Pervasive *Listeria monocytogenes* are common in the Norwegian food system and associated with increased prevalence of stress survival and resistance determinants

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

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

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

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

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

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

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

Abstract

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

38 # Importance

39 *Listeria monocytogenes* can persist in food processing environments for months to decades and
40 spread through the food system 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 system. The current study
43 presents a comprehensive overview of the prevalence of persistent clones and of the diversity of
44 *L. monocytogenes* in Norwegian food processing facilities. The results demonstrate extensive
45 spread of highly similar strains throughout the Norwegian food system, 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.

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 disinfection
56 agents, high and low pH, and low temperatures (3, 4), it is very difficult to eliminate *L.*
57 *monocytogenes* from food processing environments. Clonal populations of *L. monocytogenes* that
58 survive in the processing environment over an extended time-period (months or years) are
59 referred to as persistent *L. monocytogenes* (5, 6). In contrast, transient or sporadic *L.*
60 *monocytogenes* are those that enter the processing environment without establishing a
61 permanent presence there, but which instead are eliminated through cleaning and disinfection
62 (6, 7). Some authors also define a category of "persistent transient" *L. monocytogenes*
63 contamination which is a consequence of continual introduction of one or more subtypes into the
64 processing environment from outside reservoirs combined with a failure to apply sufficient
65 *Listeria* control measures (7). The concept of pervasive bacterial strains is sometimes used to
66 describe subpopulations of bacteria with enhanced ability to spread or migrate to new
67 geographical locations or ecological habitats (8, 9). This term has not previously been used to
68 describe subpopulations of *L. monocytogenes*, although the dissemination of persistent strains to
69 more than one food processing facility is a well-documented phenomenon (10-18).

70 In a phylogenetic context, *L. monocytogenes* comprises four separate deep-branching lineages,
71 which are further subdivided into sequence types (STs) and clonal complexes (CCs or clones) by

72 multilocus sequence typing (MLST) (19). Certain clones such as CC1 and CC4 belonging to lineage
73 I are commonly associated with clinical disease while others, often belonging to lineage II (e.g.,
74 CC9, CC121), are frequently found in food processing environments and food, but rarely among
75 clinical cases (19-21). The underlying causes behind these differences are not fully understood,
76 but thought to be linked to differences in virulence potential and the ability to survive and
77 multiply in food processing environments (11, 22-29). An increased capacity for biofilm formation
78 can contribute to the survival and persistence of *L. monocytogenes* in both natural and food
79 processing environments (3, 30-32). Resistance to stressors encountered in food processing
80 environments, e.g., biocides and alkaline pH, may also contribute to survival. Associated stress
81 resistance determinants can be spread through mobile genetic elements such as plasmids,
82 prophages and transposons (21, 33-37). For example, it has been shown that the presence of
83 *bcrABC* or *qacH* (located on plasmids and transposon Tn*6188*, respectively) results in tolerance to
84 low concentrations of quaternary ammonium compounds (QAC), biocides commonly used in the
85 food industry (26, 36, 38-40). Another genetic determinant associated with CCs commonly found
86 in food processing plants is premature stop codon (PMSC) truncation mutations in *inlA* encoding
87 the virulence factor internalin A (41, 42). Although the ecological significance of *inlA* PMSC
88 mutations is not fully understood, some studies indicate that they mediate increased adhesion
89 and biofilm formation (43-45) and increased tolerance to 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

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

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

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

133 Results

134 Diversity of isolates from the Norwegian food system

135 The basis of the current study was a collection of 769 *L. monocytogenes* isolates from the
136 Norwegian food industry from 1990 to 2020 (Table 1 andTable S1 in the supplemental material).

137 Samples were mainly from the processing environment (floors, drains, and or food processing
138 equipment) and to a minor degree from raw materials and products. In total, 460 isolates were
139 from meat processing industry, in which 413 were from environmental samples from hygienic
140 zones in meat processing factories (either raw or cooked meat departments), 21 were from meat
141 product samples, seven were from raw materials or meat sampled during processing, and 19 were
142 from zones with lower hygienic conditions associated with meat processing (i.e., animal transport
143 vehicles, animal holding pens, and slaughter departments). From salmon processing factories,
144 306 isolates were included, of which 232 isolates were environmental samples from the
145 processing environment, 27 were from product samples (ranging from gutted salmon to packaged
146 filets), and 48 were samples collected from raw materials entering the processing factories. All
147 were collected from nine different meat production plants and six different salmon processing
148 plants (Table S2), except for five isolates from other meat processing environments and three
149 isolates from other food associated sources (salmon and cheese product, domestic kitchen). A
150 subset of the isolates belonging to *L. monocytogenes* ST8 and ST9 (5 and 252 isolates,
151 respectively) were previously subjected to WGS analysis as part of earlier studies (10, 11). The
152 additional 512 isolates were subjected to WGS, *in silico* MLST, and wgMLST profiling. MLST
153 showed that 13% of the 769 studied isolates (n=97) belonged to lineage I and the remaining
154 isolates to lineage II (n=672), while lineage III or IV isolates were not detected. The isolates were
155 assigned to 33 different STs and 28 different CCs (Fig. 1).

156 Genetic distances obtained from wgMLST analysis were compared with results obtained from a
157 SNP analysis performed separately for each CC using the CFSAN SNP pipeline (57), with reference
158 genomes selected from each CC. The average number of SNPs and wgMLST loci detected within

159 each CC was 126 and 140, respectively. The results show that with default filtering settings,
160 wgMLST analysis was somewhat more sensitive than SNP analysis (Table 2).

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

162 A BLAST analysis was carried out to detect plasmids and genetic determinants associated with
163 stress survival and antimicrobial resistance. Overall, plasmids were identified in 58% of the
164 *L. monocytogenes* isolates. All were *repA*-family theta-replicating plasmids belonging to either
165 group 1 (G1) or group 2 (G2) (58), with one exception. Isolate MF6196 belonging to CC7 harbored
166 a novel RepA group protein, hereby named RepA group 12 (G12), which was 76% identical to
167 RepA G2 (Table 3). In total, 39% of the isolates harbored *repA* G1 plasmids while 18% harbored
168 *repA* G2 plasmids (Table S3). In addition, both G1 and G2 *repA* genes were identified in ten CC5
169 isolates collected in factory M6 between 2016 and 2019, indicating that they harbored two
170 plasmids. These isolates were related (13-45 wgMLST differences) to three CC5 isolates containing
171 *repA* G1 but not *repA* G2, collected in the same factory during 2010 and 2012, suggesting that this
172 strain had acquired a second plasmid harboring a *repA* G2 gene during the intervening years.

173 The search for genetic determinants associated with stress survival, resistance, and persistence
174 identified 76 core genes that were present in all genomes, and 20 accessory genes, gene loci or
175 gene variants that were present in a subset of genomes (Fig. 2A and Table S4). The only identified
176 antibiotic resistance gene, detected by searching the ResFinder database (59), was the fosfomycin
177 resistance gene (*lmo1702*, *fosX_2*) (60), present in all genomes. The accessory genetic
178 determinants comprised cadmium resistance (*cadA1C1*, *cadA2C2*, *cadA4C4*, *cadA5C5*) and
179 arsenic resistance operons (*arsA1D1R1D2R2A2B1B2* on *Listeria* Genomic Island 2 (LGI2) and
180 *arsCBADR* on a Tn*554*-like transposon) (61), QAC resistance loci (chromosomally encoded qacH

181 and plasmid encoded bcrABC (40, 62)), various additional plasmid-associated stress response
182 genes (*clpL*, *mco*, *npr*, a *gbuC*-like gene, a NiCo riboswitch, and *tmr* (63)), genes located on the
183 stress survival islets (SSIs) SSI-1 and SSI-2 (64, 65), biofilm associated genes (*bapL* and *inlL* (66,
184 67)), and PMSC and internal deletion mutations in *inlA* and *inlB* (45). Further details regarding the
185 identified accessory loci found in food environmental isolates are presented in Text S1 in the
186 supplemental material and Table 3. The unique combinations of accessory genes are presented
187 in Fig. 2A and the complete phylogeny is shown in Fig. S1.

188 For statistical analyses, the accessory stress survival genes were grouped into categories of
189 cadmium resistance, arsenic resistance, QAC resistance, stress survival islets, biofilm-associated
190 genes, and *inlA* PMSC mutations (the identified *inlA* 3-codon deletion (3CD) *inlA* mutation
191 associated with an increased Caco-2 cell invasion phenotype (68, 69) was not included in this
192 category) (Fig. 2B). The prevalence of plasmids and all tested categories of stress survival genes
193 was significantly higher in isolates from hygienic zones in meat processing factories than in salmon
194 processing environments and in low-hygienic zones associated with meat production ($p \leq 0.01$;
195 Table S5). However, it should be emphasized that the high prevalence among isolates from meat
196 processing was associated with their high prevalence within CC9 isolates, which constituted the
197 majority of isolates in this category (n=286; 65%). When CC9 isolates were excluded from the
198 analysis, the occurrence of cadmium and arsenic resistance loci was significantly higher in isolates
199 from salmon processing environments than from hygienic zones in meat processing factories
200 ($p \leq 0.001$). In contrast, the occurrence of biofilm-associated genes was significantly higher also
201 when CC9 isolates were excluded in isolates from hygienic zones in meat processing factories than

202 from salmon processing environments (p=0.001) and low-hygienic zones associated with meat
203 production (p=0.02) (Fig. 2B and Table S5).

204 Prevalence of each CC in food processing plants

205 The number of sequenced isolates from each plant varied widely, ranging from 4 to 192 for the
206 meat processing plants and from 2 to 188 for the salmon processing plants (Table S2). To assess
207 the prevalence of each CC in food processing plants, we therefore did not count the raw number
208 of collected isolates belonging to each CC, as this number would be heavily biased towards CCs
209 present in the plants where the greatest number of samples were obtained. Instead, the
210 prevalence of different CCs was assessed by counting the number of processing plants harboring
211 each CC (Fig. 3). The CC detected in the greatest number of processing plants overall was CC121
212 (found in 6/9 meat factories and 4/6 salmon factories), followed by CC7, CC8, and CC9. CC9 was
213 detected in the largest number of meat processing plants (7/9).

214 Closely related isolates were present over time within individual factories

215 We next investigated whether persistence of specific strains of *L. monocytogenes* occurred in
216 processing plants, and whether this was associated with certain CCs or the presence of plasmids
217 or stress survival genes. In total, 551 isolates (72%), belonging to 15 different CCs, were linked to
218 persistence. The proportion of persistent isolates was significantly higher in lineage II (74%) than
219 in lineage I (53%) ($p < 0.001$). A persistent isolate was here defined as an isolate that showed 20
220 or fewer wgMLST allelic differences towards an isolate collected from the same factory in a
221 different calendar year. A persistent strain was defined as a clonal population showing this level
222 of similarity towards at least one other isolate in a cluster found across more than one calendar

223 year in the same factory. The definition of persistence is irrespective of whether the isolates
224 originated from e.g., an established house strain, reintroduction from raw materials or external
225 environment, or from house strains present at a supplier's factory. Clonal clusters of isolates
226 collected within the same calendar year and thus not designated as persistent included 23 isolates
227 belonging to CC3 (0-8 wgMLST differences) collected from factory S6 between January and April
228 of 2020, and 55 CC9 isolates collected during an eight-week period in 2014 at factory M4 in
229 connection with the previously described event related to installation of a contaminated second-
230 hand slicer line (10).

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

237 Factories and typical niches where persistent *L. monocytogenes* strains were found are
238 summarized in Table 4. The most common sites found to be contaminated with persistent *Listeria*
239 were floors and drains in both meat and salmon processing plants as well as conveyor belts and
240 gutting machines in salmon processing plants. The results show that many different CCs can be
241 associated with persistence. However, some appear to have a greater tendency than others for
242 becoming persistent, e.g., CC7 and CC8, each identified as persistent in four factories, and CC9
243 and CC121, each in three factories.

244 From the salmon slaughterhouses (S1-S6), persistent strains were identified in ten CCs (11 STs;
245 Table 4). Six of these CCs were present among isolates repeatedly found in S6, where the greatest
246 number of isolates of the same persistent CC121 strain (n=59) was obtained through sampling in
247 the period from February 2019 to April 2020. The emergence of CC121 was followed by extensive
248 sampling in S6 in this period and the CC121 strain was detected in the slaughterhouse processing
249 equipment, machines and environment. Interestingly, the CC121 strain was also found onboard
250 a salmon slaughter ship supplying S6 with fresh salmon for further processing and on 26 samples
251 from fresh salmon sampled upon arrival in the factory. This indicated that fresh slaughtered
252 salmon contaminated with *L. monocytogenes* was a likely source for the introduction and
253 subsequent persistence of this particular CC121 strain in the plant. Salmon slaughtered in other
254 slaughterhouses and subsequently further processed in factory S2 was likely the source of a
255 persistent CC1 strain repeatedly and extensively isolated in samples from production
256 environment, equipment and product in S2 over a two-year period: In a cluster of 30 isolates
257 differing by 0-16 wgMLST alleles, two isolates were obtained from samples of supplied
258 slaughtered salmon. For salmon processing plant S1, we previously reported isolation of the same
259 CC8 strain ten years apart (11). This factory also harbored a persistent strain of CC7 that was
260 repeatedly isolated from the processing environment and equipment throughout a three-year
261 sampling period. A similar situation was observed in factory S5, with repeated isolation of a CC177
262 strain over a two-year sampling period. In factory S2, clonal ST732 isolates (CC7) sampled eight
263 years apart (2011-2019) were isolated from various surfaces of equipment and processing
264 environments. These observations indicate that *L. monocytogenes* strains had persisted in the

265 respective slaughterhouses or were repeatedly reintroduced between sampling events during the
266 study period.

267 For the meat processing plants, repeated isolation of the same strain over at least two different
268 years was observed for nine CCs (Table 4). The dominance and persistence of CC9/ST9 over
269 several years in meat processing plants M1 and M4 was previously described (10). In the present
270 study, persistence of a CC9 strain over a two-year period was also confirmed in factory M8. For
271 factory M1, the CC9 strains were repeatedly isolated from the department producing heat-
272 treated products (10), while in raw meat departments, persistent strains were identified for CC8,
273 CC19, CC91 and CC415. Closely related isolates of CC19 and CC415 were also repeatedly isolated
274 in factory M4, but in contrast to M1, only in the department producing heat-treated products.
275 Persistent CC7 strains were isolated from floors in poultry processing plants M2 and M6. CC7 was
276 the dominant clonal group in M2, and the only CC from which the same strain was repeatedly
277 isolated in this factory. In factory M6, a persistent strain of CC5 was repeatedly isolated in 2012
278 and during 2016-2019 from floors in a room used for processing of heat-treated products. Closely
279 related CC199 was isolated three times over an eight-month period from factory M8. Two
280 different clusters of CC121 isolates were identified in drains in factory M4, one consisting of two
281 isolates collected three years apart and differing by six wgMLST alleles, and another comprising
282 six isolates differing by 0-6 wgMLST alleles, collected over a period of four years. A cluster of
283 CC121 was also detected in different sites in factory M6, comprising six isolates differing by 3-23
284 wgMLST alleles over a two-year period. Thus, for several CCs, including CC7, CC9, CC19 and CC121,
285 persistent strains were repeatedly isolated from more than one meat processing factory, while
286 some CCs (e.g., CC5, CC199) were repeatedly isolated only at single plants.

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

288 Phylogenetic analysis revealed that many isolates were closely related despite being isolated from
289 different processing plants (Fig. 1). We hereby designate the observation of clonal populations of
290 *L. monocytogenes* found in more than one factory as pervasion. The definition does not
291 differentiate between the mode of dissemination between factories and includes isolates with a
292 common source and ancestor. In total, 433 of the isolates (56%) were pervasive, showing 20 or
293 fewer wgMLST allelic differences towards an isolate found in a different factory. A pervasive strain
294 was defined as the isolates belonging to such a cluster. The proportion of pervasive isolates
295 differed between food source, with 70% of isolates from meat processing environments
296 designated as pervasive, in contrast to only 40 % of isolates from salmon processing industry. The
297 proportion of pervasive isolates was significantly higher in lineage II (59%) than in lineage I (34%)
298 ($p<0.001$).

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

306 Factories S2 and S6 were the two most heavily sampled salmon slaughterhouses. The two
307 factories are located in different geographic regions in Norway and belong to different
308 companies. WGS was performed for 68 isolates collected during 2011-2012 and 2018-2019 from

309 S2, and for 188 isolates collected between June 2017 and April 2020 from S6. We identified three
310 different CCs (CC1, CC121, and CC88) in which closely related isolates (≤17 wgMLST allelic
311 differences) were found in both factories. In the case of CC1, one isolate from a product sample
312 collected in October 2019 from factory S6 (MF7739) differed by 7-15 wgMLST alleles from the
313 previously mentioned persistent strain comprising 30 environmental isolates from factory S2, for
314 which the likely source was slaughtered salmon. These were obtained between July 2018 and
315 May 2019. Also, one CC121 isolate collected from a drain at factory S2 in 2012 (MF4804) showed
316 5-11 wgMLST allelic differences towards the previously mentioned cluster of 59 persistent CC121
317 isolates collected in factory S6 during 2019-2020, which included one isolate collected onboard
318 the slaughter ship. However, the slaughter ship had not been in operation in 2012 when the S2
319 isolate was detected. The third case linking these two factories were two isolates separated by
320 17 wgMLST alleles belonging to CC88, obtained from the boots of a factory worker at S2 in
321 November 2018 and from salmon from an external supplier at factory S6 one year later, in
322 November 2019. Genetically similar isolates were also found at factories S2 and S3, where two
323 CC14 isolates collected in S2 in 2011-2012 showed 12-28 wgMLST allelic differences towards a
324 cluster of 13 isolates from S3 collected during 2018-2019.

325 In the meat industry, the majority of isolates from pervasive strains belonged to CC9, for which
326 we previously described a close genetic relationship between *L. monocytogenes* from four meat
327 processing plants (M1, M4, M5 and M7) collected during 2009-2017 (10). In the current study, 38
328 additional CC9 isolates, including representatives from three additional processing plants (M3,
329 M8, and M9) were sequenced (Fig. 5C). Of particular interest was a group of CC9 isolates collected
330 in 1998 and 2001 from the raw side of a meat factory (M3) that is no longer in operation (70).

331 One of the isolates from 1998 differed by only 5-8 wgMLST allelic differences to isolates from five
332 other processing plants (M1, M4, M5, M7, and M8), suggesting that this CC9 strain has circulated
333 in Norwegian meat chains for at least two decades.

334 Pervasion was significantly associated with presence of stress survival genes

335 In contrast to persistent strains (Fig. 4A), isolates identified as pervaders showed significantly
336 increased prevalence of plasmids and all examined categories of stress and persistence associated
337 genes (p<0.001) (Fig. 4B and Table S6). The prevalence of individual gene variants (*clpL*, *cadA1C1*,
338 *cadA2C2*, *bcrABC*, *qacH,* the *ars* operon on LGI2, *arsCBADR* on Tn*554*, SSI-1, SSI-2, *bapL*, and *inlL*)
339 was also significantly higher in isolates classified as pervaders compared to non-pervaders
340 (p≤0.004) (Table S6).

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

348 Stress survival determinants in clinical and environmental isolates

349 To compare the prevalence of plasmids and stress survival genes in isolates from food processing
350 environments to human clinical isolates and isolates from natural environments in Norway, the
351 BLAST analysis was also performed for the genomes of 111 Norwegian clinical isolates and 218

352 isolates from rural, urban, and farm environments in Norway, which were examined in a recent
353 study (2). The prevalence of plasmids was significantly different between isolates from the three
354 sources (p<0.001; Table S7), with 26% and 8% of clinical and environmental isolates harbouring
355 plasmids, respectively, compared to 41% for the isolates from food processing environments
356 (Table S3). The lowest prevalence of plasmids was found in the 87 isolates from dairy farms (5%).
357 In addition to *repA* G1 and *repA* G2 plasmids, one *repA* group 4 (G4) plasmid and two small non-
358 *repA* plasmids were identified among the clinical isolates (Table 3 and Text S1).

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

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

377 Discussion

378 To gain a better understanding of the population structure and genomic diversity of *L.*
379 *monocytogenes* in the Norwegian food system, genome sequences from 769 *L. monocytogenes*
380 isolates from the Norwegian food industry collected over three decades from 15 food processing
381 factories were characterized using WGS-based comparative genomic analyses. The study showed
382 that 56% of isolates were closely related (2-20 wgMLST allelic differences) to an isolate collected
383 from a different factory. These isolates were designated as *pervasive*, a term which has previously
384 been used to describe subpopulations of bacteria with enhanced ability to spread or migrate to
385 new geographical locations (8, 9). Several other studies have similarly reported that *L.*
386 *monocytogenes* isolates from geographically and temporally unrelated sources were separated
387 by equally short genetic distances (2, 12-18). WGS analyses are increasingly used in epidemiology
388 and have in recent years been essential for solving several European foodborne listeriosis
389 outbreaks (72-75). However, as we know from the use of DNA as forensic evidence in criminal
390 trials, the apparent certainty of DNA evidence can be deceptive, and there is a danger that the
391 statistical significance of a DNA match can be overstated (76, 77). This is particularly likely in the
392 case of bacteria which reproduce using binary fission and especially for *L. monocytogenes* which
393 has an extremely low evolutionary rate (16, 19). Indeed, in one case described by Lüth *et al*. (18),

394 the same *L. monocytogenes* strain identified in two different processing plants matched the same
395 CC5 outbreak cluster. There is a risk that authorities could mistakenly consider a WGS match
396 between two bacterial isolates as proof of identification of a contamination source, also in cases
397 lacking other epidemiological evidence. It is therefore of crucial importance to consider the
398 possibility of highly similar isolates being found in multiple factories when WGS analyses of *L.*
399 *monocytogenes* are performed, both for outbreak investigations and for food safety risk-based
400 decisions and risk assessment in food industry. The high prevalence of *L. monocytogenes* isolates
401 identified in more than one processing plant in the current study suggests that the occurrence of
402 pervasive strains in the Norwegian food industry – particularly in the meat distribution chain
403 where 70% of isolates were pervasive – may be significantly higher than in other countries. This
404 possibly reflects a particularly complex and interconnected Norwegian meat supply chain.
405 Regardless, mistaken identification of an outbreak source can have enormous economic impacts
406 and cause significant food waste due to unnecessary recalls. Examples include the German 2011
407 outbreak of enterohaemorrhagic *Escherichia coli* (EHEC) where initially cucumbers imported from
408 Spain were erroneously implicated as the source of the infections (78, 79), and the Norwegian
409 2006 EHEC outbreak caused by contaminated traditional cured sausage ("morrpølse"), in which
410 minced meat was initially indicated as the source (80).

411 The prevalence of genetic determinants associated with stress survival, metal and biocide
412 resistance and biofilm formation, as well as PMSC mutations in *inlA*, was higher in isolates
413 collected from food processing environments than among clinical isolates and isolates from
414 natural environments, thus supporting previous studies suggesting that these factors are involved
415 in the survival and growth in food processing environments (21, 35, 39, 64, 69, 81, 82). Notably,

416 none of the 218 isolates from natural environments contained biocide resistance genes *qacH*,
417 *bcrABC*, or *inlA* PMSC mutations. This concurs with a recent study of *L. monocytogenes* collected
418 from surface waters in California (83), in which *qacH* and *bcrABC* were detected in zero and 18
419 isolates respectively, while *inlA* PMSC mutations were present in only four of 1248 examined
420 isolates. In contrast, we identified *qacH* or *bcrABC* genes in 45% and *inlA* PMSC mutations in 51%
421 of isolates from food processing environments. These values coincide with data from numerous
422 previous studies (26, 28, 68, 81, 82). A practical implication of this finding would be to change to
423 a disinfectant with an other mechanisms of action when facing challenges with *L. monocytogenes*
424 surviving in the production environment, e.g., change to an oxidative disinfectant for factories
425 using a QAC-based disinfectant. In drains, which may be difficult to reach through regular
426 disinfection, citric acid could be used to reduce the problem (84). Antibiotic resistance
427 determinants present on mobile genetic elements appear to have low prevalence in *L.*
428 *monocytogenes* isolated in Norway, as only one such antibiotic resistance gene was detected
429 among the 1098 examined genomes (a *tet*(M) tetracycline resistance gene in a clinical isolate).
430 This is lower than that recently reported among isolates from other countries (85-90).

431 Typical niches where persistent *L. monocytogenes* strains were found in the current study were
432 largely consistent with previous studies showing that the most common sites contaminated with
433 persistent *Listeria* were floors, drains, conveyor belts, slicers, and tables (7). However, isolation
434 of persistent strains on conveyors was only observed in salmon processing plants, and in addition
435 salmon gutting machines were identified as common sites for isolation.

436 It is acknowledged that the distribution and abundance of *L. monocytogenes* CCs varies between
437 different environments, such as humans, animals, soil, water, plants, and various types of food,

438 and that this is likely driven by selective adaptation (2, 22, 24, 29, 91-94). However, the same is
439 not always true when comparing persistent and transient isolates from food processing
440 environments, although it is highly likely that inheritable genetic traits are responsible also for
441 the ability to survive long-term in food processing facilities. It is frequently reported that certain
442 CCs have been identified as persistent in food processing environments, notably CC9 and CC121,
443 however, persistent strains have been detected in many CCs including CC5, CC7, CC8, CC31,
444 CC155, and CC321 (14, 16, 18, 21, 26, 95-100). This largely corresponds with observations from
445 the current study, in which CC7, CC8, CC9 and CC121 were identified as persistent in the greatest
446 number of processing plants. Nevertheless, persistent strains were identified within 15 of the 25
447 CCs represented by more than one isolate (60%), including in three lineage I CCs and 12 lineage II
448 CCs.

449 It is likely that the genetic basis behind increased prevalence of certain *L. monocytogenes* in food
450 processing environments aligns with the genetic basis behind persistence in the ecological sense
451 of the word, i.e., an increased ability to survive long-term in food processing environments (5, 6).
452 There is some evidence of association of persistence with the presence of the *bcrABC* cassette
453 conferring resistance to QAC disinfection agents (26, 39, 101) and perhaps also biofilm formation
454 capacity (30, 100). However, several studies have failed to identify any phenotypic or genotypic
455 differences between persistent and transient *L. monocytogenes* (6, 96, 102, 103) or failed to
456 associate persistent strains with differences in stress response, sanitizer resistance, or adhesion
457 properties (16, 25, 32, 70, 103-105). In line with these reports, the current study showed a
458 relatively weak statistically significant association ($p \leq 0.04$) between persistence and QAC
459 resistance determinants and biofilm-associated genes, and none between persistence and

460 plasmids, heavy metal resistance genes, SSIs, or *inlA* PMSC mutations. In contrast, we found that
461 pervasive isolates, belonging to strains present in more than one factory, showed strong
462 statistically significant association ($p<0.001$) with all examined categories of stress survival and
463 persistence genes, as well as towards plasmids and several individual genes or gene loci. Pervasive
464 strains were identified in eight CCs: one belonging to lineage I and seven belonging to lineage II.
465 Notably, only a small proportion of pervaders were classified as non-persistent (i.e., transient). It
466 thus appeared that strains that occurred at several factories and were repeatedly isolated over
467 time in one or more of these facilities were more likely to show problematic properties, i.e., carry
468 genetic determinants that enable them to establish as house strains and disperse to new
469 environments. Our results furthermore support the hypothesis that there is not one single genetic
470 determinant responsible for survival in food processing environments, but rather an
471 accumulation of stress resistance genes, biofilm associated genes, and *inlA* PMSC mutations.

472 The identity of the isolates classified as non-persistent pervaders in the current study exemplify
473 one of the difficulties in separating between "true" persistent and "true" transient isolates using
474 operational definitions. These included 82 isolates of which 55 were from the previously
475 described contaminated second-hand slicer line isolated from factory M4 (10). These were not
476 defined as persistent, as the contamination was only detected during an eight-week period in
477 2014 in which the slicer line was installed in the factory. However, from an ecological viewpoint,
478 they obviously belonged to a house strain established in a difficult to clean niche (5, 6). Another
479 challenge with operational definitions of persistence is that strains without any specific adaptive
480 features responsible for increased survival in processing environments may survive there under
481 permissive conditions, e.g., in a period with higher temperatures or inadequate cleaning and

482 sanitation, or alternatively, reoccurrence can be due to repeated introduction from an outside
483 reservoir (7, 46). Therefore, even when using a sampling method targeted towards house strains,
484 including sampling after cleaning and disinfection and detection of reoccurrence over a longer
485 time-period, the obtained isolates may not carry specific genetic determinants for survival in the
486 factory environment. The results from the current study suggest that the operational definition
487 of pervasion is superior to those used to define persistence in identifying strains that carry
488 adaptations responsible for increased ability to survive and multiply in food processing
489 environments. This approach may contribute to further unravelling of mechanisms responsible
490 for survival of *L. monocytogenes* in the food system, which in turn could guide improvements in
491 control strategies for this important pathogen.

492 Material and methods

493 Source of isolates

494 The isolates from food processing environment, raw materials and processed foods included in
495 the study are listed in Table S1 and were from the *L. monocytogenes* strain collection at Nofima,
496 Norway. A total of 305 isolates from both meat and salmon industry were collected during 2011-
497 2015 as part of a previous study (38). The majority of these were isolated after sanitation, and
498 before start of production. Most of the isolates obtained after 2016 were from the factories' own
499 sampling programs, and for these, sampling was mainly performed during production. A subset
500 of 252 isolates belonging to CC8 and CC9 has been described in previous studies (10, 11).

501 ## WGS and genome assembly

502 Bacteria were grown on BHI agar overnight at 37 °C before a loopful of cells was suspended in
503 500 µl 2x Tris-EDTA buffer with 1.2 % Triton X-100. Cells were lysed using lysing matrix B and a
504 FastPrep instrument (both MP Biomedicals), and genomic DNA was isolated using the DNeasy
505 blood and tissue kit (Qiagen). Libraries were prepared using the Nextera XT DNA sample
506 preparation kit or the Nextera DNA Flex Library prep kit (both Illumina) and sequenced on a MiSeq
507 platform with 300-bp paired-end reads. Raw reads were filtered on q15 and trimmed of adaptors
508 before *de novo* genome assembly was performed using SPAdes v3.10.0 or v.3.13.0 (106) with the
509 careful option and six k-mer sizes (21, 33, 55, 77, 99, and 127). Contigs with sizes of <500 bp and
510 *k*-mer coverage of <5 were removed from the assemblies (a coverage cutoff 15 was used for
511 MF7896). The average coverage for the genome assemblies was calculated using BBmap v36.92
512 (107). The quality of all assemblies was evaluated using QUAST v5.0.2 (108).

513 ## MLST analyses

514 Classical MLST analysis followed the MLST scheme described by Ragon *et al*. (42) and the database
515 maintained at the Institute Pasteur's *L. monocytogenes* online MLST repository
516 (http://bigsdb.web.pasteur.fr/listeria/listeria.html). CC14 and CC91 were defined as previously
517 described (2). The wgMLST analysis was performed using a whole-genome scheme containing
518 4797 coding loci from the *L. monocytogenes* pan-genome and the assembly-based BLAST
519 approach, implemented in BioNumerics 7.6 (http://www.applied-maths.com/news/listeria-
520 monocytogenes-whole-genome-sequence-typing).

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

526 Calculation of pairwise wgMLST distances for Neighbor-Joining (NJ) trees was performed using
527 the daisy function (109) from the cluster package v2.1.1 (110) in R and selection of the gower
528 metric. NJ trees were generated using an improved version of the NJ algorithm (BION) (111)
529 implemented in the ape package v5.4-1 (112) in R v4.0.4 (113) as function bionjs. Interactive Tree
530 Of Life (iTOL) v6.5.2 (114) was used for visualization.

531 SNP analysis

532 Read mapping based SNP analysis was performed separately for each CC. An internal reference
533 genome was selected from each CC (listed in Table S1) using the following criteria: Centrally
534 positioned in CC clusters by wgMLST analysis, from larger sub-clusters if more than one cluster
535 was present in a CC, older isolates preferred to more recent isolates, higher quality assemblies
536 (high coverage, few contigs) preferred. The reference based SNP analysis was performed using
537 the CFSAN SNP pipeline v2.1.1 (57) and default filtering settings, except that regions of high-
538 density SNPs were defined for each sample individually instead of filtering a dense region found
539 in any genome from all genomes. The default filtering removes SNPs located closer than 500 bp
540 to the end of a contig on the reference genome, SNPs where there are more than 3 SNPs in a
541 1000 base window, more than 2 SNPs in a 125 base window, and more than 1 SNP in a 15 base
542 window, reads with map quality below than 30, reads with a base quality of at least 15 at a given

543 position, SNPs where less than 90% of the calls agree, and SNPs with a read coverage of less than
544 5. Results were read from the output matrix of pairwise SNP distances.

545 BLAST analyses

546 The selection of plasmids included in the current analysis was based on a study by Chmielowska
547 *et al*. (58), which included 113 unique completely sequenced plasmid replicons from *Listeria* spp.
548 strains, of which 63 were assigned to 19 groups (different only by point mutations or small indels
549 of size <1 kb). In the current analysis, one plasmid each from the 19 groups (selecting the largest
550 plasmid in each group) and the 50 plasmids that did not belong to a group were included, in total
551 69 different plasmids (Table S8). All contigs >2000 bp from the *L. monocytogenes* genomes were
552 used as queries in a BLAST search (blastn v2.10.0+) against a local nucleotide BLAST database
553 created for the 69 *Listeria* spp. plasmids. A contig was viewed as a plasmid contig if the query
554 coverage (qcov) was ≥85%, and the percentage identity of the best alignment was >90%. Hits
555 were inspected manually.

556 A BLAST search for plasmid replication genes was carried out using one *repA* gene selected from
557 each of the eleven RepA groups (G1-G11) identified by Chmielowska *et al*. (58) as queries (Table
558 S9), against a local nucleotide BLAST database created for the *L. monocytogenes* genomes. The
559 same analysis was also performed using as queries the nine plasmid-encoded/associated genetic
560 elements involved in stress response identified by Schmitz-Esser *et al*. (63), as well as additional
561 genetic elements associated with stress response and resistance (Table S4) and the entire content
562 of the ResFinder database (59) downloaded on November 11, 2021. Only the best hit for each
563 query sequence in each genome was kept. When the minimum nucleotide identity was <99%,
564 and/or the length ratio of the query sequence relative to the match in the genome (length/qlen)

565 was ≠ 1, the alignments and contigs were manually inspected before a presence/absence gene
566 call was made. In addition, the protein sequences of the BLAST hits were aligned to the query
567 protein sequences. A call was made for the presence of the gene if the protein identity was >92%.
568 Differences in the nucleotide sequences that lead to alterations of the protein sequences, such
569 as insertions, deletions, premature stop codons or elongations, were recorded.

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

575 Calculation of weighted similarity scores

576 Genetic associations between pervasive isolates from different factories, weighted by their
577 similarity, and illustrated in the chord diagram in Fig. 5B were calculated as follows: All pairs of
578 isolates from different factories separated by 20 or fewer wgMLST allelic differences were
579 counted towards the total strength of links between factories. For these pairs, the genetic
580 distance D was converted to a similarity score $S = 21 - D$. Thus, a distance of 20 wgMLST alleles
581 corresponds to a similarity score of 1, and a distance of 0 wgMLST alleles corresponds to a
582 similarity score of 21. Of note, the lowest genetic distance separating two isolates from different
583 factories was 2. Then, all similarity scores were grouped by CC and factory and summed to
584 generate the final scores, which are represented by the thickness of each arc in Fig. 5B. The image
585 was created using the R package circlize (115).

586 Data availability

587 The raw data and assembled genomes for the 512 genomes sequenced in the current study has
588 been submitted to NCBI as BioProject accession PRJNA689484. For GenBank and Sequence Read
589 Archive (SRA) accession numbers for all 769 genomes from food processing industry, see Table
590 S1. The assemblies were annotated using the NCBI Prokaryotic Genomes Automatic Annotation
591 Pipeline (PGAAP) server.

592 Acknowledgments

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

597 References

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

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

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

606 4. **NicAogáin K, O'Byrne CP.** 2016. The role of stress and stress adaptations in determining the
607 fate of the bacterial pathogen *Listeria monocytogenes* in the food chain. Front Microbiol
608 **7:**1865.

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

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

614 7. **Belias A, Sullivan G, Wiedmann M, Ivanek R.** 2022. Factors that contribute to persistent
615 *Listeria* in food processing facilities and relevant interventions: A rapid review. Food Control
616 **133:**108579.

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

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

624 10. **Fagerlund A, Langsrud S, Møretrø T.** 2020. In-depth longitudinal study of *Listeria*
625 *monocytogenes* ST9 isolates from the meat processing industry: Resolving diversity and

626 transmission patterns using whole-genome sequencing. Appl Environ Microbiol **86:**e00579-
627 20.

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

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

634 13. **Morganti M, Scaltriti E, Cozzolino P, Bolzoni L, Casadei G, Pierantoni M, Foni E, Pongolini**
635 **S.** 2016. Processing-dependent and clonal contamination patterns of *Listeria*
636 *monocytogenes* in the cured ham food chain revealed by genetic analysis. Appl Environ
637 Microbiol **82:**822-831.

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

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

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

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

652 18. **Lüth S, Halbedel S, Rosner B, Wilking H, Holzer A, Roedel A, Dieckmann R, Vincze S, Prager**
653 **R, Flieger A, Al Dahouk S, Kleta S.** 2020. Backtracking and forward checking of human
654 listeriosis clusters identified a multiclonal outbreak linked to *Listeria monocytogenes* in
655 meat products of a single producer. Emerg Microbes Infect **9:**1600-1608.

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

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

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

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

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

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

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

680 26. **Cherifi T, Carrillo C, Lambert D, Miniaï I, Quessy S, Larivière-Gauthier G, Blais B, Fravalo P.**
681 2018. Genomic characterization of *Listeria monocytogenes* isolates reveals that their
682 persistence in a pig slaughterhouse is linked to the presence of benzalkonium chloride
683 resistance genes. BMC Microbiol **18:**220.

684 27. **Hingston P, Chen J, Dhillon BK, Laing C, Bertelli C, Gannon V, Tasara T, Allen K, Brinkman**
685 **FS, Truelstrup Hansen L, Wang S.** 2017. Genotypes associated with *Listeria monocytogenes*

686 isolates displaying impaired or enhanced tolerances to cold, salt, acid, or desiccation stress.
687 Front Microbiol **8:**369.

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

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

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

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

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

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

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

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

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

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

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

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

725 40. **Dutta V, Elhanafi D, Kathariou S.** 2013. Conservation and distribution of the benzalkonium
726 chloride resistance cassette *bcrABC* in *Listeria monocytogenes*. Appl Environ Microbiol
727 **79:**6067-6074.

728 41. **Nightingale KK, Windham K, Martin KE, Yeung M, Wiedmann M.** 2005. Select *Listeria*
729 *monocytogenes* subtypes commonly found in foods carry distinct nonsense mutations in
730 *inlA*, leading to expression of truncated and secreted internalin A, and are associated with
731 a reduced invasion phenotype for human intestinal epithelial cells. Appl Environ Microbiol
732 **71:**8764-8772.

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

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

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

741 45. **Piercey MJ, Hingston PA, Truelstrup Hansen L.** 2016. Genes involved in *Listeria*
742 *monocytogenes* biofilm formation at a simulated food processing plant temperature of 15
743 °C. Int J Food Microbiol **223:**63-74.

744 46. **Spanu C, Jordan K.** 2020. *Listeria monocytogenes* environmental sampling program in
745 ready-to-eat processing facilities: A practical approach. Compr Rev Food Sci Food Saf
746 **19:**2843-2861.

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

750 48. **Jagadeesan B, Gerner-Smidt P, Allard MW, Leuillet S, Winkler A, Xiao Y, Chaffron S, Van**
751 **Der Vossen J, Tang S, Katase M, McClure P, Kimura B, Ching Chai L, Chapman J, Grant K.**
752 2019. The use of next generation sequencing for improving food safety: Translation into
753 practice. Food Microbiol **79:**96-115.

754 49. **Zamudio R, Haigh RD, Ralph JD, De Ste Croix M, Tasara T, Zurfluh K, Kwun MJ, Millard AD,**
755 **Bentley SD, Croucher NJ, Stephan R, Oggioni MR.** 2020. Lineage-specific evolution and gene
756 flow in *Listeria monocytogenes* are independent of bacteriophages. Environ Microbiol
757 **22:**5058-5072.

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

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

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

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

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

775 55. **Chen Y, Gonzalez-Escalona N, Hammack TS, Allard MW, Strain EA, Brown EW.** 2016. Core
776 genome multilocus sequence typing for identification of globally distributed clonal groups
777 and differentiation of outbreak strains of *Listeria monocytogenes*. Appl Environ Microbiol
778 **82:**6258-6272.

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

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

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

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

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

793 61. **Parsons C, Lee S, Kathariou S.** 2020. Dissemination and conservation of cadmium and
794 arsenic resistance determinants in *Listeria* and other Gram-positive bacteria. Mol Microbiol
795 **113:**560-569.

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

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

802 64. **Harter E, Wagner EM, Zaiser A, Halecker S, Wagner M, Rychli K.** 2017. Stress Survival Islet
803 2, predominantly present in *Listeria monocytogenes* strains of sequence type 121, is
804 involved in the alkaline and oxidative stress responses. Appl Environ Microbiol **83:**e00827-
805 17.

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

809 66. **Jordan SJ, Perni S, Glenn S, Fernandes I, Barbosa M, Sol M, Tenreiro RP, Chambel L, Barata**
810 **B, Zilhao I, Aldsworth TG, Adriao A, Faleiro ML, Shama G, Andrew PW.** 2008. *Listeria*
811 *monocytogenes* biofilm-associated protein (BapL) may contribute to surface attachment of
812 *L. monocytogenes* but is absent from many field isolates. Appl Environ Microbiol **74:**5451-
813 5456.

814 67. **Popowska M, Krawczyk-Balska A, Ostrowski R, Desvaux M.** 2017. InlL from *Listeria*
815 *monocytogenes* is involved in biofilm formation and adhesion to mucin. Front Microbiol
816 **8:**660.

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

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

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

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

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

836 73. **Halbedel S, Wilking H, Holzer A, Kleta S, Fischer MA, Lüth S, Pietzka A, Huhulescu S,**
837 **Lachmann R, Krings A, Ruppitsch W, Leclercq A, Kamphausen R, Meincke M, Wagner-**
838 **Wiening C, Contzen M, Kraemer IB, Al Dahouk S, Allerberger F, Stark K, Flieger A.** 2020.
839 Large nationwide outbreak of invasive listeriosis associated with blood sausage, Germany,
840 2018-2019. Emerg Infect Dis **26:**1456-1464.

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

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

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

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

854 78. **Rogers K.** 2021**.** 8 Jun. 2021. German *E. coli* outbreak of 2011. *In* Encyclopedia Britannica.
855 https://www.britannica.com/event/German-E-coli-outbreak-of-2011. Accessed 13 May
856 2022.

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

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

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

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

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

875 84. **Møretrø T, Schirmer BCT, Heir E, Langsrud S.** 2017. *In situ* evaluation of citric acid powder
876 to control *Listeria monocytogenes* on floors in meat processing plants. Arch
877 Lebensmittelhyg **68:**84–87.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

920 100. **Harrand AS, Jagadeesan B, Baert L, Wiedmann M, Orsi RH.** 2020. Evolution of *Listeria*
921 *monocytogenes* in a food processing plant involves limited single-nucleotide substitutions
922 but considerable diversification by gain and loss of prophages. Appl Environ Microbiol
923 **86:**e02493-19.

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

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

930 103. **Assisi C, Forauer E, Oliver HF, Etter AJ.** 2021. Genomic and transcriptomic analysis of biofilm
931 formation in persistent and transient *Listeria monocytogenes* isolates from the retail deli
932 environment does not yield insight into persistence mechanisms. Foodborne Pathog Dis
933 **18:**179-188.

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

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

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

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

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

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

952 110. **Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K.** 2021. cluster: Cluster analysis
953 basics and extensions. R package version 2.1.1, https://CRAN.R-
954 project.org/package=cluster.

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

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

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

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

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

965 # Tables

966 **Table 1: Overview of *L. monocytogenes* isolates from food industry included in this study[a]**

Source of sample	Meat processing factory										Salmon processing factory						Other sources[b]
	M1	M2	M3	M4	M5	M6	M7	M8	M9	other	S1	S2	S3	S4	S5	S6	
Food processing environment	141	14	12	176	4	30	4	23	8	1	24	55	2	3	20	128	1
Food processing - low hygienic zone	7						6		3	3							
Food product				15		4		1		1		7				19	2
Food sample taken during processing			4	1													
Raw materials			2								1	6				41	

967 [a] For information on individual isolates, see Supplementary Table S1.

968 [b] Other sources were salmon and cheese products and a sample from a domestic kitchen

969 **Table 2: Comparison of number of differences detected in each CC using SNP and wgMLST**
970 **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

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

973 **Table 3: 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).

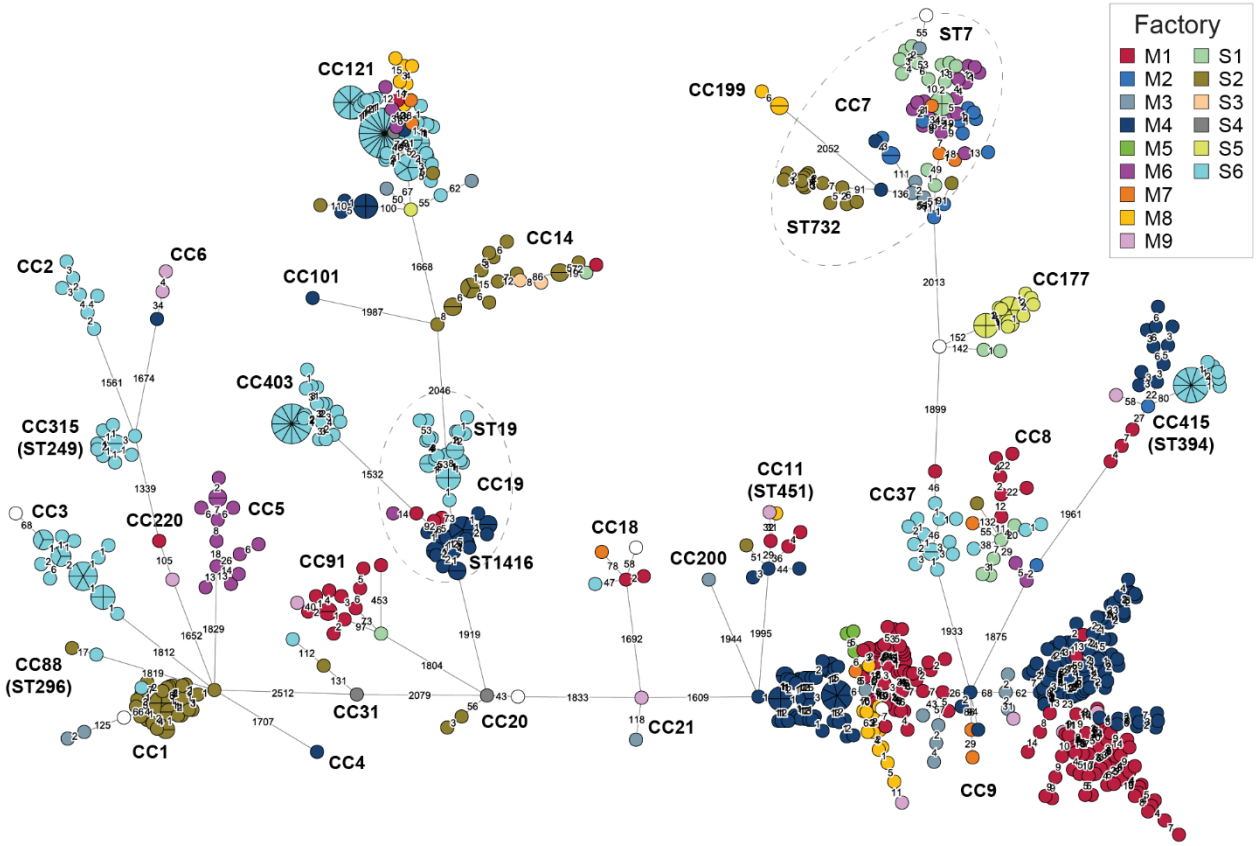
974

975 **Table 4. *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

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

980 # Figure legends



981

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

988

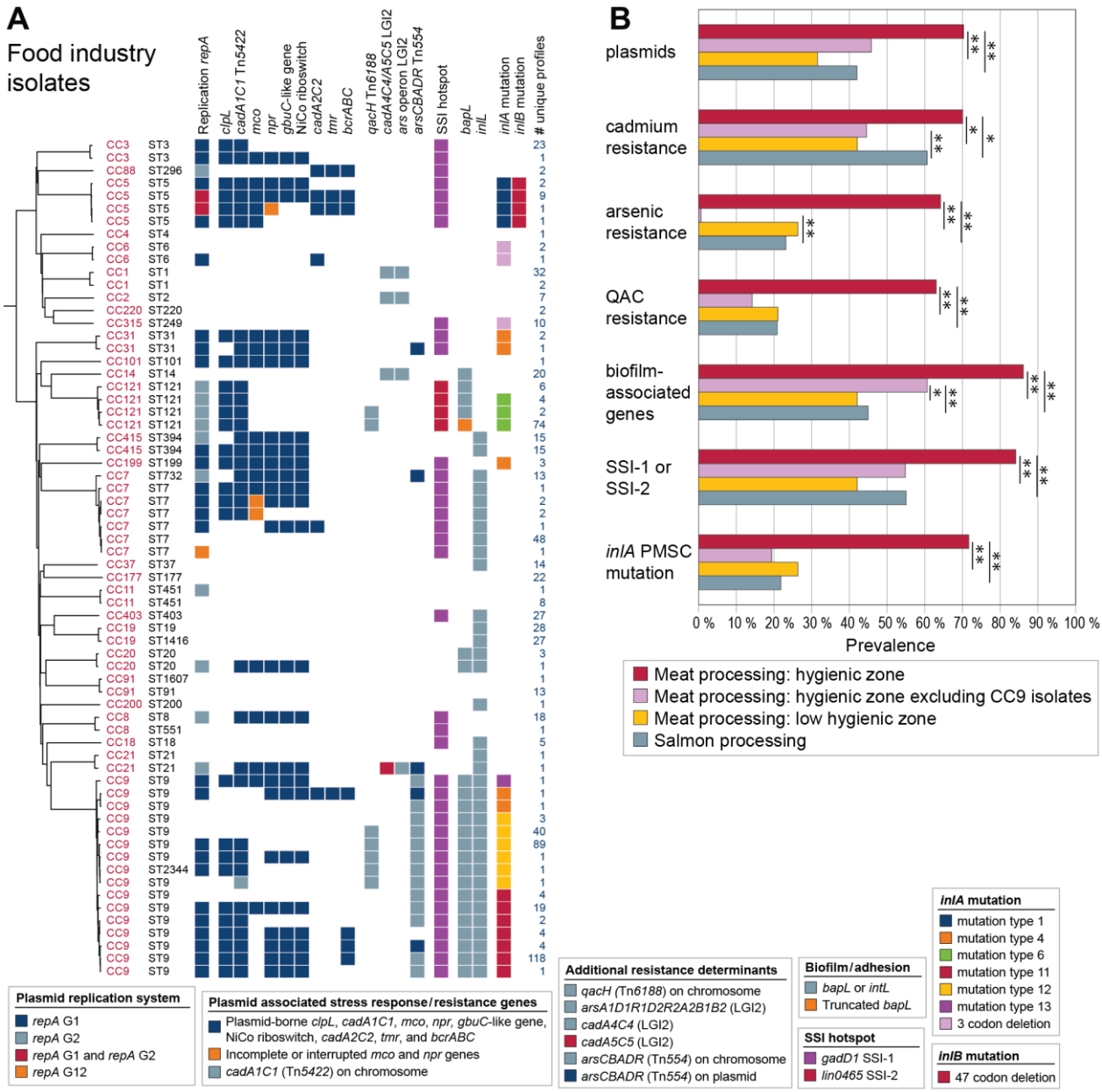
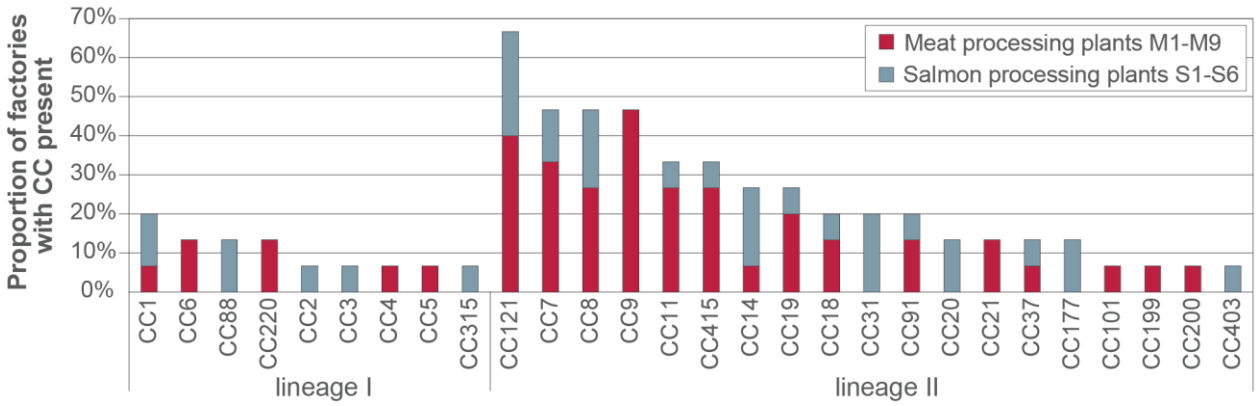


Figure 2: Presence of accessory genetic determinants associated with stress survival, resistance, or persistence in *L. monocytogenes* from food processing environments. **A)** Unique combinations of ST and the variable stress response loci in isolates from food processing environments. All 769 isolates are shown individually in Fig. S1. The phylogeny is a midpoint-rooted Neighbor-Joining tree based on wgMLST analysis and shows one arbitrarily selected genome from each of the groups of genomes containing the same unique gene combination. The number of genomes harboring the same unique combination is indicated in the right column. **B)** Prevalence of genetic determinants in isolates from different sources. Cadmium resistance; *cadA1C1*, *cadA2C2*, *cadA4A4* or *cadA5C5*, arsenic resistance; *arsA1D1R1D2R2A2B1B2* on LGI2 or *arsCBADR* on Tn554-like transposon, QAC resistance; *bcrABC* or *qacH*, biofilm-associated genes; *bapL* and/or *inlL*. Asterisks represent significant differences (Pearson's chi-squared test. *, $p < 0.03$; **, $p \leq 0.001$, see also Table S5).

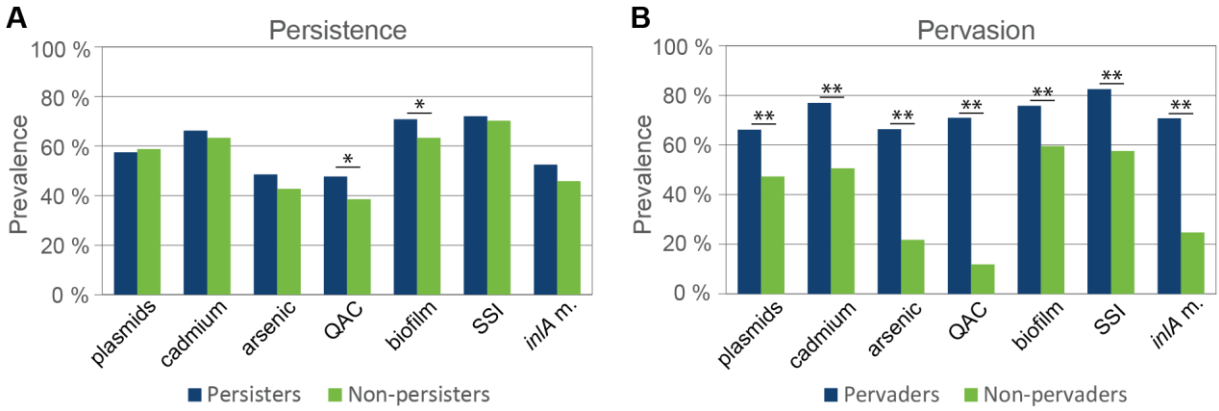


1002

1003 **Figure 3:** Different CCs detected in varying proportions of factories. The data is presented as a

1004 stacked bar plot showing the percentage of examined meat and salmon processing plants in which

1005 each CC was detected. A total of 15 factories were included.



1006

1007 **Figure 4**: Prevalence of subgroups of accessory genetic determinants associated with stress

1008 survival, resistance, and persistence in *L. monocytogenes* classified as **A)** persistent and/or **B)**

1009 pervasive. Cadmium; cadmium resistance loci *cadA1C1*, *cadA2C2*, *cadA4C4*, or *cadA5C5*, arsenic;

1010 arsenic resistance operons *arsA1D1R1D2R2A2B1B2* on LGI2 or *arsCBADR* on the Tn554-like

1011 transposon, QAC; QAC resistance loci *qacH* or *bcrABC*, biofilm; biofilm-associated genes *bapL*

1012 and/or *inlL*, SSI; stress survival islets SSI-1 or SSI-2, and *inlA* m.; PMSC mutation in the *inlA* gene.

1013 Asterisks represent significant differences (Pearson's chi-squared test. *, $p < 0.05$; **, $p < 0.001$,

1014 see also Table S6).

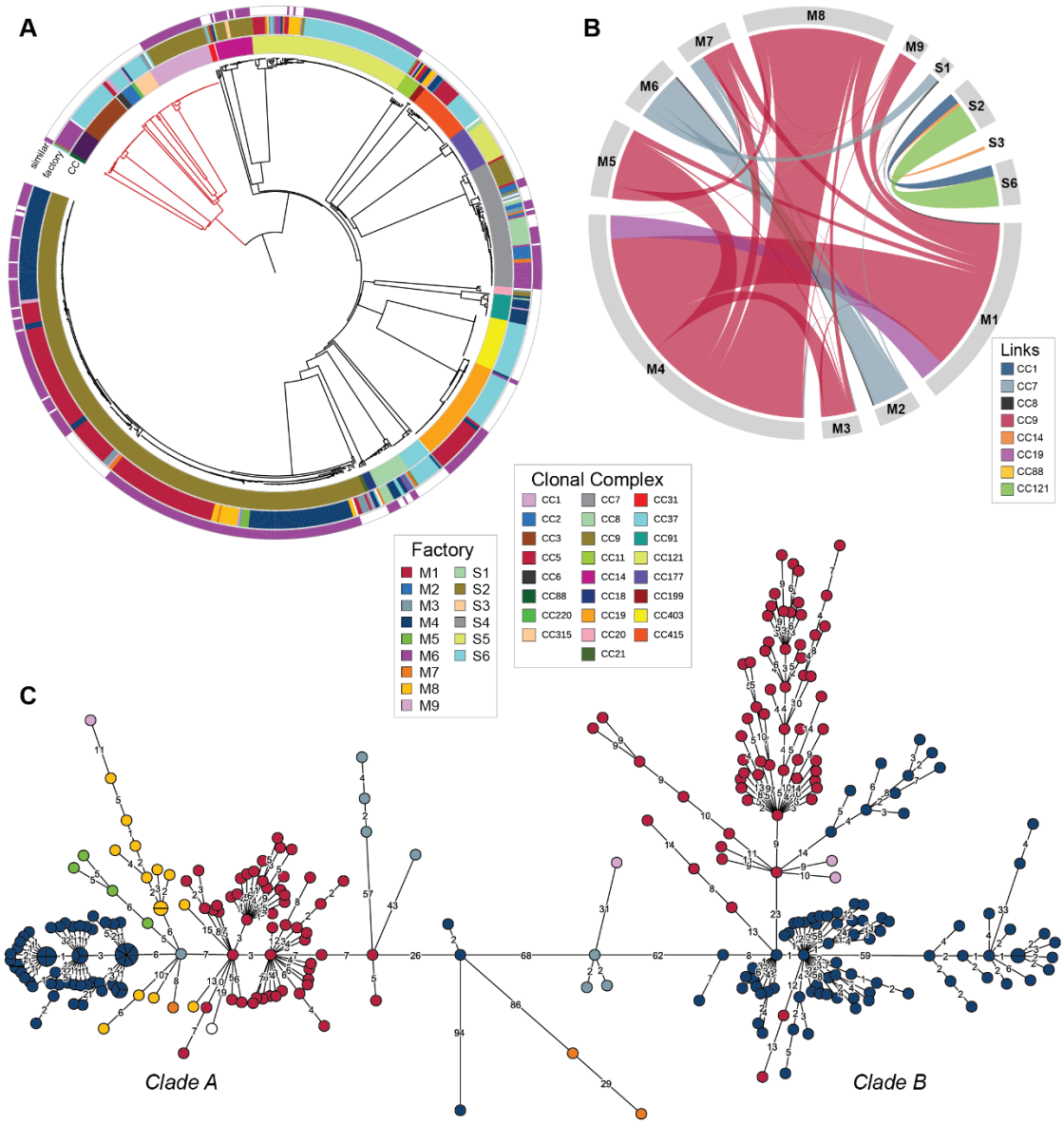


Figure 5: Pervasive strains, present in more than one processing plant. **A)** Neighbor-Joining phylogenetic tree based on wgMLST analysis, showing CC (inner ring), factory of origin (middle ring), and isolates for which genetically similar isolates (≤20 wgMLST allelic differences) were found in at least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

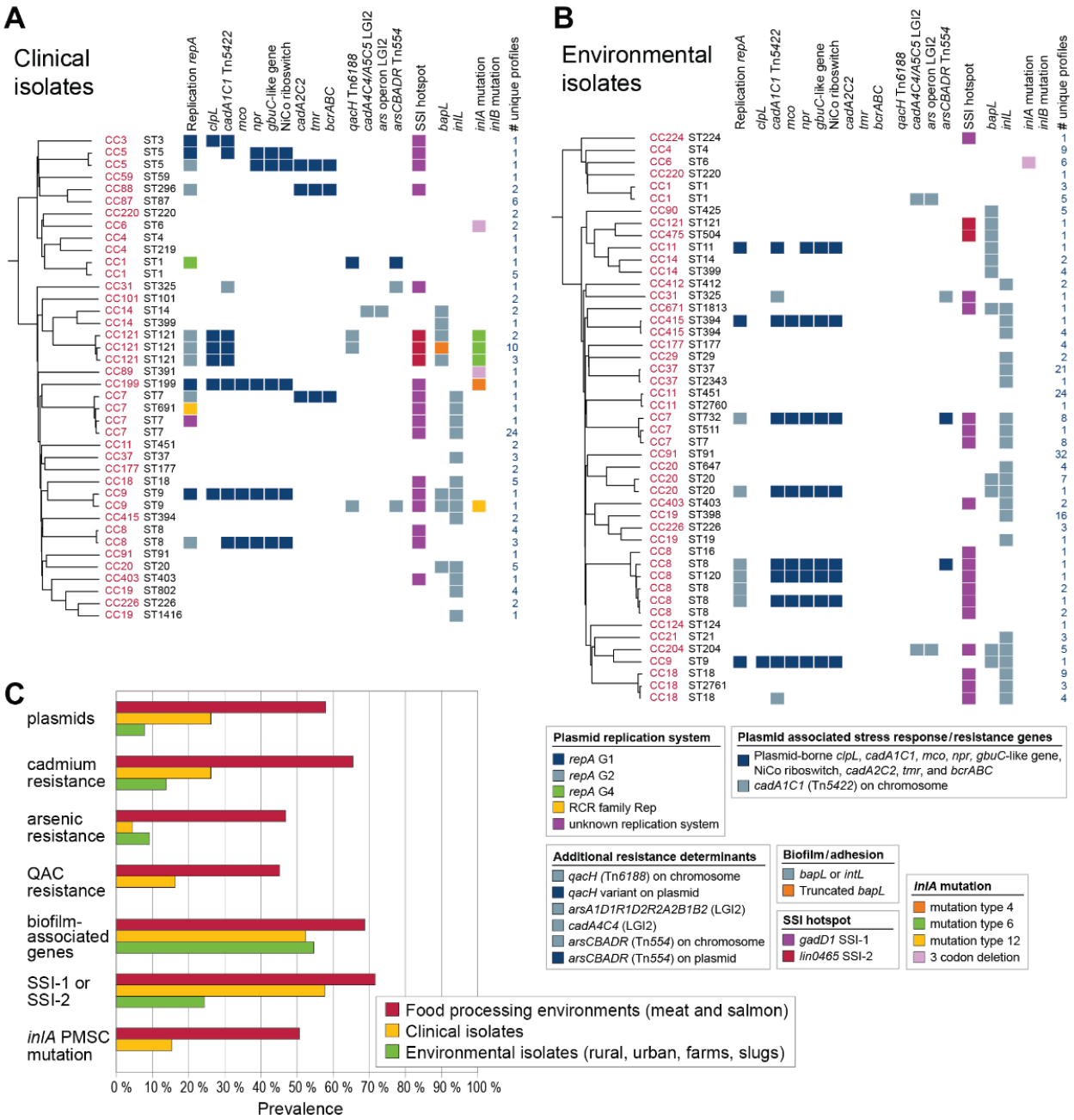


Figure 6: Presence of accessory genetic determinants associated with stress survival, resistance, or persistence in *L. monocytogenes* from different sources. **A)** Norwegian clinical isolates 2010-2015 (56), and **B)** isolates from natural (rural/urban/farm/slug) environments in Norway (2). For each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