International Journal of Food Microbiology 410 (2024) 110482

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Genomic analysis of *Listeria monocytogenes* CC7 associated with clinical infections and persistence in the food industry

Trond Møretrø [*], Eva Wagner [1], Even Heir, Solveig Langsrud, Annette Fagerlund

Nofima, Norwegian Institute of Food, Fisheries and Aquaculture Research, N-1430 Aas, Norway

ARTICLE INFO

Keywords:
Listeriosis
ST7
Whole genome sequencing
WGS
Pervasion
Virulence
Food processing

ABSTRACT

Listeria monocytogenes clonal complex 7 (CC7), belonging to lineage II, is the most common subtype among clinical listeriosis isolates in Norway, and is also commonly found in Norwegian food industry and outdoor environments. In the present study, the relative prevalence of CCs among clinical isolates of *L. monocytogenes* in European countries during 2010–2015 was determined. Then, phylogenomic and comparative genomic analyses was performed for 115 Norwegian and 255 international reference genomes from various sources, to examine potential explanations underlying the high prevalence of CC7 among Norwegian listeriosis cases. Selected isolates were also compared using *in vitro* virulence assays.

The results showed a high relative prevalence of CC7 in clinical isolates from Norway and the neighboring Nordic countries Sweden and Finland. In contrast to in most other European countries, lineage II dominated among clinical isolates in these countries. Phylogenetic analysis of the 370 CC7 isolates separated the genomes into four clades, with the majority of Norwegian isolates (69 %) clustered in one of these clades, estimated to have diverged from the other clades around year 1830. The Norwegian isolates within this clade were widely distributed in different habitats; several (poultry) meat processing factories, a salmon processing plant, in nature, farms, and slugs, and among human clinical isolates. In particular, one pervasive CC7 clone was found across three poultry processing plants and one salmon processing plant, and also included three clinical isolates.

All analysed CC7 isolates harbored the same set of 72 genes involved in both general and specific stress responses. Divergence was observed for plasmid-encoded genes including genes conferring resistance against arsenic (Tn*554-arsCBADR*), cadmium (*cadA1C1* and *cadA2C2*), and the biocide benzalkonium chloride (*bcrABC*). No significant difference in prevalence of these genes was seen between isolates from different habitats or sources. Virulence attributes were highly conserved among the CC7 isolates. *In vitro* virulence studies of five representative CC7 isolates revealed a virulence potential that, in general, was not significantly lower than that of the control strain EGDe, with isolate-dependent differences that could not be correlated with genetic determinants.

The study shows that CC7 is widespread in Norway, and that a pervasive CC7 clone was present in food processing plants. The study highlights the importance of CC7 and lineage II strains in causing listeriosis and shows that more research is needed to understand the reasons behind geographical differences in CC prevalence.

1. Introduction

As the causative agent of listeriosis, the foodborne bacterial human pathogen *Listeria monocytogenes* is of significant concern for food and consumer safety. Listeriosis is acquired upon the consumption of insufficiently heated or ready-to-eat food products, mostly involving unpasteurized milk and dairy products, soft cheeses, deli meats, fish products and fresh produce (Desai et al., 2019). *L. monocytogenes* may establish itself in food processing plants and contaminate food during processing. *L. monocytogenes* is challenging to eradicate and may persist in food processing environments for years (Ferreira et al., 2014).

L. monocytogenes can be divided into groups using genetic typing methods. Multilocus sequence typing (MLST) has become a common tool to group genetically similar strains into sequence types (STs) based on the sequences of seven housekeeping genes. Similar STs are in turn grouped into clonal complexes (CCs) (Ragon et al., 2008).

* Corresponding author.
E-mail address: trond.moretro@nofima.no (T. Møretrø).
[1] Present address: Institute of Food Science, University of Natural Resources and Life Sciences, Vienna, Muthgasse 18, 1190 Vienna, Austria.

https://doi.org/10.1016/j.ijfoodmicro.2023.110482
Received 7 May 2023; Received in revised form 3 November 2023; Accepted 5 November 2023
Available online 10 November 2023

L. monocytogenes strains of lineage I that belong to CCs which are predominantly isolated from human listeriosis cases (CC1, CC2, CC4, CC6) have been described to be infection-associated and hypervirulent (FAO and WHO, 2022; Maury et al., 2016) and appear to have a genetic advantage during the passage in the human host as they encode accessory virulence factors in addition to the core virulence factors. Other CCs (*e.g.*, CC9, CC121) belonging to lineage II, have been described to be mostly food-associated (Maury et al., 2016) and have been suggested to be genetically better equipped for food processing conditions as compared to passage in the human host (Elhanafi et al., 2010; Harter et al., 2017; Müller et al., 2014). CCs that are neither associated with infections nor with food have been described to be intermediate. In the study conducted by Maury et al. (2016), these CCs included but were not limited to CC3, CC5 (lineage I), CC7, CC8, CC16 or CC155 (lineage II). Based on above mentioned and similar data, a subtype specific virulence ranking was proposed by the FAO and WHO (2022).

In Norway, CC7, belonging to lineage II, has been involved in a cheese related outbreak (Johnsen et al., 2010), and was also the most common CC among clinical isolates of *L. monocytogenes* during the period 2010–2015 (Fagerlund et al., 2022a), as well as during recent years: Of the 53 *L. monocytogenes* isolates collected and analysed at the Norwegian Institute of Public Health during 2020 and 2021, CC7 was the most common, constituting 17 % of the isolates (The Norwegian Directorate of Health, 2022). Among isolates collected from Norwegian meat and salmon processing environments, CC7 was the third most commonly detected CC following CC9 and CC121, and only CC121 was found in a higher number of individual food processing plants (Fagerlund et al., 2022b). Recently we described certain strains of CC7, CC9, CC121 and other CC-types as pervasive, as these strains were found to persist in multiple food processing plants (Fagerlund et al., 2022b). In Norwegian rural and urban environments, dairy farms, and slugs, CC7 was reported as the fourth most common CC following CC91, CC11 and CC37 (Fagerlund et al., 2022a).

Whole genome sequencing (WGS) is the new gold standard for typing of bacteria, including *L. monocytogenes*. WGS makes it possible to extract the genetic information required to determine the ST, establish phylogenetic relationships between strains based on their whole genome sequence, and determine the content of virulence- and stress-associated genes (Pietzka et al., 2019). In many industrialized countries, authorities have implemented WGS-based surveillance and outbreak detection for foodborne pathogens (European Food Safety Authority (EFSA) et al., 2022; Stevens et al., 2022). However, within the food industry, there has been limited implementation of WGS. There is still limited published information regarding the diversity among *L. monocytogenes* within and between processing plants and those isolated from other environments, including *L. monocytogenes* from human patients. Such information is valuable for correct interpretation of WGS data and to further understand transmission routes and the virulence potential of *L. monocytogenes*.

L. monocytogenes is able to tolerate and adapt to a broad range of conditions to which it is exposed to in food production plants during processing, preservation, or implementation of hygienic measures. These include temperature ranges, acidic, alkaline, oxidative, desiccation and osmotic stresses and the exposure to cleaning and disinfection agents like quaternary ammonium compounds (QACs), peracetic acid or sodium hypochlorite (Bucur et al., 2018). It has been suggested that certain clones or strains have certain genotypic or phenotypic traits that enable them to persist in the food industry (Fagerlund et al., 2022b; Ferreira et al., 2014; Palaiodimou et al., 2021).

In the present study, we first compared the relative prevalence of *L. monocytogenes* CC7 among clinical isolates from different European countries, using the available datasets from the period 2010 to 2015, to determine whether the prevalence of CC7 was higher in Norway compared with other European countries. As the results confirmed this hypothesis, we tried to determine whether this was associated with i) phylogenetically distinct Norwegian CC7 clones relative to those from other countries, ii) a propensity for the Norwegian CC7 isolates to persist and spread in food processing environments (and if so), iii) whether persistence was associated with specific adaptations enabling survival in these environments, and/or iv) a relatively high virulence potential of the Norwegian CC7 isolates, as determined by virulence gene content and *in vitro* virulence potential. As part of this work, the population structure and content of relevant genetic determinants was determined for a dataset containing 115 isolates of *L. monocytogenes* CC7 isolates from the Norwegian meat and salmon industry, dairy farms, urban and rural environments and clinical cases, as well as for 255 publicly available international CC7 reference genomes included for comparative purposes. Furthermore, transmission patterns for CC7 clones within Norwegian food chains were analysed.

2. Materials and methods

2.1. Genomic MLST analyses

Classical MLST analysis, used to determine STs and CCs, followed the MLST scheme described by Ragon et al. (2008) and the database maintained at the Institute Pasteur's *L. monocytogenes* online MLST repository (https://bigsdb.pasteur.fr/listeria/). *In silico* MLST typing was performed for raw sequencing data using the program available at https://bitbucket.org/genomicepidemiology/mlst (Larsen et al., 2012), and for genome assemblies using the program available at https://github.com/ tseemann/ mlst. The whole genome MLST analysis (wgMLST) analysis was performed using a whole-genome schema containing 4797 coding loci from the *L. monocytogenes* pan-genome and the assembly-based BLAST approach, implemented in BioNumerics 7.6 (Applied Maths, 2015). The core genome MLST (cgMLST) analysis was performed using the scheme described by Moura et al. (2017) and the allele database maintained at the Institute Pasteur's *L. monocytogenes* online MLST repository. This schema is a 1748 loci subset of the employed wgMLST schema. cgMLST results were either extracted from the wgMLST data *via* a tool in BioNumerics which synchronizes the allele mapping nomenclature with that used in the public BIGSdb-*Lm* database (synchronized on 25.10.2023), or obtained directly by sequence query of assembled genomes at the BIGSdb-*Lm* web server (https://bigsdb.pasteur.fr/listeria) (performed in August 2021). Both approaches use BLAST to identify cgMLST alleles. As the cgMLST analysis was based on queries against a public cgMLST database, any allelic variant present at a specific cgMLST locus in a genome, but not recorded as an allele variant in the public database at the time of query, will be part of the wgMLST profile for the genome in question, but not the cgMLST profile. Minimum spanning trees were constructed using BioNumerics 7.6 based on the categorical differences in the allelic cgMLST profile ("cgMLST_Core Pasteur" character set) or wgMLST profile ("wgMLST (wgMLST loci)" character set) for each isolate. Loci with no allele calls were not considered in the pairwise comparison between two genomes. The number of pairwise allelic differences between isolates was read from genetic distance matrices computed from the absolute number of categorical differences between genomes.

2.2. Selection of genomes for comparative genomic analyses

To determine the prevalence of CCs in European Union (EU) and European Economic Area (EEA) countries during 2010–2015, raw Illumina genome sequencing data was downloaded from the NCBI Sequence Read Archive (SRA) for the BioProjects listed in Supplementary Table S1 on 10.06.2022, mainly from the study by Van Walle et al. (2018). Only genomes annotated to be from human clinical sources and from the time-period 2010–2015 were included. In total 3629 genomes were analysed, of which the CC could be determined for all except 11 genomes (Supplementary Table S2).

The sources of the 370 CC7 genomes used in cgMLST and BLAST analyses are listed in Table 1 and Supplementary Table S3. Of these, 115

Table 1
Overview of the genomes analysed in the current study.

Location	Source	Number of genomes included in cgMLST and BLAST analysis	Number of genomes included in SNP and BEAST analysis*	BioProject, data source and/or reference
Norway	Meat industry	37	37	PRJNA689484 (Fagerlund et al., 2022b)
	Salmon industry	30	30	PRJNA689484 (Fagerlund et al., 2022b)
	Soft cheese outbreak	1	1	PRJNA689484 (Fagerlund et al., 2022b)
	Salmon plants	3	0	PRJNA609411 (Løvdal et al., 2021)
	Farm	3	3	PRJNA744724 (Fagerlund et al., 2022a)
	Rural or urban	11	11	PRJNA689486 (Fagerlund et al., 2022a)
	Invading garden slugs (*Arion vulgaris*)	3	3	PRJNA689487 (Fagerlund et al., 2022a)
	Clinical	27	27	PRJEB26061 (Van Walle et al., 2018)
Other countries	BIGSdb-*Lm* database cgMLST data	51	32	https://bigsdb.pasteur.fr/listeria/
	NCBI GenBank	120	54	https://www.ncbi.nlm.nih.gov/genome/browse#!/prokaryotes/listeria%20monocytogenes
	ListAdapt	84	82	PRJEB38828 (Félix et al., 2022)
Total		*370*	*280*	

* Excluded genomes (relative to genomes included in cgMLST) were those for which raw Illumina fastq data and/or a collection year or date was not available. For details see Supplementary Table S3.

were from Norway, including genomes from isolates from food industry (six meat (M) processing plants, four salmon (S) processing plants, and a small-scale cheese processing plant), rural/urban/natural environments, and human clinical cases (Fagerlund et al., 2022a; Fagerlund et al., 2022b; Løvdal et al., 2021; Van Walle et al., 2018). A total of 255 international reference genomes were retrieved. For the genomes from the BIGSdb-*Lm* database the isolates for which cgMLST typing data was available on 25.08.2021 were included. For the genomes from NCBI Genbank, the included genomes were the CC7 genomes identified among the 3926 *L. monocytogenes* genomes present as of 27.08.2021, except those already included as part of the Norwegian or BIGSdb-*Lm* datasets or representing resequencing of the same isolate. For the ListAdapt project dataset (Félix et al., 2022), the 1552 genomes found in the Sequence Read Archive (SRA) for ListAdapt on 04.05.2021 were typed using MLST-KMA (Clausen et al., 2018). The genomes identified as belonging to CC7 were then subjected to *de novo* genome assembly using SPAdes v3.13.0 (Bankevich et al., 2012) as previously described (Fagerlund et al., 2022b). The three *Norwegian* CC7 ListAdapt genomes were excluded as they represented resequencing of garden slug isolates.

2.3. Character-based phylogenetic inference

2.3.1. Generation of a complete reference genome for SNP analysis

To generate a complete genome for MF2133 for use as a reference genome in the SNP analysis, a hybrid genome assembly was generated using Unicycler v0.3.0b (https://github.com/rrwick/Unicycler) (Wick et al., 2017) and default settings, using the Illumina data previously generated on a MiSeq platform with 300-bp paired-end reads (Fagerlund et al., 2022b) and Nanopore data generated as follows: DNA prepared as previously described (Fagerlund et al., 2022b) was prepared using the SQK-RAD004 Rapid Sequencing Kit (Oxford Nanopore Technologies) and sequenced on a FLO-FLG001 Flongle flow cell in a MinION Mk1 sequencing device with MinKNOW UI 4.0.20 software. Reads in fast5 format obtained from MinKNOW were basecalled using guppy v4.2.2 software (Oxford Nanopore Technologies) with default settings including a qscore filtering of 7. Read quality was assessed using NanoComp (De Coster, 2020; De Coster et al., 2018). PoreChop (Wick, 2018) was used to remove adapters. Reads >1000 bp were used in generation of the assembly. Coverage was calculated using BBmap v.36.92 for Illumina reads (Bushnell, 2014) and GraphMap v0.5.2 for Nanopore reads (Sovic et al., 2016). Raw Nanopore read data and an updated genome was submitted to SRA and GenBank (BioSample SAMN17224255). The assembly confirmed our previous result (Fagerlund et al., 2022b) that MF2133 lacked plasmids, while analysis using the Phaster tool (Arndt et al., 2016) showed that MF2133 contained two intact prophages, inserted into the *comK* and the tRNA-Arg^{TCT} loci.

2.3.2. SNP analysis

All genomes for which raw Illumina data were available and for which an isolation date or year was available were included in the SNP analysis. This included all the Norwegian isolates except three from salmon industry, and 168 of the international reference genomes; 280 in total (Table 1 and Supplementary Table S3). Raw Illumina reads were filtered on q15 and trimmed of adaptors using fastq-mcf from the ea-utils package (Aronesty, 2011). The reference based SNP analysis was performed using the CFSAN SNP pipeline v2.1.1 (Davis et al., 2015) and default filtering settings (including removal of SNPs where there are >3 SNPs in a 1000 base window, >2 SNPs in a 125 base window, and >1 SNP in a 15 base window), except that regions of high-density SNPs were defined for each sample individually instead of filtering a dense region found in any genome from all genomes. The concatenated SNP alignment (of length 5110 bases) and the matrix of pairwise SNP distances generated after removal of abnormal SNPs was used in the downstream analyses. The degree of genetic variation was calculated from the pairwise SNP distance matrix.

2.3.3. Model selection and ML analysis

Evaluation of models for sequence evolution was performed using the advanced model selection option in the ModelFinder program (Kalyaanamoorthy et al., 2017) implemented in the software package IQ-TREE2 v2.1.2 (Minh et al., 2020), with the nucleotide frequency count for the MF2133 reference genome input using the -fconst option. The best model selected using the Akaike and corrected Akaike information criteria (AIC and AIC_c) and the Bayesian information criterion (BIC), *i.e.*, K3Pu + F + R7, was used in the downstream ML analysis. Maximum Likelihood (ML) phylogenetic inference was reconstructed from the concatenated SNP alignment using IQ-TREE2 with 1000 rounds of non-parametric bootstrap test run to generate node supports (−b 1000 option), and with input of the nucleotide frequency count for the reference genome.

2.3.4. Bayesian tip-dated phylogenetic analysis

Timed phylogeny was reconstructed from the concatenated SNP alignment using the Bayesian Markov chain Monte Carlo (MCMC) method available in BEAST v1.10.4 (Suchard et al., 2018). The sampling date for each isolate was entered as the tip date. For isolates where only the sampling year was known, the sampling date was set to July 1st and the uncertainty value associated with the date was set to 1. For isolates where only the sampling month and year were known, the 1st of the month was used. The most recently collected isolate was sampled on

November 15th, 2020, which was set as the current date. Selected taxon sets were defined based on the ML phylogeny (monophyletic clades of interest with ≥90 % bootstrap values) in order to estimate the time of the most recent common ancestor (MRCA) for subclades. Settings used were the GTR + F + R7 nucleotide substitution model (the best model available in BEAST selected by BIC in ModelFinder), a strict molecular clock, and a coalescent constant size tree prior with a UPGMA starting tree. The analysis was run with a chain length of 10^9, a sampling frequency of 1000, and a 5 % burn-in. The effective sample size (ESS) values and trace files were monitored to ensure sufficient mixing (ESS of >700 for all statistics). The MCMC output was summarized as a maximum clade credibility tree, *i.e.*, the tree with the highest product of the posterior probabilities of all its nodes. The Interactive Tree Of Life (iTOL) v6.5.2 (Letunic and Bork, 2021), FigTree v1.4.4 (https://github.com/rambaut/figtree/releases), and Adobe Illustrator v26.3.1 was used for visualization.

2.4. BLAST analyses

In a previous study (Wagner et al., 2022), we performed BLAST analysis to search for 99 virulence-associated genes and *inlA* premature stop codon (PMSC) variants in the 68 Norwegian CC7 genomes from food processing environments (Table 1; (Fagerlund et al., 2022b)). In another study, (Fagerlund et al., 2022b), we performed BLAST analysis to identify plasmids, plasmid replication, stress response, and resistance genes in these 68 genomes, as well as for the 17 Norwegian CC7 genomes from natural environments (Fagerlund et al., 2022a) and the 27 Norwegian clinical CC7 isolates (Van Walle et al., 2018). This analysis included a BLAST search for the eleven plasmid replication genes identified by Chmielowska et al. (2021), the nine plasmid-encoded/associated genetic elements involved in stress response identified by Schmitz-Esser et al. (2021) as well as around 90 additional genetic elements associated with stress response and resistance (Fagerlund et al., 2022b) and the entire content of the ResFinder database (Zankari et al., 2012) downloaded on November 11th, 2021. In the current study we extended these analyses to cover all 370 CC7 genomes from both Norwegian and international sources examined in the current study. The genes used as queries in the BLAST analyses are listed in Supplementary Tables S4 and S5. The analysis was performed using blastn v2.10.0+ and a local nucleotide BLAST database created for the *L. monocytogenes* genomes. Only the best hit for each query sequence in each genome was kept. When the minimum nucleotide identity was <99 %, and/or the length ratio of the query sequence relative to the match in the genome (length/qlen) was ≠ 1, the alignments and contigs were manually inspected before a presence/absence gene call was made. In these cases, a BLAST search was performed manually in CLC Main Workbench 20.0.3 (Qiagen) and the protein sequences of the BLAST hits were aligned to the query protein sequences. If the gene was found in the manual BLAST search and the protein identity was >95 %, a call was made for the presence of the gene. Alterations in the nucleotide sequences that lead to alterations in the length of the protein sequences, such as insertions, deletions, premature stop codons or elongated sequences were described as such.

2.5. In vitro virulence assays

For the *in vitro* virulence gentamicin protection assays, five *L. monocytogenes* isolates from different Norwegian food sources and food processing factories and the reference lab strain EGDe were selected (Table 2). The *in vitro* virulence gentamicin protection assays using human intestinal epithelial Caco2 cells, human HEPG2 hepatocytes and human macrophage-like THP1 cells and single *L. monocytogenes* isolates were performed as previously described (Wagner et al., 2022) with the following modification: for the incubation step at 37 °C for 2 h prior to the infection of the cells, the OD_{600} of the bacterial cultures was adjusted to 0.1 in 5 ml BHI supplemented with 5 g/l yeast extract. Colony forming units/ml (CFU/ml) were determined by plating serial dilutions of the inoculum and of lysed and harvested cells after 45 min (invasion) and 4 h (intracellular proliferation) in medium containing gentamicin on tryptic soy agar (TSA). The invasion efficiency and the intracellular proliferation were calculated as CFU/ml recovered after 45 min and 4 h of incubation in medium containing gentamicin, respectively, divided by the mean CFU/ml of the inoculum and multiplied by 100 to obtain percentage values.

Statistical analysis was performed using Minitab v.19 software. Percentage values for the invasion efficiency and intracellular proliferation of each strain were $\log_{10}$-transformed to obtain normal distribution of the data prior to analysis and compared separately for each used cell line using one-way ANOVA with the strain as factor. Tukey's *post hoc* test for pairwise comparisons was performed to identify significant differences between the strains. Significant differences were reported at the level of significance of $p < 0.05$.

Table 2
Isolates included in *in vitro* virulence assays.

Isolate	Source	Isolation year	Phylogeny*	Other characteristics
MF3638	Camembert cheese	2007	Subclade A9	Outbreak-associated isolate from 2007 (Johnsen et al., 2010)
MF3995	Factory S2	2011	Clade C and ST732	Carries a RepA G2 plasmid
MF4991	Factory M2	2012	Subclade A12	Carries a RepA G1 plasmid
MF5050	Factory S1	2013	Subclade A17	Carries a shorter Ami variant
MF7586	Factory M7	2019	Subclade A15	Within the main poultry-associated cluster
EGDe	Lab strain		CC9	Used as reference strain

* See Supplementary Fig. S1 for location of individual CC7 isolates in the phylogenetic tree.

3. Results

3.1. Prevalence of CCs across European countries

L. monocytogenes CC7 was previously found to be the most frequent CC detected among clinical isolates in Norway during 2010–2015 (24 %) (Fagerlund et al., 2022a; Van Walle, et al., 2018). In order to examine whether this dominance was unique to Norway, the CCs of available clinical genomes from other European countries during the same years were determined (Fig. 1 and Supplementary Tables S1 and S2). The results showed that for countries where WGS data was available for >20 % of the clinical isolates, Norway had the highest relative prevalence of CC7 followed by its two neighbor Nordic countries, Sweden and Finland, which had a relative prevalence of CC7 of 21 % and 15 %, respectively. For Norway, Finland, and Sweden, <20 % of the listeriosis cases were caused by lineage I isolates, while for most other countries this proportion was much higher, with the highest proportions for France (99 %) and Cyprus (100 %). Norway was also characterized with a high relative prevalence of CC121 (13 %) and an absence of CC155 among clinical isolates.

3.2. Examination of phylogenetic relationship between CC7 isolates

To examine whether the high prevalence of CC7 among clinical isolates in Norway was associated with genetically distinct Norwegian CC7 clones, we analysed a set of CC7 genomes including 115 from various types of environments in Norway isolated during 1991–2020, and 255 reference genomes from other countries (Table 1 and Supplementary Table S3). The Norwegian CC7 isolates were from the meat and salmon industry, dairy farms, urban and rural environments and clinical

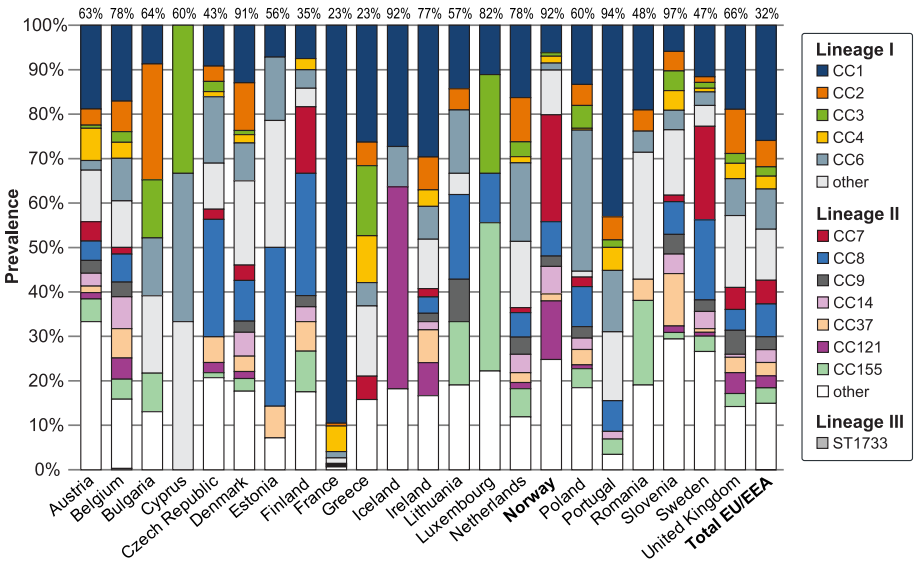

Fig. 1. Prevalence of CCs among clinical isolates of *L. monocytogenes* from EU/EEA countries during 2010–2015. The proportion of analysed genomes among all reported cases for each country is given above each bar. Data from individual countries where the proportion of analysed genomes was <20 % is not displayed in the chart, but these data are included in in the bar "Total EU/EAA" and in Supplementary Table S1. CCs with prevalence <2 % in the total dataset were combined in the "other" categories (split by lineage I and lineage II). One single lineage III isolate was detected (from Belgium).

cases. The majority of these isolates belonged to ST7, while 21 were ST732 isolates from a salmon processing factory and the surrounding urban environment, one ST511 was from a dairy farm, and one clinical isolate belonged to ST691. The international CC7 reference genomes were retrieved from the BIGSdb-*Lm* and NCBI Genbank databases and from a reference dataset comprising European isolates from animal and environmental sources and from food (Félix et al., 2022).

3.2.1. *cgMLST and wgMLST analysis*

The result from the cgMLST analysis is shown in Fig. 2A. Overall, the majority of the Norwegian food industry isolates clustered to one part of the tree, while the Norwegian clinical isolates and the isolates from natural environments (rural, urban, farm, and garden slugs) were more evenly distributed throughout the phylogeny. In several cases, cgMLST was unable to distinguish between individual isolates, for example in the case of a set of 21 isolates from two poultry processing factories; 15 from factory M6 and six from factory M2. These belonged to a cluster of 29 Norwegian isolates differing by maximum 3 cgMLST alleles, which in Fig. 2 is labelled as "poultry-associated cluster" and shown on blue shaded background. The other eight isolates in this cluster were: one additional isolate from M2, two from poultry processing plant M7, three clinical isolates, and two from natural environments.

Further differentiation using wgMLST was performed for 85 isolates shown on pink shaded background in Fig. 2A, comprising the majority of the Norwegian food industry isolates (Fig. 2B). In this analysis, all genomes could be distinguished except for three isolates from processing factory S1, two from M2, and two from Swedish fish products. The wgMLST analysis also confirmed the clustering of isolates from the poultry processing plant M2, M6 and M7.

A major difference between the two analyses concerned a cluster of six isolates indistinguishable by cgMLST, labelled with asterisks (*) in Fig. 2. The wgMLST analysis showed that one isolate in this cluster; MF2199 from meat processing factory M3, was differentiated by 58 to 78 wgMLST differences from the other five. These five isolates comprised four Norwegian clinical isolates, differentiated by 1 to 5 pairwise wgMLST allelic differences, and a natural environment isolate (MF7833) separated from the four clinical isolates by 11 to 15 wgMLST differences), as well as three isolates that were apparently not closely related to the remaining five genomes, as they were differentiated from the other five isolates by 58 to 79 wgMLST differences, and by between 56 and 94 wgMLST allelic differences from each other. A closer look at the 84 wgMLST loci differentiating these isolates (Supplementary Table S6) showed that 30 were present in the cgMLST scheme but the alleles present in the isolates were not identified upon synchronization with the allele mapping nomenclature at the BIGSdb-*Lm* database. If these alleles had been present in the database, MF2199 would have been differentiated from the other five isolates by 28 to 29 cgMLST alleles. For the remaining 54 wgMLST loci, 19 represented genes present only in MF2199 and MF7833, of which eight were likely to be prophage genes and nine were hypothetical proteins with no predicted function. The final 35 wgMLST loci represented genes found in all six isolates, of which 12 were likely prophage genes, four encoded internalins, while the remaining 19 had various other functions. In light of the commonly employed threshold of 7 to 10 cgMLST allelic differences for identification of clusters or "strains" from the same contamination source (Moura et al., 2017; Ruppitsch et al., 2015; Zamudio et al., 2020), this represents a case where the cgMLST analysis was not able to accurately differentiate between closely and not-so-closely related isolates.

3.2.2. *BEAST and ML phylogenies*

To further understand the phylogeny and temporal spread of clones, the genomes in the dataset with available raw Illumina sequencing data and known year of isolation (n = 280) were analysed by time-scaled Bayesian statistical analysis (BEAST) (Fig. 3 and Supplementary Fig. S1). Well-supported monophyletic clades from a maximum likelihood (ML) phylogeny were used to specify taxon sets for which the time of the most recent common ancestor (MRCA) was determined (Supplementary Table S7). The analysis indicated that CC7 originated around 500 years ago (date, 1537; 95 % CI, 1450 to 1619). The analysis furthermore separated the isolates into four distinct major monophyletic clades, referred to as Clades A, B, C, and D in Fig. 3. A comparison of the ML and Bayesian phylogenies showed that while the genomes in Clades B and D clustered similarly in the two trees, there were notable incongruences between the phylogenies within the two other clades, especially within Clade A. These differences were in most cases reflected by low bootstrap support values and low posterior probabilities for individual branch points (Supplementary Fig. S1).

Clade D was the most distant clade, composed of 13 isolates from Finland, France and USA. As it lacked Norwegian isolates, and for clarity, this clade was not shown in the cgMLST tree in Fig. 2A. The

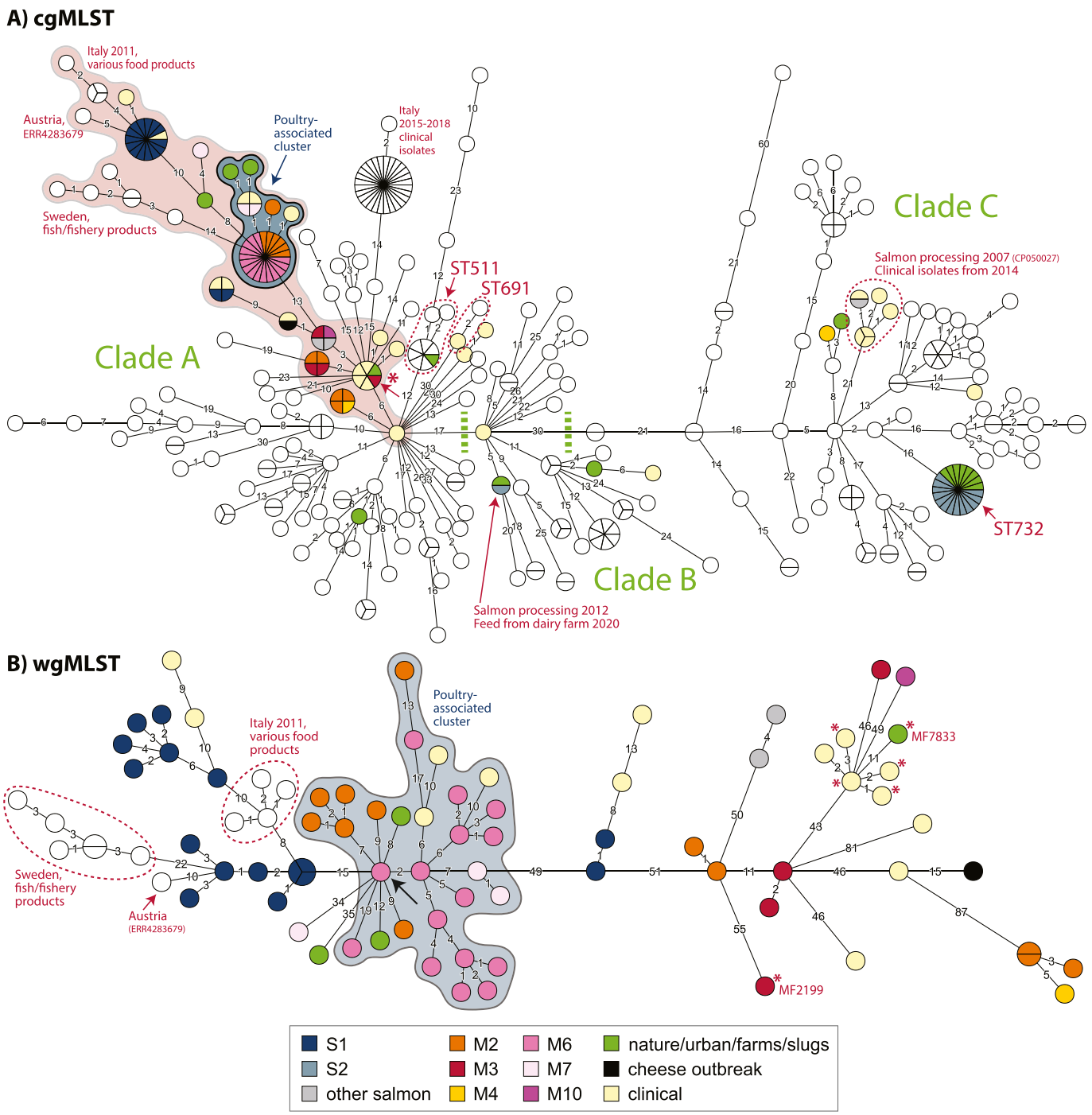

Fig. 2. Minimum spanning trees for CC7 genomes based on genomic MLST analyses. The area of each circle is proportional to the number of isolates represented, and the number of allelic differences between isolates is indicated on the edges connecting two nodes. A) cgMLST analysis of 115 Norwegian genomes (shown in colour) and reference genomes from other countries (in white). Different colours indicate the source of the isolates; S1-S2 are salmon industry isolates, M2-M10 are from meat industry. Norwegian isolates are ST7 unless otherwise indicated. Thirteen of the 255 selected publicly available reference genomes listed in Table 1 were excluded for clarity, as they belonged to a distinct clade without any Norwegian isolates (shown as Clade D in Fig. 3) which was separated from the remaining isolates by >340 cgMLST alleles. Clades A, B, and C as defined in Fig. 3 are indicated, and the boundaries between the three clades are shown as green dashed lines. B) wgMLST analysis for the 85 ST7 genomes shown on shaded background in A). The black arrow indicates MF2133 used as reference genome in subsequent SNP analysis. Isolates belonging to the cgMLST cluster labelled with an asterisk and an arrow in A) are also labelled with an asterisk in B).

MRCA of Clades A, B, and C combined was estimated to have existed in the middle of the 18th century, while that of Clades A and B combined was estimated to have existed early in the 19th century. Clade A, with an estimated time of the MRCA of around 1830, included the genomes from the wgMLST analysis presented in Fig. 2B, and was thus the clade containing the majority of the Norwegian food industry isolates. After comparing the SNP-based and cgMLST trees to assign clades for the 90 isolates not included in the SNP-based phylogeny, we found that 39 % of the Clade A isolates were Norwegian, 58 % originated in other European countries, and 3 % were from outside Europe (Supplementary Table S3). Of the 58 isolates in the subclade named A15 (Fig. 3), 81 % of isolates were Norwegian. In contrast, Clades B and C contained many isolates

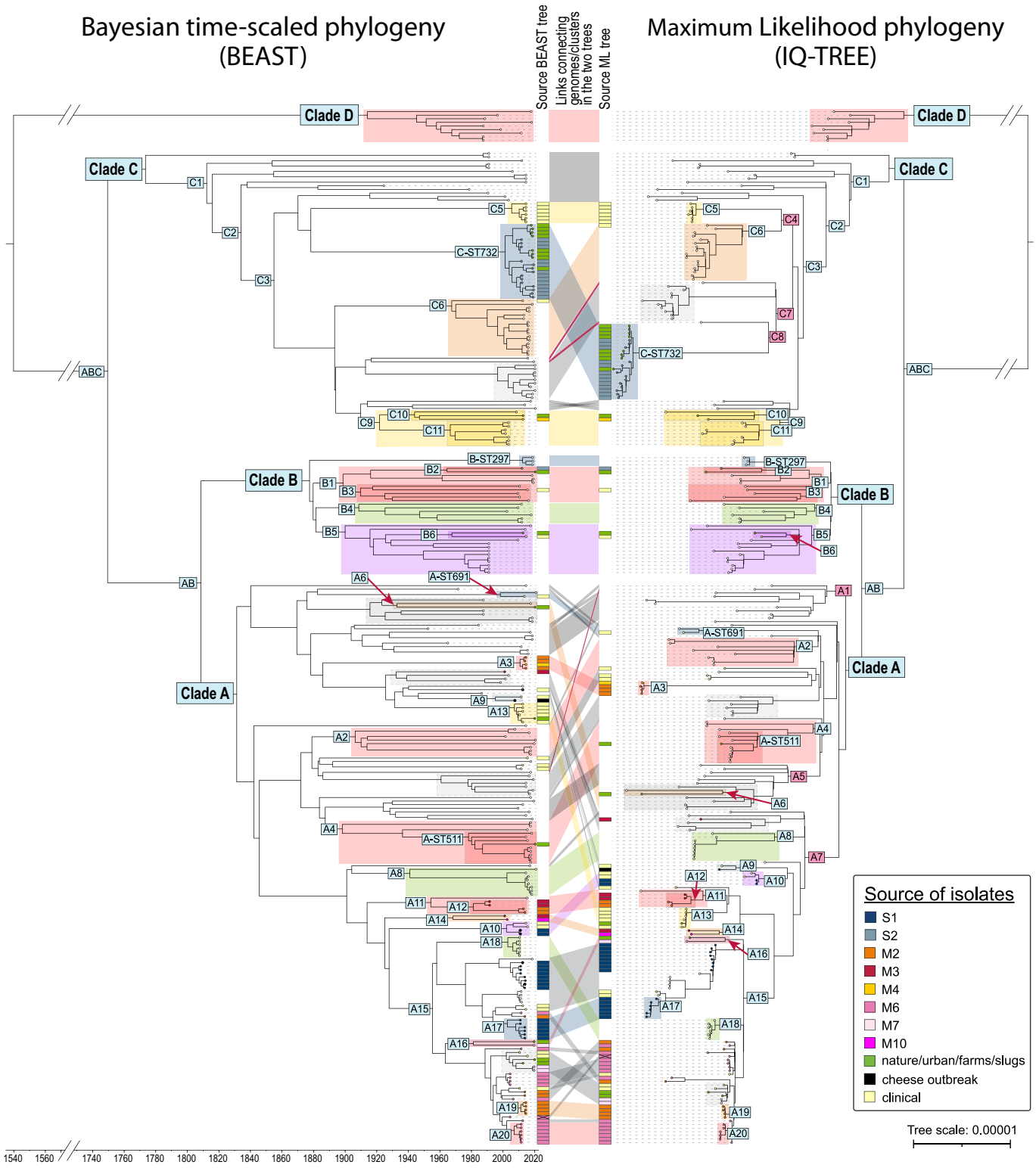

Fig. 3. Comparison of genome SNP trees of 280 CC7 *L. monocytogenes* obtained using time-scaled Bayesian and Maximum Likelihood phylogenetic approaches. The trees were plotted facing each other (on the left Bayesian and on the right ML) with corresponding genomes and clusters connected by coloured links (center). In the Bayesian tree, the tips of the branches correspond to the sampling date and the scale is in years. The ML tree was rooted at midpoint and nodes were rotated to optimize matching between corresponding strains in the Bayesian tree as closely as possible. The source of the Norwegian isolates is indicated with colours in columns between the trees and the links, and the SNP analysis reference genome MF2133 is marked with an X in the source columns. Clade labels on coloured background on the branches leading to selected nodes refer to the monophyletic clusters selected from the ML tree for estimation of MRCA during the Bayesian analysis – blue background indicates clades found in both trees (see Supplementary Table S7 for estimates of time of MCRA and 95 % HPD intervals) while pink background indicates clusters that did not represent monophyletic clades in the Bayesian analysis. The same tree including names of each isolate and posterior probability values or bootstrap supports for each split is shown in Supplementary Fig. S1.

originating from outside Europe (53 % and 47 % respectively, when all 370 isolates were counted). Overall, the phylogenetic analyses support the hypothesis that the genomic and geographic diversity differs within *L. monocytogenes* CC7, with a predominance of Norwegian isolates found within a specific subclade within Clade A, which again was populated predominantly by isolates of European origin.

3.3. Spread of CC7 clones in Norwegian food chains

A potential explanation for the high relative prevalence of CC7 among clinical isolates in Norway could be the successful establishment and subsequent persistence of CC7 isolates in the food industry. To examine this further, the patterns of contamination and transmission of CC7 clones within the Norwegian food industry was examined.

3.3.1. Poultry meat associated clusters

The main poultry-associated cluster mentioned above (Section 3.2.1) comprised 29 Norwegian isolates from poultry processing plants M2, M6, and M7 plus three clinical isolates (from 2010 and 2015) and two isolates from natural environments; a park in Oslo and a slug isolate, respectively. Of these, two isolates (MF6190 from M6 and MF6198 from M2) were not part of the monophyletic clade defined in the BEAST analysis, and separated from the remaining 27 isolates in the cluster by pairwise distances of 0–3 cgMLST alleles, 17–58 wgMLST alleles, and 20–39 SNPs. The pairwise distances separating the remaining 27 isolates, which formed a monophyletic clade in the BEAST phylogeny with an estimated time of the MRCA of around 1988, was 0–3 cgMLST alleles, 1–37 wgMLST alleles, and 0–27 SNP differences. The six isolates from M2 were from 2013 and 2014, the 16 isolates from M6 were from 2004, 2006, 2011, and 2012, while the two from M7 were from 2019. These 27 isolates comprise a pervasive strain, which we previously defined as a cluster of isolates collected over time with the same level of genetic similarity as persistent strains but isolated from different factories (Fagerlund et al., 2022b), which has likely circulated in the Norwegian poultry chain for over 30 years.

The above discussed isolates from poultry processing plant M2 were all found in floor-associated samples in the high-risk zone in which cooked meat was handled. Two additional persistent and pervasive CC7 clones were also found in factory M2, in a drain in the raw meat zone. The first clone was two isolates collected in 2012 and 2013, which were closely related to two isolates found in raw bovine meat at another meat processing plant in 1991; plant M3. The pairs of isolates from different factories were undistinguishable by cgMLST but separated by 11 or 12 wgMLST alleles and 13 to 15 SNPs. The second clone comprised three isolates from 2013 and 2014, which were closely related (indistinguishable by cgMLST, 4 to 7 wgMLST allelic differences and 1 to 2 SNPs) to a fourth isolate found in 2014 in yet another meat processing plant; plant M4. Plant M4 predominantly processed red meats but also had some poultry products. The estimated times of the MRCAs of these two clusters was 1981 and 2011, respectively (Supplementary Table S7). Thus, three different pervasive CC7 clones were identified in poultry processing plant M2, with close genetic links to isolates found in four other meat processing factories. All these factories were located >40 km apart from each other.

3.3.2. Salmon associated clusters

The phylogenies presented in Figs. 2 and 3 show that the main poultry-associated cluster was part of the larger subclade named A15, which also harbours two clusters of CC7 isolates from salmon processing plant S1, as well as a group of four isolates originating from various Italian foods in 2011, six isolates from Swedish salmon products collected in 2010, and an Austrian isolate from a fish product collected in 2018. The estimated time of the MRCA for subclade A15 was 1955 (Supplementary Table S7).

One of the two S1 clusters comprised eight isolates from 2011 to 2014, four of which were collected in different months during 2013 and 2014 from a conveyor belt associated with a fileting machine. The second S1 cluster (clade A17 in Fig. 3) comprised six isolates from 2011 and 2013. It is not clear whether these clusters persisted in factory S1 or whether they were persistent in an upstream part of the food chain and repeatedly reintroduced to S1 *via* gutted salmon. The two S1 clusters were indistinguishable by cgMLST analysis, but separated from each other by pairwise genetic distances of between 10 and 18 wgMLST alleles (median, 15) and between 12 and 22 SNPs (median, 16). The two clusters could thus potentially be considered a single strain in the context of *e.g.*, an outbreak investigation, as it has been suggested that a threshold of fewer than 7–10 cgMLST differences (Moura et al., 2017; Ruppitsch et al., 2015; Zamudio et al., 2020) and/or fewer than 20 SNPs indicates an epidemiological link between *L. monocytogenes* isolates (Allard et al., 2019; Wang et al., 2018).

The median pairwise genetic distances between the factory S1 isolates in clade A15 and the 27 isolates in the main poultry-associated cluster with isolates from factories M2, M6, and M7 was 11 cgMLST alleles (range 11–13), 24 wgMLST alleles (range 15–57), and 26 SNPs (range 18–57). Dependent on the selected threshold for inclusion of isolates, one could consider all isolates in subclade A15 to be members of a pervasive clone, circulating in both meat and salmon processing environments for over half a century.

Clade C harbored two subclades of Norwegian isolates associated with salmon. One was a large cluster of 21 ST732 isolates, found in salmon processing factory S2 during 2011–2019 (primarily in the stunning and gutting area) as well as during sampling performed during 2017–2018 from the outside urban environment (concrete, asphalt *etc.*) in the vicinity of factory S2. The estimated time of the MRCA for ST732 isolates was 2001 (Supplementary Table S7). The genomes were indistinguishable by cgMLST (Fig. 2A) but separated by pairwise distances of 0 to 24 wgMLST alleles (median, 9) and 0 to 17 SNPs (median, 8). The isolates from outside and inside the factory were intermingled in the phylogenetic trees, indicating that the same persistent population was present in both habitats, but it was not clear if the primary niche(s) for persistence was inside or outside of the factory.

In sum, these data show that CC7 clones had persisted in both meat and salmon processing environments in Norway.

3.4. Association with genes conferring stress tolerance and resistance

It has been suggested that *L. monocytogenes* strains recurrently isolated from food processing environments have characteristics enabling them to persist in stressful conditions (Ferreira et al., 2014; Palaiodimou et al., 2021). To examine whether the pervasion of CC7 isolates in Norwegian food processing chains correlated with an increase in the prevalence of plasmids and genetic determinants associated with stress survival and antimicrobial resistance, a genome-wide BLAST analysis was performed.

Of the analysed stress response or resistance associated genes, 72 core genes were found in all 370 analysed CC7 genomes (Supplementary Table S4). The core genes comprise genetic loci involved in the general stress response, such as the *lisRK* and the *sigB* operon, but also loci corresponding to more specific stress responses, such as stress survival islet 1 (SSI-1), conferring tolerance towards acidic, salt, bile and gastric stress (Ryan et al., 2010), the *argCJBDF* and *argGH* operons, both conferring acid stress tolerance, the *gbuABC* and the *opuC* operon, both conferring osmotic stress tolerance or the *dnaK* operon conferring heat stress tolerance, and the biofilm-associated gene *inlL* (Popowska et al., 2017). In accordance with the presence of SSI-1 in the hypervariable genetic hotspot between homologs of the genes *lmo0443* and *lmo0449*, the other two possible insertions in this locus, SSI-2 and a *LMOf2365_0481* homolog (Harter et al., 2017; Hein et al., 2011), were absent in all of the analysed CC7 isolates. The biofilm-associated gene *bapL*, the cadmium resistance cassette *cadA3C3*, the genes *emrE* and *qacH*, both encoding quaternary ammonium compound efflux pumps, and the arsenic resistance operon present on the *Listeria* Genomic Island

2 (LGI2) were not found in any of the analysed isolates.

Ten genetic elements associated with stress or resistance showed variable presence/absence patterns between genomes, and all these were found to be located on plasmids. They comprised the Tn*554*-associated arsenic resistance operon *arsCBADR* plus the nine plasmid-encoded genetic elements involved in stress response identified by Schmitz-Esser et al. (2021), including the cadmium and benzalkonium chloride resistance cassettes *cadA1C1*, *cadA2C2*, and *bcrABC* (Supplementary Table S3 and Fig. 4). All identified plasmids carried either RepA G1 or RepA G2 replication systems, with the exception of the previously identified plasmids carrying the novel variant RepA G12 (factory M2 isolate MF6196), the small pAUSMDU00000235 plasmid found in the clinical ST7 isolate ERR2522330, and the small rolling-circle replication (RCR) plasmid pLMST6 (pLmN12-0935) found in the two ST691 isolates ERR2522310 and ERR4284584.

The 21 ST732 isolates within Clade C, which were found both in salmon processing factory S2 and in the urban environment in the vicinity of the factory, all contained a RepA G2 plasmid containing several resistance determinants including *cadA1C1* and *arsCBADR*. None of the other Norwegian CC7 isolates from natural/urban environments contained plasmids. When the ST732 isolates were disregarded, there was no significant difference in plasmid contents between the isolates from Norwegian food, clinical, or natural/urban environments ($p > 0.49$, Fisher's exact test), and also no significant difference between isolates from meat and salmon processing environments ($p = 0.24$). Overall for CC7, there was no significant difference in plasmid content between the Norwegian isolates and the international references, as 27 % and 25 % of genomes in the two groups, respectively, carried plasmids. Excluding ST732, there was lower ($p = 0.03$) prevalence of plasmids in Norwegian isolates (11 %), compared to international references (25 %). No significant difference was seen between these two groups within Clade A, in which 10 % of the Norwegian genomes and 18 % of the reference

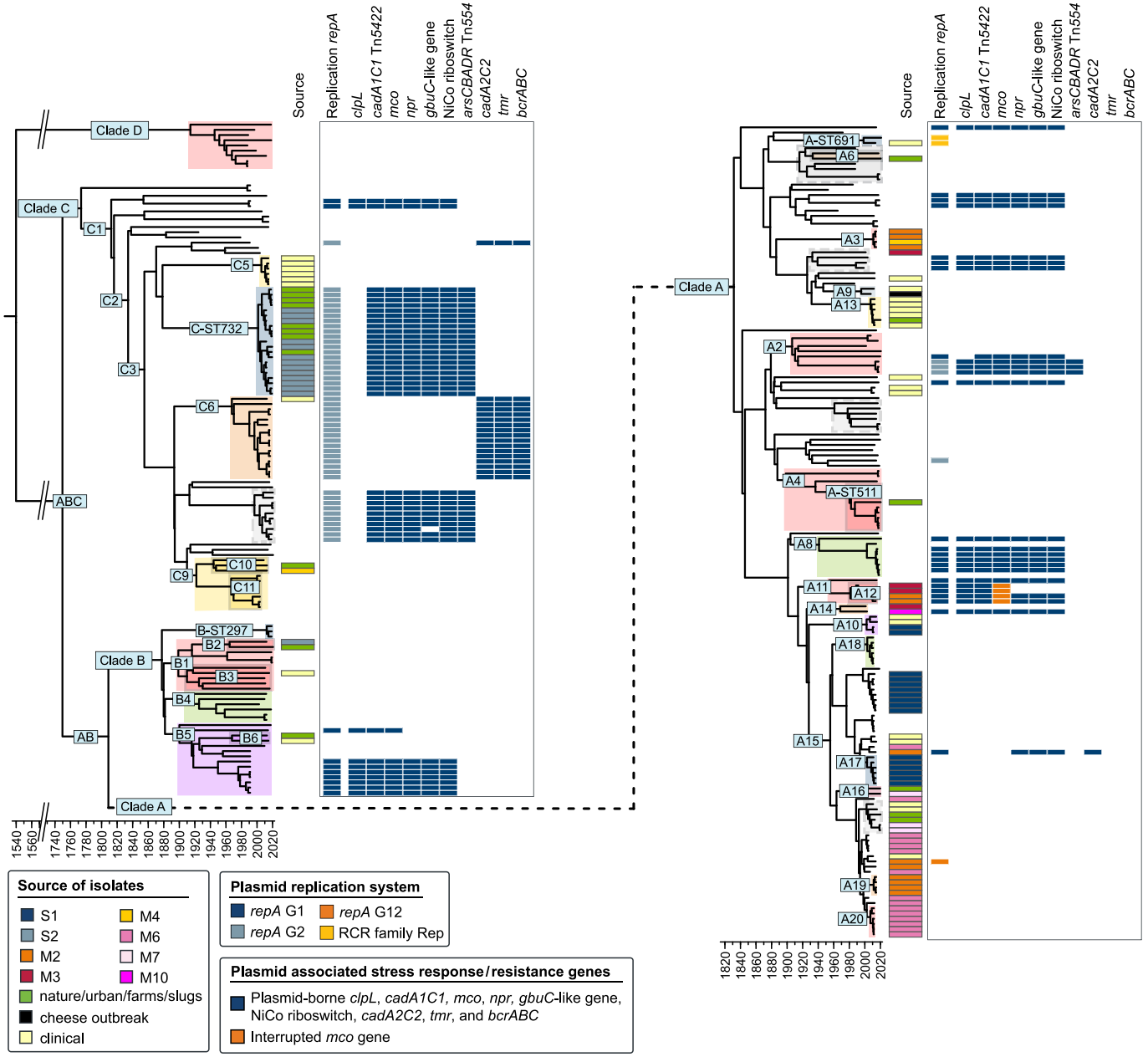

Fig. 4. Presence of accessory genetic determinants associated with stress survival or resistance in *L. monocytogenes* CC7. The phylogeny is the Bayesian time-scaled tree also shown in Fig. 3, but split in two sections with Clade A on the right. The same tree including names of each isolate and posterior probability values for each split is shown in Supplementary Fig. S1.

genomes carried plasmids ($p = 0.12$). In particular, only two of the 58 isolates in the pervasive clone comprising the A15 subclade contained plasmids (Supplementary Table S3 and Fig. 4). In total, the Norwegian CC7 isolates did not seem to have a higher prevalence of genetic elements conferring resistance to stressful conditions encountered in the food industry than CC7 isolates from other countries.

3.5. Virulence characteristics of the CC7 isolates

Another hypothesis is that the high relative prevalence of CC7 among Norwegian clinical isolates could be due to a high virulence potential of the Norwegian CC7 isolates. The virulence potential was assessed by *in vitro* virulence assays of selected Norwegian isolates and by analyses of known genetic virulence determinants in the genome sequences of Norwegian as well as international reference isolates.

3.5.1. Analysis of virulence gene content in CC7 genomes

The results from the BLAST search for virulence-associated genes and *inlA* premature stop codon (PMSC) variants showed that the genetic virulence attributes were highly conserved in all 370 analysed CC7 genomes, independent of *e.g.*, the source and country of isolation. Out of the 99 analysed virulence genes, 83 were present while 16 were absent in all strains (Supplementary Table S5). Furthermore, internalins A and B, which are the most prominent virulence factors mediating receptor-dependent host cell invasion, were both present as full-length variants in all genomes. Listeria Pathogenicity Island (LIPI) 1 was present in all isolates, while LIPI-3 and LIPI-4 were absent. All strains harbored *inlC2* and *inlD* located between *inlE* and *inlG* and therefore lacked *inlH*. *InlP1* and *inlP3*, which were described to contribute to the invasion of intestinal epithelial Caco2 cells *in vitro* (Harter et al., 2019) and *vip* (virulence protein) were consistently absent in the CC7 *L. monocytogenes* strains analysed in this study.

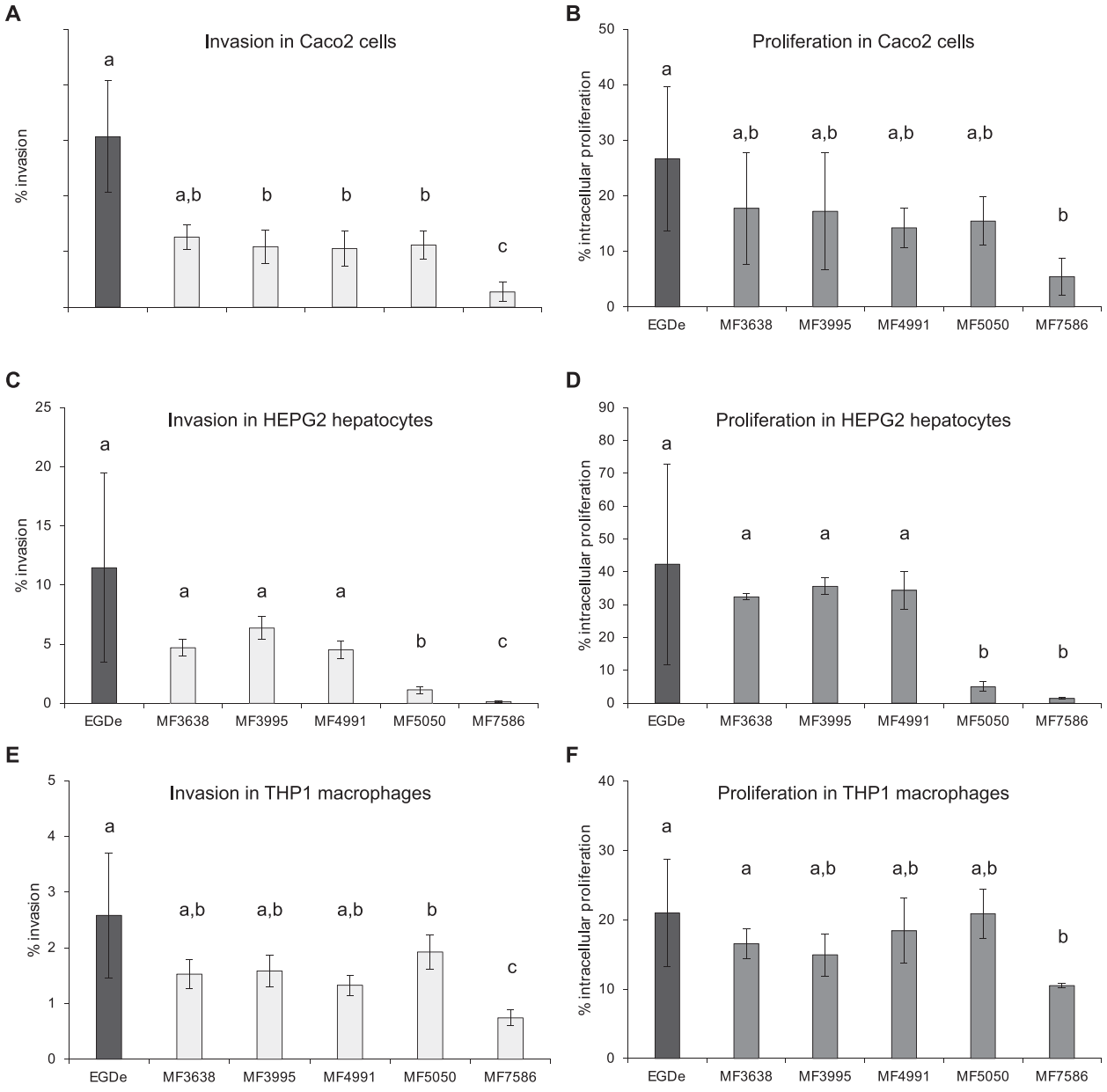

Fig. 5. Virulence potential of five *L. monocytogenes* CC7 isolates and the reference strain EGDe in human cell lines. (A) Invasion efficiency (%) and (B) intracellular proliferation (%) in human intestinal epithelial Caco2 cells. (C) Invasion efficiency (%) and (D) intracellular proliferation (%) in human HEPG2 hepatocytes. (E) Invasion efficiency (%) and (F) intracellular proliferation (%) in human THP1 macrophages. Data are shown as mean ± SD of three biological replicates determined in duplicates. Means that do not share a letter are significantly different ($p < 0.05$).

The only difference found between the analysed CC7 isolates were variants with different number of GW repeat domains in the cell wall anchored autolysin protein Ami. While most genomes had an *ami* variant containing eight GW domains, the following four carried a variant encoding four domains: MF5050 (from subclade A17 and salmon factory S1), ERR2522310 (clinical isolate from Norway), ERR4284584 (food isolate from France), and ID52872 (food isolate from UK).

3.5.2. *In vitro virulence assays*

As prediction of virulence potential based on virulence gene profiles has limitations and should be combined with phenotypic data (Wagner et al., 2022), the virulence potential of five selected isolates from Norwegian food processing environments was characterized using *in vitro* virulence assays. Clinical CC7 isolates from Norway were not available and could not be tested.

The isolates were selected among the Norwegian food or food processing industry isolates and the selection was based on phylogenetic diversity and inclusion of isolates from different factories and food sources, including an isolate from cheese obtained in association with a soft cheese outbreak in 2007 (Johnsen et al., 2010) and one ST732 isolate from Clade C (Table 2). In addition, different gene profiles were included where possible, however, the virulence gene profile of the five selected isolates was identical, with the exception of the MF5050 isolate harboring the shorter *ami* variant. As a reference, the EGDe lab strain was included in all assays. The virulence gene profile of EGDe differed from the CC7 isolates by the presence of *inlH* instead of *inlC2* and *inlD* and the presence of *vip*. In addition, EGDe carried a variant of the *iap* gene, encoding an invasion-associated secreted endopeptidase, with 18 variable TN (threonine and asparagine) repeats compared to a variant with 14 repeats found in the CC7 genomes (Wagner et al., 2022).

The invasion efficiency and intracellular proliferation was determined in human intestinal epithelial Caco2 cells (Fig. 5AB), HEPG2 hepatocytes (Fig. 5CD) and THP1 macrophages (Fig. 5EF) using gentamicin protection assays. The results showed that invasion and intracellular proliferation of the CC7 isolates, compared to the reference isolate EGDe, were generally lower in Caco2 cells but similar in HEPG2 and THP1 cells. Relatively homogeneous values for both the invasion and intracellular proliferation were observed among the CC7 isolates in all three assessed cell types except for the MF7586 isolate (from the main poultry-associated cluster), which consistently showed lower invasion and intracellular proliferation, and the MF5050 isolate (with the shorter Ami variant), which showed lower invasion and intracellular proliferation in HEPG2 hepatocytes (Fig. 5).

Conclusively, although the assessed isolates did not show exceptionally high *in vitro* invasion and intracellular proliferation, they were able to invade all three different host cell types and proliferate within them to different extents, depending on the isolate.

4. Discussion

The relative prevalence of *L. monocytogenes* CC7 among clinical isolates was higher in Norway and in the neighboring countries Sweden and Finland compared to in most other European countries in the period 2010 to 2015. In these countries, *L. monocytogenes* lineage II CCs (including CC7) dominated, comprising >75 % of the clinical isolates, while lineage I isolates dominated in the majority of the other European countries. This is in line with Orsi et al. (2011), who previously pointed out that lineage II clinical isolates dominated in Northern Europe. In the present study, for some of the countries the CC prevalence was rather uncertain since WGS data was only available for a low fraction of the total number of isolates, but it is still clear that Norway had a high relative prevalence of CC7 among clinical isolates compared to other European countries.

In a study of *L. monocytogenes* isolated from 2005 to 2013 in France (Maury et al., 2016), CC7 was considered intermediate regarding its distribution between food isolates and clinical isolates. The lineage I CCs CC1, CC2, CC4 and CC6 were frequently associated with clinical cases and not often associated with food, while the lineage II CCs CC9 and CC121 were frequently isolated from food and rarely associated with clinical cases. Also, another European study showed a higher prevalence of lineage I among clinical isolates and of lineage II among food isolates (Painset et al., 2019). CC7 was the only CC of lineage II that was more frequently isolated from clinical cases than from food in both the above-mentioned French and European studies (Maury et al., 2016; Painset et al., 2019). Although lineage I CCs are associated with many of the largest listeriosis outbreaks (Büla et al., 1995; Cantinelli et al., 2013; Halbedel et al., 2020; Maury et al., 2016; Thomas et al., 2020) several have also involved CC7. The large cantaloupe outbreak in USA involved several CCs, including CC7 (Lomonaco et al., 2013). CC7 was also involved in two outbreaks associated with RTE pork in Italy and USA (CDC, 2011; Chiaverini et al., 2021), an outbreak from soft cheese in Norway (Johnsen et al., 2010), an outbreak associated with whipped cream in Canada (Pagotto et al., 2006), and recently in an outbreak with fish cakes in Denmark (Anonymous, 2022). Some subtypes of *L. monocytogenes* seem to be more commonly associated with certain food types or food industries than others, like CC9 with the meat industry and CC1 with dairy products (Fagerlund et al., 2022b; Maury et al., 2019). In contrast, outbreaks with CC7 have been reported within all the top four risk foods for listeriosis: RTE meat, soft cheese, fresh produce or fruit, and lightly processed fish products. In summary, it can be concluded that CC7 is more frequently associated with listeriosis than many other lineage II CCs, but that the relative clinical prevalence of CC7 varies between countries and seems to be especially high in Norway and some other Nordic countries.

L. monocytogenes belonging to sublineage 7 (SL7), corresponding to CC7, has been shown to have a low geographical transition rate (Moura et al., 2017), which may have led to a more uneven international distribution of CC7 compared to other subtypes. Since listeriosis is foodborne, the prevalence of CC7 among clinical cases of listeriosis in a region could potentially be influenced by its relative prevalence in the environment and food chains. In France, Maury et al. (2016) reported low relative prevalence of CC7 both among food isolates (0.9 %) and clinical isolates (2.9 %), while in a study from 10 European countries (Papić et al., 2019), CC7 was not ranked among the 11 most common CCs in the natural environment. In contrast, CC7 is relatively common in Norway both in the environment, where it ranked fourth among CCs, and the food chain, where it ranked third (Fagerlund et al., 2022a; Fagerlund et al., 2022b). Potentially, a higher risk of intake of CC7 *via* food may be responsible for the high relative prevalence of CC7 among clinical isolates in Norway.

As only a minority of the Norwegian clinical isolates were closely related to identified persistent or pervasive clones, it is not possible to directly link the prevalence of persistent clones with a higher clinical frequency. Furthermore, it is difficult to conclude whether persistence of CC7 is more common in Norway than in other countries, as relatively few studies on persistence (and almost none before 2010) includes information about CCs or STs. However, in addition to the current study, which showed that CC7 persisted in the production environment in several Norwegian poultry and salmon processing plants, persistence of CC7 has been reported in a Danish (Knudsen et al., 2017) and a French (Palma et al., 2020) fish processing plant, while Chiaverini et al. (2021) reported isolation of the same clone of CC7 over time in pork ready to eat products from Italy, likely indicating persistence of the clone. The current study did not identify any significant difference in prevalence of genetic determinants of stress response or persistence between the Norwegian and international isolates, thus no genetic basis for a higher prevalence of persistence of CC7 in Norway was identified. Overall, persistence of CC7 in the food industry – both in Norway and other countries – is reported in several cases but likely not as common as for CC121 and CC9, where persistence has been frequently reported (Cherifi et al., 2018; Fagerlund et al., 2020; Luo et al., 2017; Melero et al., 2019; Schmitz-Esser et al., 2015).

A high prevalence of a certain CC among clinical isolates could be related to its relative virulence potential. CC7 has been proposed to show intermediate virulence characteristics, with lower potential than the hypervirulent lineage I isolates belonging to CC1, CC4, and CC6, and higher potential than the hypovirulent lineage II CC9 and CC121 subtypes, in a mouse model (Maury et al., 2016). It was also of intermediate virulence in previous work comparing virulence between CCs in a Caco2 *in vitro* virulence assay (Wagner et al., 2022). CC7 was also ranked in the intermediate group (group 2 out of 3), in a virulence ranking proposed by the FAO and WHO (2022), in which intermediate strains were characterized by the lack of LIPI-3 or LIPI-4, and the presence of complete and functional LIPI-1 and full length internalin A genes. These criteria were fulfilled by all of the analysed CC7 isolates. The virulence gene content of Norwegian CC7 isolates did not vary from isolates from other countries, making it unlikely that observed differences in the relative prevalence of clinical CC7 between countries was due to variations in the content of specific known virulence factors. For the *in vitro* virulence studies performed in the present work, we did not have access to and could therefore not use clinical isolates, but instead used closely related CC7 isolates from the food industry. According to our studies, the CC7 isolates from the Norwegian food industry were able to invade all three types of human cell lines tested, with isolate-dependent differences. Three of the four food industry isolates had similar invasion capacities as an isolate from a small cheese producer associated with an outbreak in 2007 (Johnsen et al., 2010), while isolate MF5050, harboring a shorter *ami* variant, showed lower invasion and intracellular proliferation in hepatocytic HEPG2 cells, but not in the intestinal and macrophage cell lines. Our results imply that the virulence potential of CC7 represented by the phylogenetically diverse isolates selected in this study is comparable, unless an unambiguous mutation occurs in a gene that plays a substantial role in invading host cells. Although it cannot be ruled out that differences in virulence potential between CC7 isolates exist, the findings in the current study suggests that factors other than differences in virulence among CC7 isolates may explain the large differences between countries in the relative prevalence of CC7 among clinical isolates.

Many of the Norwegian isolates were closely related. The present study shows the superiority of wgMLST and SNP analyses over cgMLST to differentiate between closely related strains, and that cgMLST was not sufficient for distinguishing between CC7 isolates from different food processing plants. Several factors contributed to the large difference in number of variable genes obtained using cgMLST analysis relative to wgMLST analysis observed in a subset of the clusters. The first was related to the dependence of standardized cgMLST analyses on curated databases. The consistent allele nomenclature available with fixed and publicly available MLST schemes is an advantage with respect to portability of results between studies or laboratories, but results in lower phylogenetic resolution than expected in cases where the query genome contains allelic variants not present in (or which are not upon query automatically submitted to) the employed database. In the case of the presently employed BIGSdb-*Lm* cgMLST database, new alleles are included upon deposition of their corresponding genome sequences. The full potential of a cgMLST analysis generated through query of this database can thus only be ensured by prior deposition of all genomes included in the analysis in the database. Another factor contributing to the difference in resolution obtained with the different WGS-based subtyping tools was the presence of prophage genes. In general, the choice of subtyping tool and whether to include horizontally transferred sequences such as prophages in a phylogenetic analysis will depend on the degree of relatedness of the isolates in the dataset as well as on the goal of the analysis. Such sequences are usually confounding factors when the desired result is to determine the best estimate for a phylogeny representing the most likely ancestor-descendent evolutionary relationships. However, when comparing very closely related genomes, such as during outbreak investigations, variations present in mobile genetic elements may constitute valuable information. Indeed, the employed wgMLST scheme (Applied Maths, 2015) specifically includes accessory loci such as prophage genes in order to increase the discriminatory power of the analysis and allow detection of subtype- or outbreak-specific markers.

Overall, CC7 isolates from different Norwegian food processing plants were often more closely related to each other than to isolates from other countries. This indicates that the majority of the Norwegian isolates have a common origin. In the main cluster with isolates from three poultry processing plants (M2, M6, M7), the differences within the cluster were so small (<20 wgMLST alleles) that it indicates that this pervasive strain had a relatively recent common source, estimated in the BEAST analysis to have existed around year 1988. The potential common source of these isolates is not known. We have previously reported the transfer of an ST8 strain from M6 to M2 after transfer of used equipment (conveyor) (Fagerlund et al., 2016). We are aware that also other equipment was transferred between these processing plants, but it is not known if this led to transfer of *L. monocytogenes* CC7 between the plants. Common for the meat processing plants M2, M6 and M7 are that they process only poultry, and not other types of meat. We have recently speculated whether crossing distribution patterns of raw red meat is the reason for overlapping ST9 strains in different plants (Fagerlund et al., 2020). There are fewer slaughterhouses than meat processing plants in Norway, meaning that different processing plants can receive raw meat (fresh or frozen) from the same slaughterhouse. Although CC7 was isolated from several poultry processing plants, and to a lesser extent from red meat processing plants, where CC9 dominated (Fagerlund et al., 2020; Fagerlund et al., 2022b), it cannot be concluded that CC7 was more frequently associated with poultry than meat processing, as the number of sampled processing plants in each category was low. Also, we have not found support for this hypothesis in the literature, and to the contrary, the two meat-associated CC7 outbreaks we are aware of were both associated with pork (CDC, 2011; Chiaverini et al., 2021). We are not aware that there are systematic differences in hygienic routines between the two types of processing (poultry *vs* red meat) in Norway.

5. Conclusion

Although the relative prevalence of *L. monocytogenes* CC7 among clinical isolates was higher in Norway compared to in most European countries, the reason behind this is still not clear. A pervasive clone was identified in the Norwegian salmon and meat industry, however, most clinical isolates were not related to this clone. Also, no major differences were seen in virulence and stress response determinants when comparing Norwegian isolates and those from other countries. CC7 appears to be a subtype with potential both for persistence in the food industry and potential to infect humans. CC7 contains the determinants for virulence and stress responses shared by most CC types of *L. monocytogenes*, without the additional virulence genes shared by hypervirulent strains or truncated internalin genes shared by hypovirulent strains. Further work is still needed to explain differences in prevalence between countries, both for CC7 and for CCs in general.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijfoodmicro.2023.110482.

Declaration of competing interest

None.

Data availability

All sequence data have been deposited to public available databases

Acknowledgements

We thank Merete Rusås Jensen at Nofima for excellent technical assistance and the Institut Pasteur teams for the curation and

maintenance of BIGSdb-Pasteur databases at https://bigsdb.pasteur.fr/. This study was supported by The Research Council of Norway through grant no. 294910.

References

Allard, M.W., Strain, E., Rand, H., Melka, D., Correll, W.A., Hintz, L., Stevens, E., Timme, R., Lomonaco, S., Chen, Y., Musser, S.M., Brown, E.W., 2019. Whole genome sequencing uses for foodborne contamination and compliance: discovery of an emerging contamination event in an ice cream facility using whole genome sequencing. Infect. Genet. Evol. 73, 214–220.

Anonymous, 2022. Denmark–Outbreak of Invasive Listeria Infection Sequence Type 7 in Denmark Caused by Fish Meatballs. FoodWorld. https://kswfoodmicro.com/2022/12/19/denmark-outbreak-of-invasive-listeria-infection-sequence-type-7-in-denmark-caused-by-fish-meatballs/.

Applied Maths, 2015. *Listeria monocytogenes* whole genome sequence typing. https://www.applied-maths.com/news/listeria-monocytogenes-whole-genome-sequence-typing.

Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., Wishart, D.S., 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44 (W1). Jul 8. W16-21.

Aronesty, E., 2011. ea-utils: "Command-line Tools for Processing Biological Sequencing Data". https://github.com/ExpressionAnalysis/ea-utils.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Son, P., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.

Bucur, F.I., Grigore-Gurgu, L., Crauwels, P., Riedel, C.U., Nicolau, A.I., 2018. Resistance of *Listeria monocytogenes* to stress conditions encountered in food and food processing environments. Front. Microbiol. 9, 2700.

Büla, C.J., Bille, J., Glauser, M.P., 1995. An epidemic of food-borne listeriosis in western Switzerland: description of 57 cases involving adults. Clin. Infect. Dis. 20, 66–72.

Bushnell, B., 2014. BBMap: a fast, accurate, splice-aware aligner. In: Conference: 9th Annual Genomics of Energy & Environment Meeting,, March 17–20, 2014, Walnut Creek, CA.

Cantinelli, T., Chenal-Francisque, V., Diancourt, L., Frezal, L., Leclercq, A., Wirth, T., Lecuit, M., Brisse, S., 2013. "Epidemic clones" of *Listeria monocytogenes* are widespread and ancient clonal groups. J. Clin. Microbiol. 51, 3770–3779.

CDC, 2011. Outbreak of invasive listeriosis associated with the consumption of hog head cheese–Louisiana, 2010. MMWR Morb. Mortal. Wkly Rep. 60, 401–405.

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

Chiaverini, A., Guidi, F., Torresi, M., Acciari, V.A., Centorotola, G., Cornacchia, A., Centorame, P., Marfoglia, C., Blasi, G., Di Domenico, M., Migliorati, G., Roussel, S., Pomilio, F., Sevellec, Y., 2021. Phylogenetic analysis and genome-wide association study applied to an Italian *Listeria mono- cytogenes* outbreak. Front. Microbiol. 12, 750065.

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

Clausen, P.T.L.C., Aarestrup, F.M., Lund, O., 2018. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinformatics 19, 307.

Davis, S., Pettengill, J.B., Luo, Y., Payne, J., Shpuntoff, A.H.R., Strain, E., 2015. CFSAN SNP pipeline: an automated method for constructing SNP matrices from next-generation sequence data. PeerJ Comput. Sci. 1, e20.

De Coster, W., 2020. NanoComp. https://github.com/wdecoster/nanocomp.

De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., Van Broeckhoven, C., 2018. NanoPack: visualizing and processing long-read sequencing data. Bioinformatics 34, 2666–2669.

Desai, A.N., Anyoha, A., Madoff, L.C., Lassmann, B., 2019. Changing epidemiology of *Listeria monocytogenes* outbreaks, sporadic cases, and recalls globally: a review of ProMED reports from 1996 to 2018. Int. J. Infect. Dis. 84, 48–53.

Elhanafi, D., Dutta, V., Kathariou, S., 2010. Genetic characterization of plasmid-associated benzalkonium chloride resistance determinants in a *Listeria monocytogenes* strain from the 1998-1999 outbreak. Appl. Environ. Microbiol. 76, 8231–8238.

European Food Safety Authority (EFSA), Costa, G., Di Piazza, G., Koevoets, P., Iacono, G., Liebana, E., Pasinato, L., Rizzi, V., Rossi, M., 2022. Guidelines for reporting whole genome sequencing-based typing data through the EFSA One Health WGS system. EFSA Supporting Publications 19, 7413E.

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

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

Fagerlund, A., Idland, L., Heir, E., Møretrø, T., Aspholm, M., Lindback, T., Langsrud, S., 2022a. Whole-genome sequencing analysis of *Listeria monocytogenes* from rural, urban, and farm environments in Norway: genetic diversity, persistence, and relation to clinical and food isolates. Appl. Environ. Microbiol. 88, e02136-21.

Fagerlund, A., Wagner, E., Møretrø, T., Heir, E., Moen, B., Rychli, K., Langsrud, S., 2022b. Pervasive *Listeria monocytogenes* is common in the Norwegian food system and is associated with increased prevalence of stress survival and resistance determinants. Appl. Environ. Microbiol. 88, e0086122.

FAO, WHO, 2022. *Listeria monocytogenes* in Ready-to-Eat (RTE) Foods: Attribution, Characterization and Monitoring – Meeting Report. Rome. Microbiological Risk Assessment series, Rome, Italy, p. 202.

Félix, B., Sevellec, Y., Palma, F., Douarre, P.E., Felten, A., Radomski, N., Mallet, L., Blanchard, Y., Leroux, A., Soumet, C., Bridier, A., Piveteau, P., Ascensio, E., Hébraud, M., Karpíšková, R., Gelbíčová, T., Torresi, M., Pomilio, F., Cammà, C., Di Pasquale, A., Skjerdal, T., Pietzka, A., Ruppitsch, W., Canelhas, M.R., Papić, B., Hurtado, A., Wullings, B., Bulawova, H., Castro, H., Lindström, M., Korkeala, H., Šteingolde, Ž., Kramarenko, T., Cabanova, L., Szymczak, B., Gareis, M., Oswaldi, V., Marti, E., Seyfarth, A.-M., Leblanc, J.-C., Guillier, L., Roussel, S., 2022. A European-wide dataset to uncover adaptive traits of *Listeria monocytogenes* to diverse ecological niches. Sci. Data 9, 190.

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

Halbedel, S., Wilking, H., Holzer, A., Kleta, S., Fischer, M.A., Lüth, S., et al., 2020. Large Nationwide outbreak of invasive Listeriosis associated with blood sausage, Germany, 2018–2019. Emerg. Infect. Dis. 26, 1456–1464.

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

Harter, E., Lassnig, C., Wagner, E.M., Zaiser, A., Wagner, M., Rychli, K., 2019. The novel Internalins InlP1 and InlP4 and the Internalin-like protein InlP3 enhance the pathogenicity of *Listeria monocytogenes*. Front. Microbiol. 10, 1644.

Hein, I., Klinger, S., Dooms, M., Flekna, G., Stessl, B., Leclercq, A., Hill, C., Allerberger, F., Wagner, M., 2011. Stress survival islet 1 (SSI-1) survey in *Listeria monocytogenes* reveals an insert common to *Listeria innocua* in sequence type 121 *L. monocytogenes* strains. Appl. Environ. Microbiol. 77, 2169–2173.

Johnsen, B.O., Lingaas, E., Torfoss, D., Strøm, E.H., Nordøy, I., 2010. A large outbreak of *Listeria monocytogenes* infection with short incubation period in a tertiary care hospital. J. Inf. Secur. 61, 465–470.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin, L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14, 587–589.

Knudsen, G.M., Nielsen, J.B., Marvig, R.L., Ng, Y., Worning, P., Westh, H., Gram, L., 2017. Genome-wide-analyses of *Listeria monocytogenes* from food-processing plants reveal clonal diversity and date the emergence of persisting sequence types. Environ. Microbiol. Rep. 9, 428–440.

Larsen, M.V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R.L., Jelsbak, L., Sicheritz-Ponten, T., Ussery, D.W., Aarestrup, F.M., Lund, O., 2012. Multilocus sequence typing of Total-genome-sequenced bacteria. J. Clin. Microbiol. 50, 1355–1361.

Letunic, I., Bork, P., 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 49, W293–W296.

Lomonaco, S., Verghese, B., Gerner-Smidt, P., Tarr, C., Gladney, L., Joseph, L., Katz, L., Turnsek, M., Frace, M., Chen, Y., Brown, E., Meinersmann, R., Berrang, M., Knabel, S., 2013. Novel epidemic clones of *Listeria monocytogenes*, United States, 2011. Emerg. Infect. Dis. 19, 147–150.

Løvdal, T., Brandal, L.T., Sundaram, A.Y.M., Naseer, U., Roth, B., Lunestad, B.T., 2021. Small-scale comparative genomic analysis of *Listeria monocytogenes* isolated from environments of salmon processing plants and human cases in Norway. Hygiene 1, 43–55.

Luo, L.J., Zhang, Z.D., Wang, H., Wang, P.F., Lan, R.T., Deng, J.P., Miao, Y.M., Wang, Y., Wang, Y., Xu, J.G., Zhang, L., Sun, S.S., Liu, X., Zhou, Y., Chen, X., Li, Q., Ye, C.Y., 2017. A 12-month longitudinal study of *Listeria monocytogenes* contamination and persistence in pork retail markets in China. Food Control 76, 66–73.

Maury, M.M., Tsai, Y.H., Charlier, C., Touchon, M., Chenal-Francisque, V., Leclercq, A., Criscuolo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E.P.C., Brisse, S., Lecuit, M., 2016. Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. Nat. Genet. 48, 308–313.

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

Melero, B., Manso, B., Stessl, B., Hernandez, M., Wagner, M., Rovira, J., Rodriguez-Lazaro, D., 2019. Distribution and persistence of Listeria monocytogenes in a heavily contaminated poultry processing facility. J. Food Prot. 82, 1524–1531.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R., 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. 37, 1530–1534.

Moura, A., Criscuolo, A., Pouseele, H., Maury, M.M., Leclercq, A., Tarr, C., Björkman, J. T., Dallman, T., Reimer, A., Enouf, V., Larsonneur, E., Carleton, H., Bracq-Dieye, H., Katz, L.S., Jones, L., Touchon, M., Tourdjman, M., Walker, M., Stroika, S., Cantinelli, T., Chenal-Francisque, V., Kucerova, Z., Rocha, E.P.C., Nadon, C., Grant, K., Nielsen, E.M., Pot, B., Gerner-Smidt, P., Lecuit, M., Brisse, S., 2017. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. Nat. Microbiol. 2, 16185.

Müller, A., Rychli, K., Zaiser, A., Wieser, C., Wagner, M., Schmitz-Esser, S., 2014. The *Listeria monocytogenes* transposon Tn6188 provides increased tolerance to various quaternary ammonium compounds and ethidium bromide. FEMS Microbiol. Lett. 361, 166–173.

Orsi, R.H., Bakker, H.C.D., Wiedmann, M., 2011. *Listeria monocytogenes* lineages: genomics, evolution, ecology, and phenotypic characteristics. Int. J. Med. Microbiol. 301, 79–96.

Pagotto, F., Ng, L.-K., Clark, C., Farber, J., Network, C.P.H.L., 2006. Canadian Listeriosis Reference Service, 3, pp. 132–137.

Painset, A., Björkman, J.T., Kiil, K., Guillier, L., Mariet, J.-F., Félix, B., Amar, C., Rotariu, O., Roussel, S., Perez-Reche, F., Brisse, S., Moura, A., Lecuit, M., Forbes, K., Strachan, N., Grant, K., Møller-Nielsen, E., Dallman, T.J., 2019. LiSEQ – whole-genome sequencing of a cross-sectional survey of *Listeria monocytogenes* in ready-to-eat foods and human clinical cases in Europe. Microb. Genom. 5, e000257.

Palaiodimou, L., Fanning, S., Fox, E.M., 2021. Genomic insights into persistence of *Listeria* species in the food processing environment. J. Appl. Microbiol. 131, 2082–2094.

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

Papić, B., Pate, M., Félix, B., Kušar, D., 2019. Genetic diversity of *Listeria monocytogenes* strains in ruminant abortion and rhombencephalitis cases in comparison with the natural environment. BMC Microbiol. 19, 299.

Pietzka, A., Allerberger, F., Murer, A., Lennkh, A., Stöger, A., Cabal Rosel, A., Huhulescu, S., Maritschnik, S., Springer, B., Lepuschitz, S., Ruppitsch, W., Schmid, D., 2019. Whole genome sequencing based surveillance of *L. monocytogenes* for early detection and investigations of listeriosis outbreaks. Front. Public Health 7, 139.

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

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

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

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

Schmitz-Esser, S., Müller, A., Stessl, B., Wagner, M., 2015. Genomes of sequence type 121 *Listeria monocytogenes* strains harbor highly conserved plasmids and prophages. Front. Microbiol. 6, 380.

Schmitz-Esser, S., Anast, J.M., Cortes, B.W., 2021. A large-scale sequencing-based survey of plasmids in *Listeria monocytogenes* reveals global dissemination of plasmids. Front. Microbiol. 12, 653155.

Sovic, I., Sikic, M., Wilm, A., Fenlon, S.N., Chen, S., Nagarajan, N., 2016. Fast and sensitive mapping of nanopore sequencing reads with GraphMap. Nat. Commun. 7.

Stevens, E.L., Carleton, H.A., Beal, J., Tillman, G.E., Lindsey, R.L., Lauer, A.C., Pightling, A., Jarvis, K.G., Ottesen, A., Ramachandran, P., Hintz, L., Katz, L.S., Folster, J.P., Whichard, J.M., Trees, E., Timme, R.E., Mcdermott, P., Wolpert, B., Bazaco, M., Zhao, S., Lindley, S., Bruce, B.B., Griffin, P.M., Brown, E., Allard, M., Tallent, S., Irvin, K., Hoffmann, M., Wise, M., Tauxe, R., Gerner-Smidt, P., Simmons, M., Kissler, B., Defibaugh-Chavez, S., Klimke, W., Agarwala, R., Lindsay, J., Cook, K., Austerman, S.R., Goldman, D., Mcgarry, S., Hale, K.R., Dessai, U., Musser, S.M., Braden, C., 2022. Use of whole genome sequencing by the Federal Interagency Collaboration for genomics for food and feed safety in the United States. J. Food Prot. 85, 755–772.

Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A., 2018. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus. Evolution 4 vey016.

The Norwegian Directorate of Health, 2022. Referansefunksjoner i medisinsk mikrobiologi. https://www.helsedirektoratet.no/tema/smittevern/referansefunksjoner-i-medisinsk-mikrobiologi#rapporteringogevaluering.

Thomas, J., Govender, N., McCarthy, K.M., Erasmus, L.K., Doyle, T.J., Allam, M., Ismail, A., Ramalwa, N., Sekwadi, P., Ntshoe, G., Shonhiwa, A., Essel, V., Tau, N., Smouse, S., Ngomane, H.M., Disenyeng, B., Page, N.A., Govender, N.P., Duse, A.G., Stewart, R., Thomas, T., Mahoney, D., Tourdjman, M., Disson, O., Thouvenot, P., Maury, M.M., Leclercq, A., Lecuit, M., Smith, A.M., Blumberg, L.H., 2020. Outbreak of listeriosis in South Africa associated with processed meat. N. Engl. J. Med. 382, 632–643.

Van Walle, I., Björkman, J.T., Cormican, M., Dallman, T., Mossong, J., Moura, A., Pietzka, A., Ruppitsch, W., Takkinen, J., European Listeria WGS typing group, 2018. Retrospective validation of whole genome sequencing-enhanced surveillance of listeriosis in Europe, 2010 to 2015. Eurosurveillance 23, 14–24.

Wagner, E., Fagerlund, A., Thalguter, S., Jensen, M.R., Heir, E., Møretrø, T., Moen, B., Langsrud, S., Rychli, K., 2022. Deciphering the virulence potential of *Listeria monocytogenes* in the Norwegian meat and salmon processing industry by combining whole genome sequencing and in vitro data. Int. J. Food Microbiol. 383, 109962.

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

Wick, R.R., 2018. PoreChop v0.2.4. https://github.com/rrwick/Porechop.

Wick, R.R., Judd, L.M., Gorrie, C.L., Holt, K.E., 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput. Biol. 13.

Zamudio, R., Haigh, R.D., Ralph, J.D., De Ste Croix, M., Tasara, T., Zurfluh, K., Kwun, M. J., Millard, A.D., Bentley, S.D., Croucher, N.J., Stephan, R., Oggioni, M.R., 2020. Lineage-specific evolution and gene flow in *Listeria monocytogenes* are independent of bacteriophages, 22, 5058–5072.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., Larsen, M.V., 2012. Identification of acquired antimicrobial resistance genes. J. Antimicrob. Chemother. 67, 2640–2644.