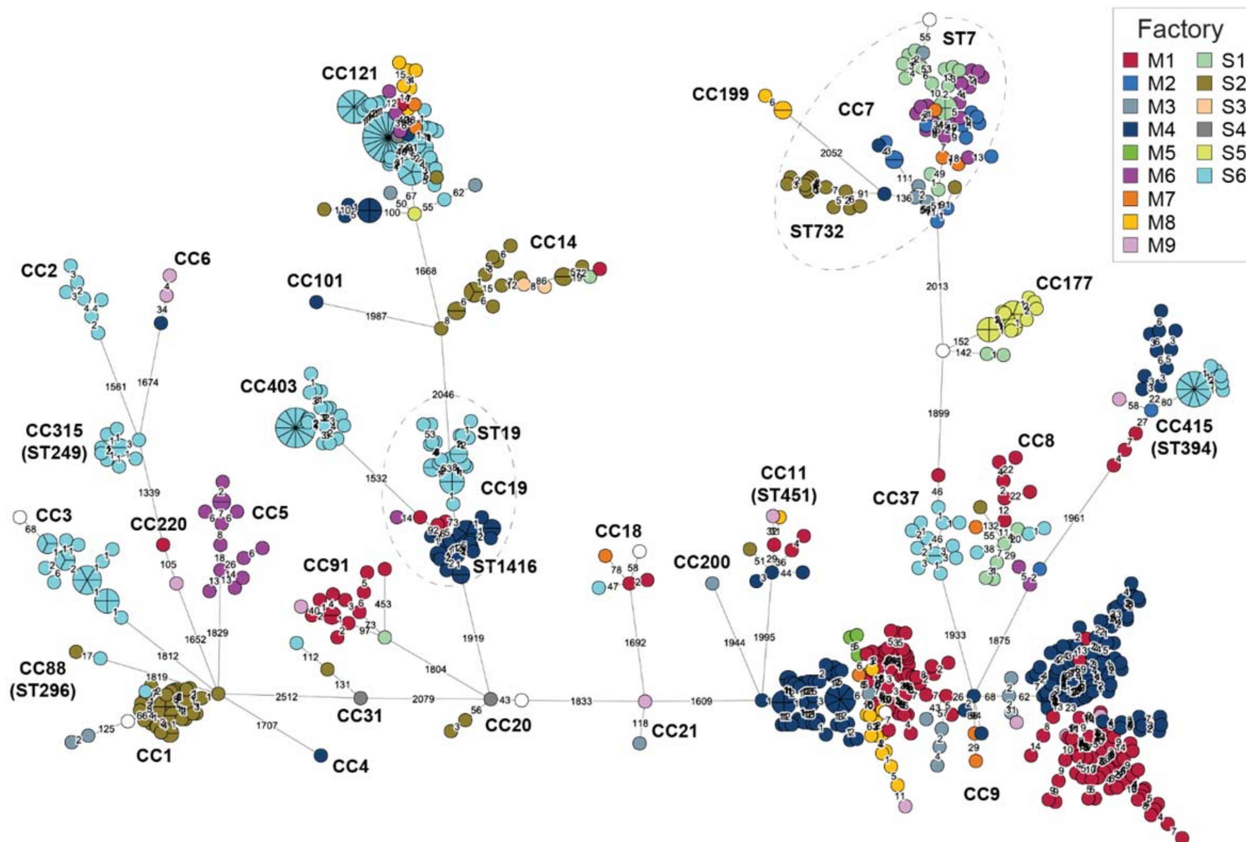
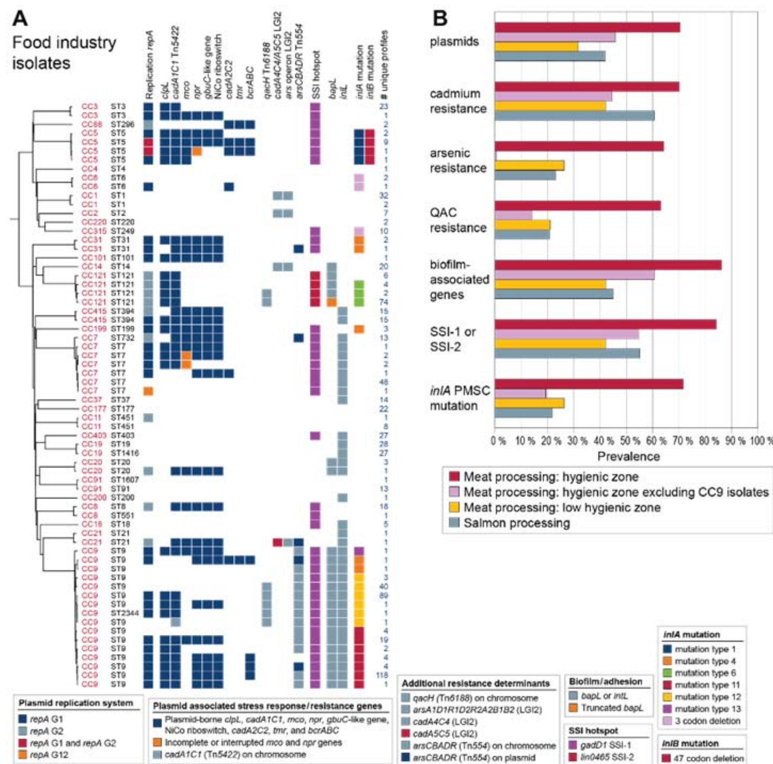


Figure 1: Minimum spanning tree based on wgMLST analysis for the 769 *L. monocytogenes* from food processing environments. The area of each circle is proportional to the number of isolates represented, and the number of allelic differences between isolates is indicated on the edges connecting two nodes. The nodes are colored by factory of origin (meat production plants M1-M9; salmon processing plants S1-S6), and CCs and STs are indicated (the ST number is the same as the CC unless specified).

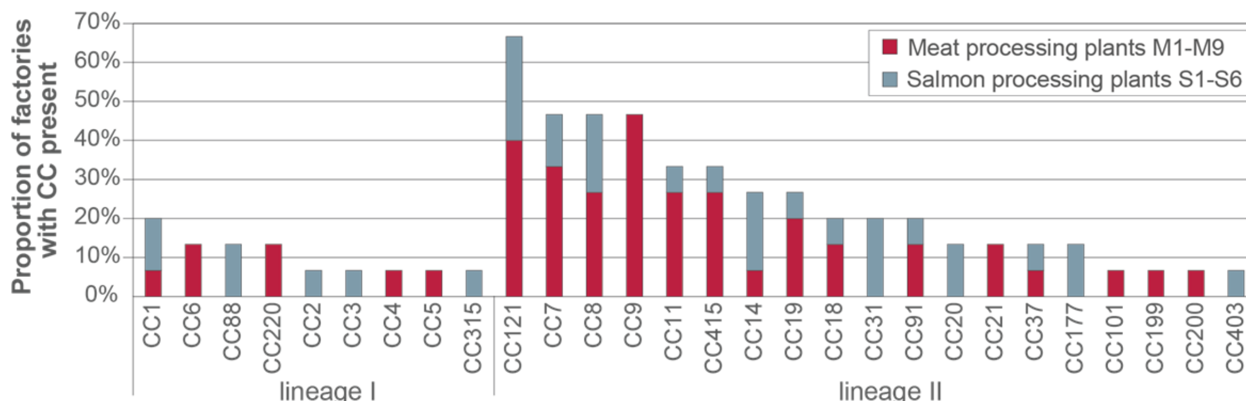
bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



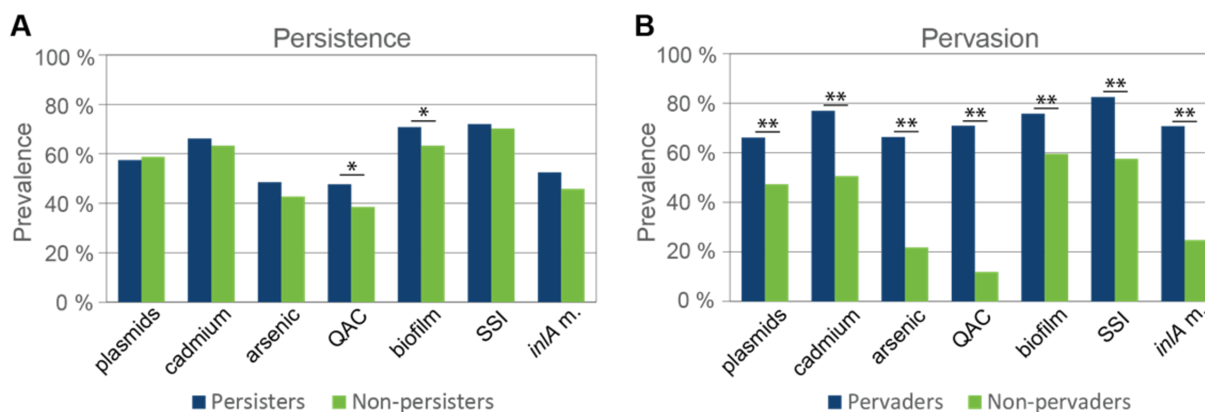
977

978 **Figure 2**: Presence of accessory genetic determinants associated with stress survival, resistance,
979 or persistence in *L. monocytogenes* from food processing environments. **A)** Unique
980 combinations of ST and the variable stress response loci in isolates from food processing
981 environments. All 769 isolates are shown individually in Fig. S1. The phylogeny is a midpoint-
982 rooted Neighbor-Joining tree based on wgMLST analysis and shows one arbitrarily selected
983 genome from each of the groups of genomes containing the same unique gene combination.
984 The number of genomes harboring the same unique combination is indicated in the right
985 column. **B)** Prevalence of genetic determinants in isolates from different sources. Cadmium
986 resistance; *cadA1C1*, *cadA2C2*, *cadA4A4* or *cadA5C5*, arsenic resistance;
987 *arsA1D1R1D2R2A2B1B2* on LGI2 or *arsCBADR* on Tn554-like transposon, QAC resistance; *bcrABC*
988 or *qacH*, biofilm-associated genes; *bapL* and/or *inlL*. Statistical analysis for differences in
989 categories using Pearson's chi-squared test is presented in Table S5.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

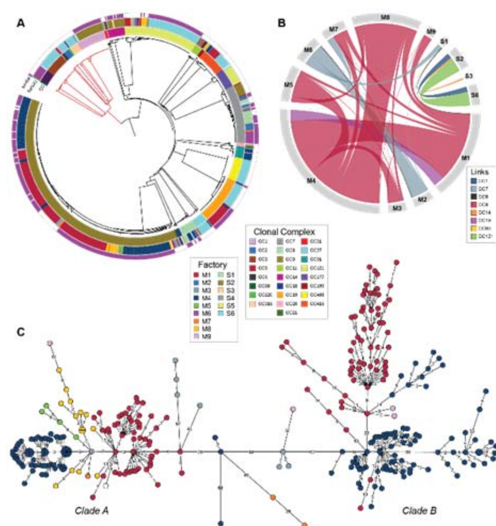


990
991 **Figure 3:** Different CCs detected in varying proportions of factories. The data is presented as a
992 stacked bar plot showing the percentage of examined meat and salmon processing plants in
993 which each CC was detected. A total of 15 factories were included.



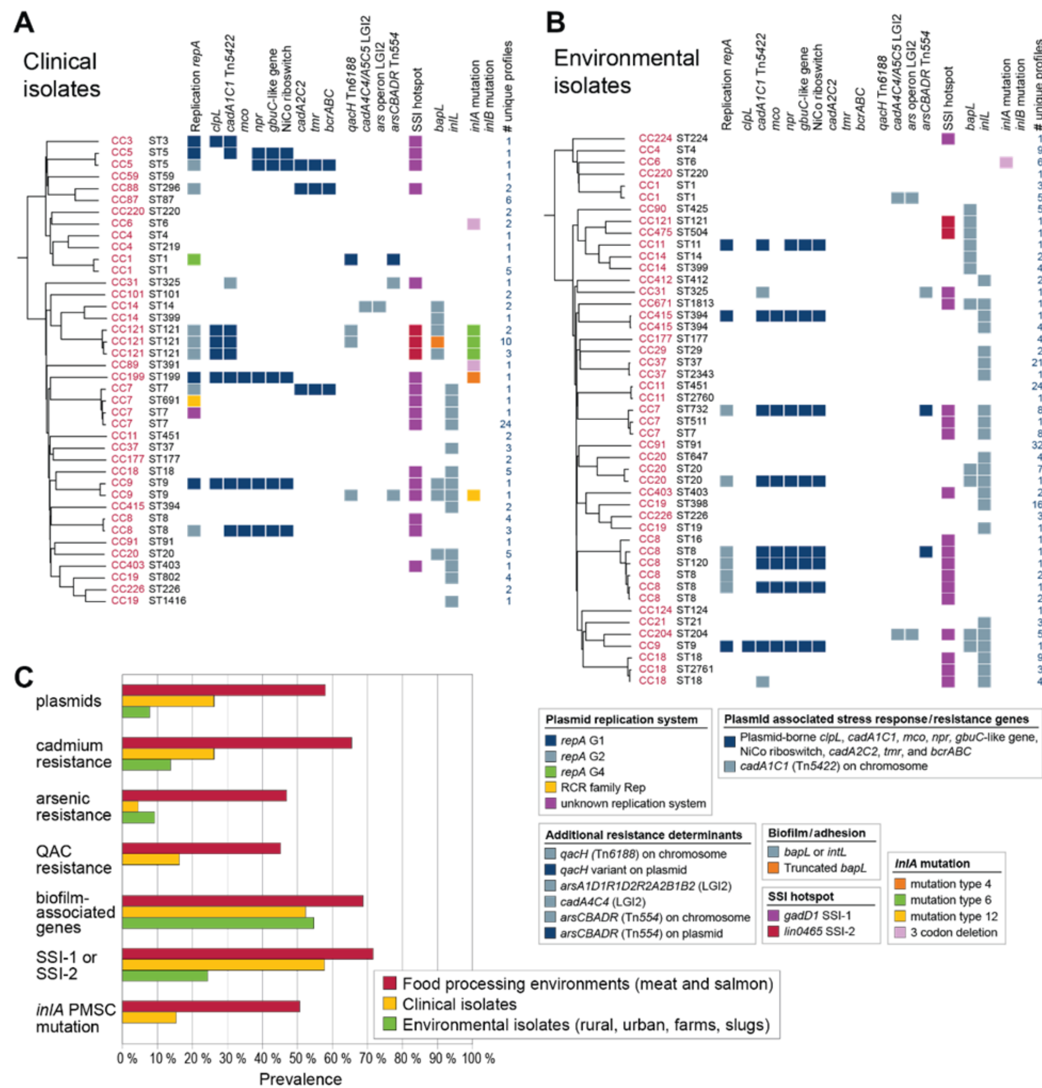
994
995 **Figure 4**: Prevalence of subgroups of accessory genetic determinants associated with stress
996 survival, resistance, and persistence in *L. monocytogenes* classified as **A)** persistent and/or **B)**
997 pervasive. Cadmium; cadmium resistance loci *cadA1C1*, *cadA2C2*, *cadA4C4*, or *cadA5C5*, arsenic;
998 arsenic resistance operons *arsA1D1R1D2R2A2B1B2* on LGI2 or *arsCBADR* on the Tn554-like
999 transposon, QAC; QAC resistance loci *qacH* or *bcrABC*, biofilm; biofilm-associated genes *bapL*
1000 and/or *inlL*, SSI; stress survival islets SSI-1 or SSI-2, and *inlA* m.; PMSC mutation in the *inlA* gene.
1001 Asterisks represent significant differences (Pearson's chi-squared test. *, $p < 0.05$; **, $p < 0.001$,
1002 see also Table S6).

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1003 

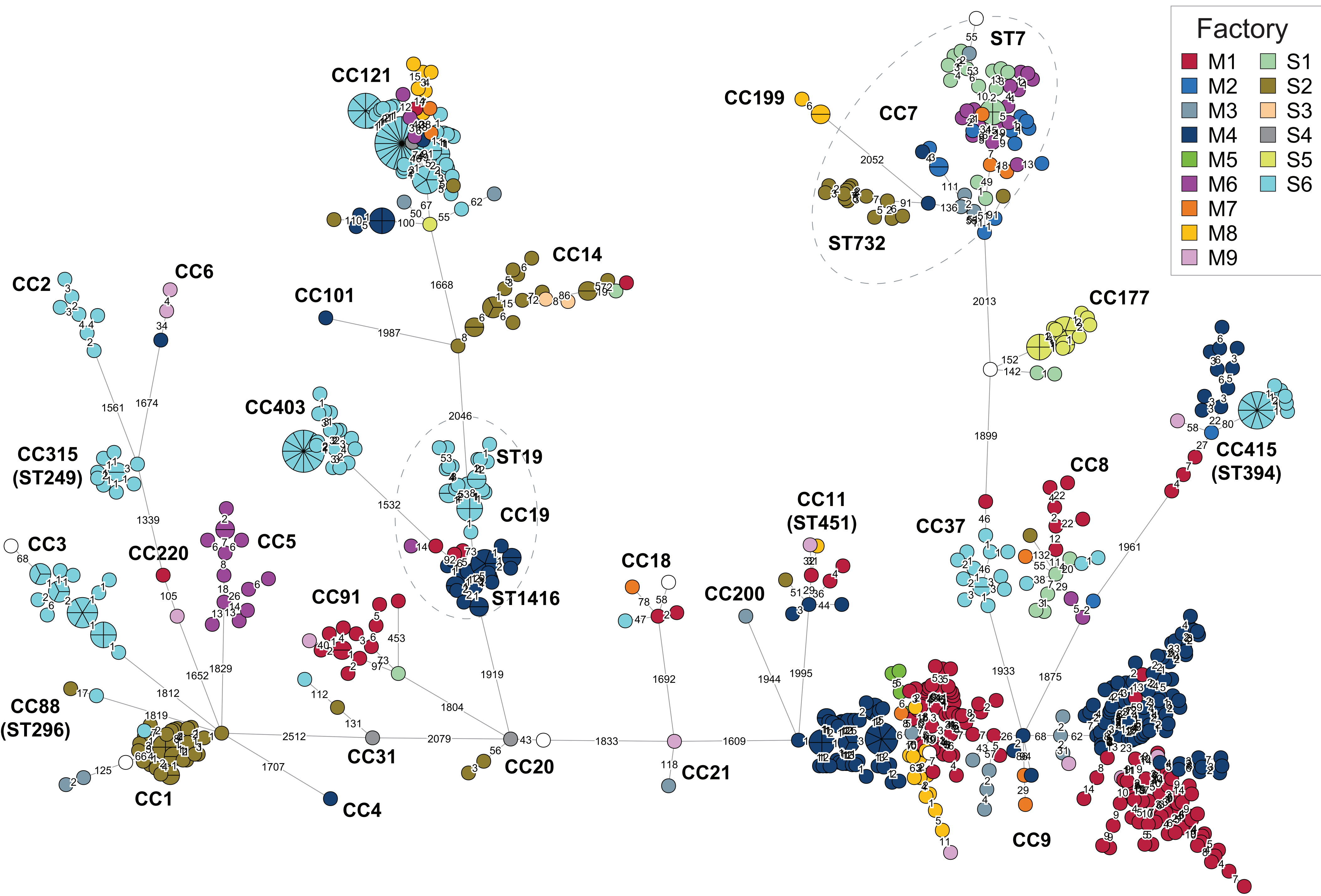
1004 **Figure 5:** Pervasive strains, present in more than one processing plant. **A)** Neighbor-Joining
1005 phylogenetic tree based on wgMLST analysis, showing CC (inner ring), factory of origin (middle
1006 ring), and isolates for which genetically similar isolates (≤20 wgMLST allelic differences) were
1007 found in at least one other factory (outer ring). Red branches are lineage I, black branches
1008 lineage II. **B)** The genetic associations between isolates from different factories illustrated by a
1009 chord diagram. The outer sectors represent the factories for which genetically similar isolates
1010 (≤20 wgMLST allelic differences) were found in at least one other factory, and the links between
1011 the factories represent the pairs of genetically similar isolates, colored by CC group. The
1012 thickness of each arc represents the sum of the similarity scores for each pair of isolates found
1013 in the two factories, weighted by their similarity. **C)** Minimum spanning tree showing the
1014 relationship between the 290 CC9 isolates. Nodes are colored by factory of origin (same colors
1015 as in **A**; the white-colored isolate is from a domestic kitchen). The area of each node is
1016 proportional to the number of isolates represented, and the number of allelic differences
1017 between isolates is indicated on the branches connecting two nodes. Branch lengths are square
1018 root scaled.

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

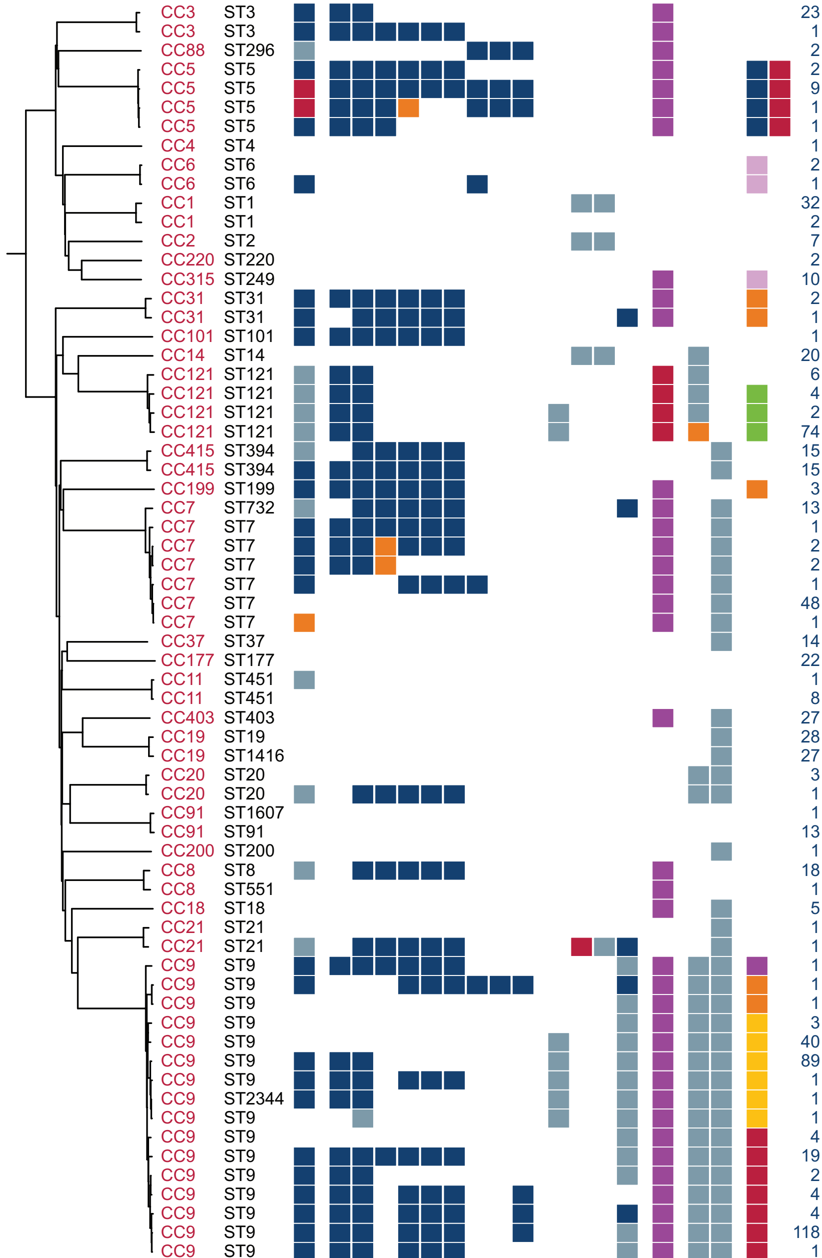


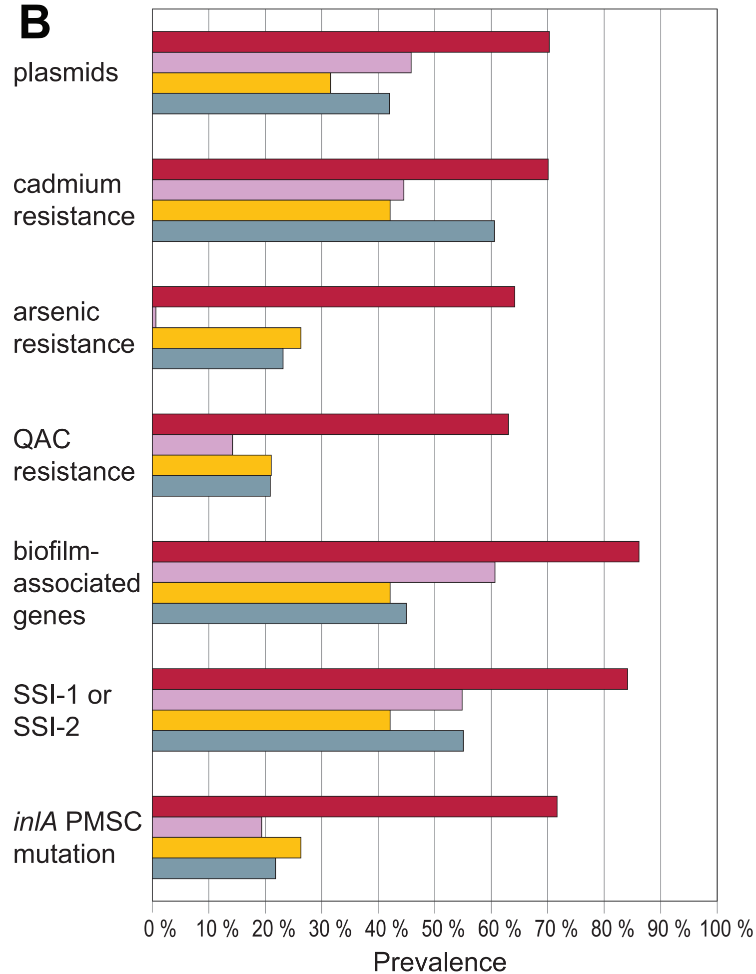
1019

1020 **Figure 6**: Presence of accessory genetic determinants associated with stress survival, resistance,
1021 or persistence in *L. monocytogenes* from different sources. **A)** Norwegian clinical isolates 2010-
1022 2015 (56), and **B)** isolates from natural (rural/urban/farm/slug) environments in Norway (2). For
1023 each ST, one arbitrarily selected genome from each of the groups of genomes containing the
1024 same unique combination of stress response loci is shown. **C)** Prevalence of genetic
1025 determinants in isolates from different sources. See legend to Fig. 2 for details on categories.
1026 Statistical analysis for differences in categories using Pearson's chi-squared test is presented in
1027 Table S7.



Factory
M1
M2
M3
M4
M5
M6
M7
M8
M9
S1
S2
S3
S4
S5
S6
ST7
CC7
CC199
ST732
CC121
CC14
CC101
CC2
CC6
CC177
CC403
ST19
CC19
CC315 (ST249)
CC415 (ST394)
CC8
CC3
CC220
CC5
CC11 (ST451)
CC37
CC18
CC200
ST1416
CC91
CC88 (ST296)
CC1
CC4
CC31
CC20
CC21
CC9





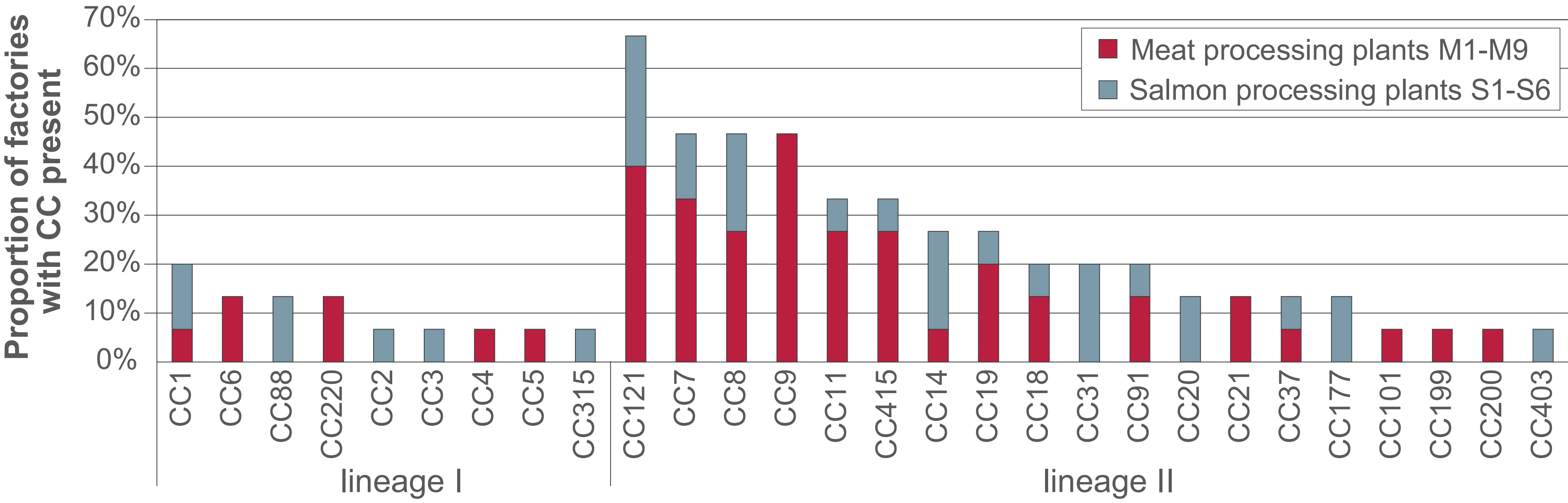
Additional resistance determinants
- qacH (Tn6188) on chromosome
- arsA1D1R1D2R2A2B1B2 (LGI2)
- cadA4C4 (LGI2)
- cadA5C5 (LGI2)
- arsCBADR (Tn554) on chromosome
- arsCBADR (Tn554) on plasmid

Biofilm/adhesion
- bapL or intL
- Truncated bapL

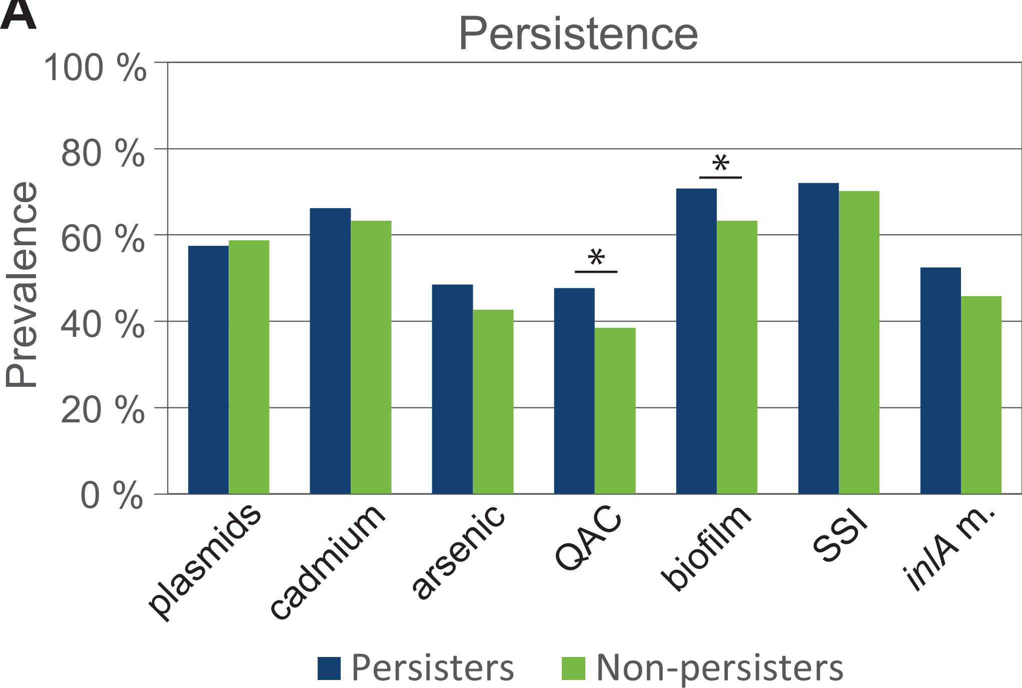
SSI hotspot
- gadD1 SSI-1
- lin0465 SSI-2

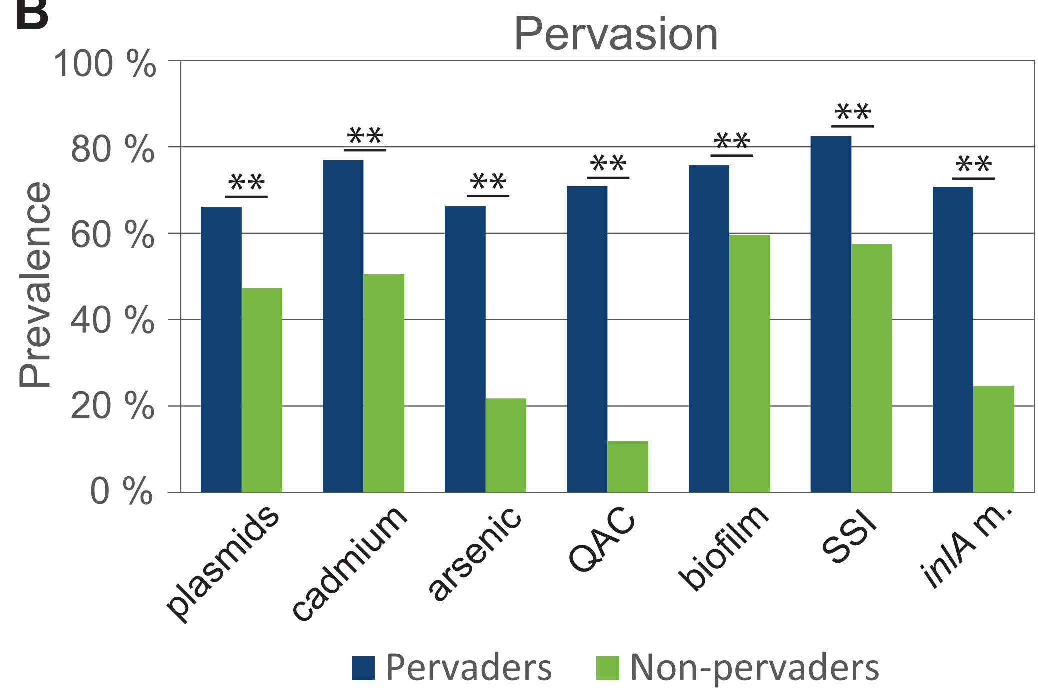
inlA mutation
- mutation type 1
- mutation type 4
- mutation type 6
- mutation type 11
- mutation type 12
- mutation type 13
- 3 codon deletion

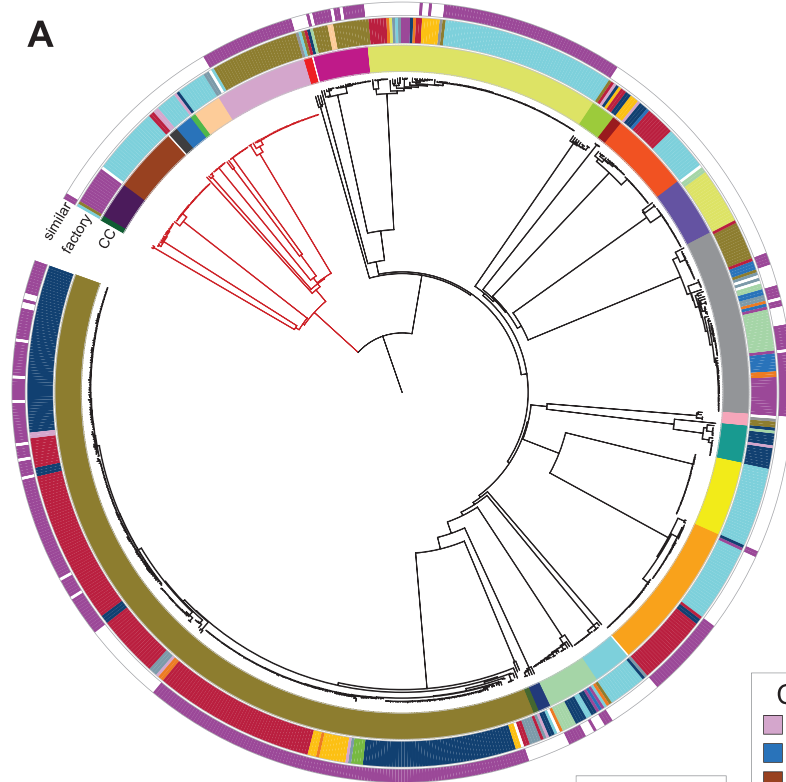
inlB mutation
- 47 codon deletion

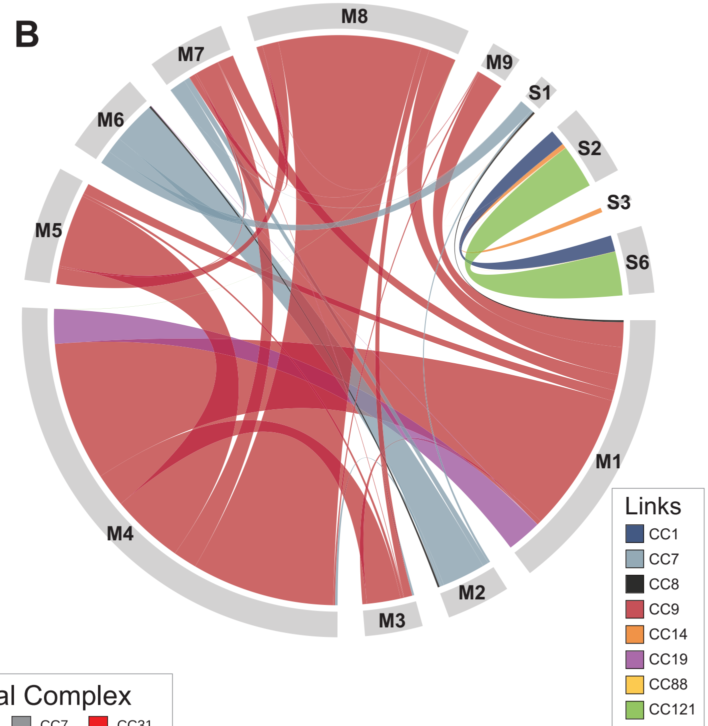


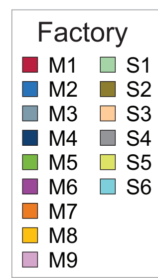
Proportion of factories with CC present
70%
60%
50%
40%
30%
20%
10%
0%
Meat processing plants M1-M9
Salmon processing plants S1-S6
CC1
CC6
CC88
CC220
CC2
CC3
CC4
CC5
CC315
CC121
CC7
CC8
CC9
CC11
CC415
CC14
CC19
CC18
CC31
CC91
CC20
CC21
CC37
CC177
CC101
CC199
CC200
CC403
lineage I
lineage II

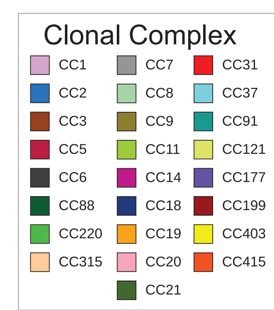

A
Persistence
100 %
80 %
60 %
40 %
20 %
0 %
Prevalence
*
*
plasmids
cadmium
arsenic
QAC
biofilm
SSI
inlA m.
Persisters
Non-persisters


B
Pervasion
100 %
80 %
60 %
40 %
20 %
0 %
Prevalence
**
**
**
**
**
**
**
plasmids
cadmium
arsenic
QAC
biofilm
SSI
inlA m.
Pervaders
Non-pervaders









C

Clade A

Clade B

bioRxiv preprint doi: https://doi.org/10.1101/2022.05.25.493524; this version posted May 27, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.