This postprint is published in:
Infection, Genetics and Evolution, 2020, Volume 77, Article 104047, Pages 1–9

DOI: 10.1016/j.meegid.2019.104047

Journal: Infection, Genetics and Evolution

Retrospective investigation of listeriosis outbreaks in small ruminants using different analytical approaches for whole genome sequencing-based typing of *Listeria monocytogenes*

Bojan Papić[a,*], Darja Kušar[a], Irena Zdovc[a], Majda Golob[a], Mateja Pate[a]

[a]University of Ljubljana, Veterinary Faculty, Institute of Microbiology and Parasitology, Gerbičeva 60, SI-1000 Ljubljana, Slovenia

[*]*Corresponding author*: Bojan Papić
Phone: +386 1 4779 810
Fax: +386 1 4779 352
E-mail: bojan.papic@vf.uni-lj.si

E-mail addresses: bojan.papic@vf.uni-lj.si (B. Papić), darja.kusar@vf.uni-lj.si (D. Kušar), irena.zdovc@vf.uni-lj.si (I. Zdovc), majda.golob@vf.uni-lj.si (M. Golob), mateja.pate@vf.uni-lj.si (M. Pate)

Running title: WGS typing of *Listeria monocytogenes* from animal outbreaks

Abstract

Listeria monocytogenes is the causative agent of listeriosis, a serious disease affecting both humans and animals. While listeriosis outbreaks in humans are commonly investigated in detail, routine typing of *L. monocytogenes* is generally not performed in animal outbreaks. Here, seven presumable listeriosis outbreaks in small ruminants were retrospectively identified based on the pulsed-field gel electrophoresis (PFGE) profiles. Outbreaks were further characterised using three different analytical approaches based on the whole-genome sequencing (WGS) data: core-genome multilocus sequence typing (cgMLST), whole-genome MLST (wgMLST) and whole-genome single nucleotide polymorphism (wgSNP) typing. A monoclonal pattern of all seven outbreaks was identified using all three approaches, indicating common-source outbreaks. The outbreak strains belonged to sequence types (STs) 1 (*n*=3), ST18 (*n*=1), ST21 (*n*=2) and ST184 (*n*=1). Two epidemiologically linked ST1 outbreaks with indistinguishable PFGE profiles showed a polyphyletic nature and differed in >78 SNPs; thus, they were classified as separate outbreaks according to WGS. In ST184, the outbreak strain was also found in faeces of apparently healthy ruminants, silage and water collected from the trough, which were the most likely source(s) of infection. The outbreak-associated isolates differed in 0–7 cgMLST alleles, 0–12 wgMLST alleles and 1–13 SNPs. The minimum genetic diversity between outbreak-associated isolates and epidemiologically unrelated isolates of the same ST was low in all analysed cases, approaching the maximum diversity within the outbreak cluster. The results

suggest that a fixed threshold to define the outbreak cluster should only be considered as a guide and highlight the role of epidemiological data for outbreak confirmation. The identified cgMLST clusters may be further investigated by wgMLST and/or wgSNP typing to increase confidence during investigations of outbreaks caused by highly clonal *L. monocytogenes* groups. This study gives an overview of the inter- and intra-outbreak genetic diversity of *L. monocytogenes* strains involved in animal outbreaks, hence improving their investigation.

Keywords: whole-genome sequencing (WGS), single nucleotide polymorphism (SNP), multilocus sequence typing (MLST), *Listeria monocytogenes*, outbreak, animal

1. Introduction

Listeria monocytogenes is ubiquitously present in the natural environment including the decaying vegetation, soil, surface water and the farm environment (Linke et al., 2014; Vivant et al., 2013). It is the causative agent of listeriosis, a potentially fatal infection in both humans and animals (Vázquez-Boland et al., 2001). Although the disease usually occurs sporadically, outbreaks are occasionally reported (Borucki et al., 2004; Walland et al., 2015). In animals, listeriosis outbreaks are primarily described in small ruminants (goats, sheep) and cattle (Akpavie and Ikheloa, 1992; Bundrant et al., 2011; Dreyer et al., 2015; Garcia et al., 2016; Vandegraaff et al., 1981; Wagner et al., 2005; Wardrope and MacLeod, 1983; Wiedmann et al., 1999).

Genotyping is an indispensable part of the surveillance of infectious diseases; however, there are only a few publications available on the animal listeriosis outbreaks that employed high-discriminatory genotyping methods such as ribotyping and pulsed-field gel electrophoresis (PFGE) (Bundrant et al., 2011; Dreyer et al., 2015; Stein et al., 2018; Wiedmann et al., 1999). For the past two decades, PFGE was considered a "gold standard" method for the typing of *L. monocytogenes* and other bacterial pathogens (Ribot et al., 2019; Roussel et al., 2014). Recently, whole-genome sequencing (WGS) has started to replace PFGE due to its unprecedented discriminatory power and the ability to discern the phylogenetic associations in comparison with PFGE and other genotyping methods (Jackson et al., 2016; Moura et al., 2016; Ribot et al., 2019). *L. monocytogenes* genome has a size of ~3 Mbp and exhibits an evolutionary rate of ~2.5×10^{-7} substitutions/site/year, indicating its highly stable nature (Moura et al., 2016).

The farm environment is an important reservoir of *L. monocytogenes* strains, but the knowledge on their transmission routes in this environment is limited (Castro et al., 2018; Ho et al., 2007; Nightingale et al., 2005; Nightingale et al., 2004). Apparently healthy cattle frequently shed *L. monocytogenes* into the farm environment, contributing to the maintenance and dispersal of the pathogen (Castro et al., 2018; Esteban et al., 2009; Ho et al., 2007; Nightingale et al., 2005). Animals, and the farm environment in general, also harbour the genotypes that are associated with human listeriosis (Borucki et al., 2004; Haley et al., 2015; Nightingale et al., 2004; Rocha et al., 2013; Vela et al., 2001). However, whether the animal strains are in fact able to cause disease in humans remains to be confirmed, as a large-scale, WGS-based comparison of human and animal *L. monocytogenes* strains has not been performed yet.

Different WGS-based analytical approaches have been widely used for the typing of *L. monocytogenes*, including the core-genome or whole-genome data, either at the level of allelic differences (multilocus sequence typing, MLST) or single nucleotide polymorphisms (SNPs). Most commonly applied approaches are core-genome MLST (cgMLST), whole-genome MLST (wgMLST) and whole-genome SNP (wgSNP) typing (Chen et al., 2017a;

Schürch et al., 2018). These were also used in the present study for the retrospective investigation of seven presumable listeriosis outbreaks in small ruminants previously identified by PFGE. Particular emphasis was given to the on-farm source investigation to establish the genetic relatedness of epidemiologically linked *L. monocytogenes* isolates of different origin (clinical, environmental) and to investigate the possible modes of transmission. To the best of our knowledge, this is the first use of WGS typing to investigate animal listeriosis outbreaks.

2. Methods

2.1. Isolate selection

Database of the Slovenian National Reference Laboratory for *L. monocytogenes* (SI NRL *Lm*), which mostly comprises isolates originating from animals, farm environment and food, was screened for presumable outbreak clusters based on the associated PFGE data and metadata (farm owner, year of isolation). In total, 24 isolates from six farms (A–F) were selected (Table 1). All isolates were streaked from frozen stocks onto sheep blood agar plates and incubated for 24 h at 37 ºC. To ensure strain purity, a single colony was subcultured on a new plate and used for PFGE and WGS typing. Outbreaks were defined according to the epidemiological (two or more geographically and temporally linked, confirmed cases of animal listeriosis; abortion, neurolisteriosis or septicaemia) and microbiological (>90 % similarity in combined *Asc*I-*Apa*I PFGE profiles to the isolate of an index case) criterion. Genome sequences of the genetically closely related, but epidemiologically unrelated isolates from the SI NRL *Lm* database and the NCBI Sequence Read Archive (SRA) database (https://www.ncbi.nlm.nih.gov/sra/) were also included in the WGS-based analysis.

2.2. PFGE

All isolates (*n*=24) were PFGE-typed according to the PulseNet standardised protocol (Graves and Swaminathan, 2001). Briefly, genomic DNA of *L. monocytogenes* isolates was digested with *Apa*I and *Asc*I restriction endonucleases. The generated fragments were separated using the CHEF-DR II system (Bio-Rad, USA) according to the recommended electrophoretic protocol. PFGE profiles were analysed with BioNumerics v7.6.3 (Applied Maths, Belgium). A combined *Asc*I-*Apa*I profile analysis was performed as described previously (Félix et al., 2018).

2.3. WGS

All isolates (*n*=24) were subjected to WGS with the Illumina technology (Illumina, USA). Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen, Germany) with the recommended pre-treatment for Gram-positive bacteria. WGS libraries were prepared using the Illumina TruSeq Nano DNA or NEBNext Ultra II DNA library prep kit. Paired-end sequencing (2×100 or 2×150 bp) was performed on the NextSeq 500 Illumina platform to a minimum coverage of 100×. The raw reads were deposited into the NCBI BioProject database (https://www.ncbi.nlm.nih.gov/bioproject/) under the accession number PRJNA532857. In addition, genome sequences of the 19 most closely related isolates were retrieved from the SI NRL *Lm* and NCBI SRA databases for comparison. Raw read data were available for 18/19 isolates, whereas for the KSM2 isolate (ST21), only the assembled genomes were available; thus, it was excluded from the wgSNP analysis. See Supplementary Table S1 for isolate metadata.

2.3.1. WGS-based typing

All isolates were typed with three analytical approaches implemented in BioNumerics v7.6.3 (Applied Maths, Belgium): cgMLST (based on profiles of 1784 coding loci in the BigsDB Pasteur cgMLST scheme), wgMLST (based on profiles of 4797 coding loci in the BioNumerics scheme) and wgSNP typing. For cg/wgMLST analysis, assembly-based and assembly-free allele calling was performed using the default settings and dendrograms based on summary calls were constructed by applying the categorical similarity coefficient and unweighted pair group method with arithmetic average (UPGMA). Raw read, *de novo* assembly and wgMLST quality statistics were calculated using the quality assessment window implemented in BioNumerics v7.6.3 and are reported in Supplementary Table S2. For wgSNP typing, a closely related complete or draft genome of the same sequence type (ST) was used as a reference genome (Table 2) to maximise the size of the shared genome and the 'Strict SNP filtering (closed SNP set)' template was used with the following settings: minimum inter-SNP distance of 12 bp, minimum total coverage of 5, minimum forward coverage of 1, minimum reverse coverage of 1. Positions with ambiguous bases, unreliable bases or gaps were excluded. A concatenated SNP alignment was used to construct a maximum-likelihood phylogenetic tree with RAxML v8.1.22 (Stamatakis, 2014) by applying the -m GTRGAMMA -# 1000 parameters. *In silico* 7-gene MLST typing, which classifies *L. monocytogenes* isolates into STs and clonal complexes (CCs), was performed using the Sequence query tool implemented in the Institut Pasteur BIGSdb-Lm database (https://bigsdb.pasteur.fr/). Prophages were identified using PHASTER (Arndt et al., 2016); only the intact prophage sequences present in the outbreak-associated isolates were included into the analysis. The presence of known *L. monocytogenes* virulence genes in prophage sequences was examined using VirulenceFinder v2.0 (https://cge.cbs.dtu.dk/services/VirulenceFinder/). Putative plasmid sequences were assembled using plasmidSPAdes v3.13.1 (Antipov et al., 2016) and identified by aligning against the NCBI NT database using BLASTn.

3. Results

3.1. Genetic diversity of the outbreak-associated isolates

All genomes passed the quality check with the BioNumerics software (Supplementary Table S2). Within each of the seven presumable outbreaks, the outbreak-associated isolates exhibited indistinguishable *Asc*I-*Apa*I profiles (data not shown). The isolates were of four STs: ST1 (*n*=3), ST18 (*n*=1), ST21 (*n*=2) and ST184 (*n*=1) (Table 1). Within a single outbreak, all clinical and environmental isolates originated from the same farm and all outbreaks were of short duration (less than one month), suggesting common-source, point outbreaks. The outbreak-associated isolates differed in 0–7 cgMLST alleles, 0–12 wgMLST alleles and 1–13 SNPs (Figs. 1–4, Table 3); all presumable outbreaks demonstrated a monoclonal pattern and were confirmed by all three WGS analytical approaches. The outbreak 1 and 3 clusters encompassed the ST1 isolates with indistinguishable PFGE profiles (data not shown), which originated from the same farm and dated to the same year, but exhibited a polyphyletic branching and a pairwise distance of >78 SNPs. Therefore, they were considered as two different outbreaks.

The wgSNP and wgMLST approaches identified a similar amount of genetic variation within an outbreak cluster, indicating their comparable discriminatory power. Both approaches had a higher discriminatory power compared with cgMLST (Figs. 1–4, Table 3). All seven outbreaks were supported by the associated metadata, i.e. isolates originated from the same farm and were isolated in the same year (Table 1). The SNP-based phylogenetic trees showed that all seven

outbreaks formed well-supported monophyletic groups; the outbreak 6 cluster was the only one characterised by a low bootstrap support (Fig. 3).

For outbreaks 1–3 (ST1), the minimum inter-outbreak diversity was higher than in outbreaks 5 and 6 (ST21) (Table 4). The highest intra-outbreak diversity was observed in outbreak 6 (ST21; cgMLST and wgSNP) and outbreak 7 (ST184; wgMLST) (Table 3). In outbreaks 5 and 6, which occurred eight months apart in two different farms of the same municipality (Table 1), the minimum inter-outbreak diversity approached (cg/wgMLST) or surpassed (wgSNP) the maximum intra-outbreak diversity (Tables 3 and 4). The observed genetic distances and geographical proximity of both farms suggest these two outbreaks are part of the same transmission cluster; however, no epidemiological data were available to reliably confirm their linkage.

3.2. ST184 outbreak

In outbreak 7, the investigation encompassed the animal clinical (*n*=5; six animals were showing clinical signs, but isolates from five animals were obtained) and environmental (*n*=6) isolates, including isolates from the faeces of asymptomatic ruminants, silage and water collected from the trough. All isolates belonged to ST184 and differed in up to 6 SNPs (Table 3). Thus, most likely, the asymptomatic carriers shed the outbreak strain into the environment, contaminating the silage and water, which were the most likely sources of infection for the animals (Fig. 4).

3.3. Mobile genetic element distribution

The outbreak 2 isolates (ST1) harboured an identical ϕtRNA-Lys prophage; a highly similar prophage (97.5 % identity) was present in one sporadic isolate (L1364). In other ST1 outbreak-associated isolates, intact prophages were not identified. The outbreak 4 isolates (ST18) harboured two intact prophages (ϕ*comK* and ϕtRNA-Ser), which were identical in all three outbreak-associated isolates, but absent in all sporadic isolates. In the outbreak 5 and 6 isolates (ST21), intact prophages were not identified. The outbreak 7 isolates (ST184) all harboured one identical ϕtRNA-Arg prophage; a highly similar prophage (96.6 % identity) was also present in one sporadic isolate (L1174). None of the identified prophage sequences harboured any of the known *L. monocytogenes* virulence traits. No plasmid sequences were identified in the analysed isolates.

4. Discussion

In the present study, cgMLST, wgMLST and wgSNP were used for the retrospective investigation of seven listeriosis outbreaks in small ruminants. All three approaches indicated a monoclonal pattern of the investigated outbreaks; the outbreak-associated isolates differed in up to 7 cgMLST alleles, 12 wgMLST alleles and 13 SNPs. Although these approaches have been widely used in listeriosis surveillance and outbreak investigations and a high concordance between them has been shown (Henri et al., 2017), a discordant clustering of closely related isolates has also been reported (Chen et al., 2017a; Jagadeesan et al., 2019; Li et al., 2017). The use of multiple WGS analytical approaches maximises the discovery of genetic differences in closely related isolates and hence reduces the effect of their limitations (Chen et al., 2017a).

Several epidemiologically unlinked *L. monocytogenes* isolates, which we added to the analysis of outbreaks on the basis of a close genetic similarity/relatedness, clustered within the outbreak

lineage in the case of ST1 and ST21 outbreaks. In addition, the pairwise genetic distance among the isolates originating from different outbreaks of the same ST was low (ST1 isolates associated with outbreaks 1–3, and ST21 isolates associated with outbreaks 5 and 6); the minimum inter-outbreak diversity approached and even surpassed the maximum intra-outbreak diversity in the case of outbreaks 5 and 6 (both ST21). A high genetic relatedness and geographical proximity of outbreaks 5 and 6 suggest their linkage; however, no between-farm contract data were available to confirm their linkage. Thus, they may be linked through an undetected transmission network such as feed or animal exchange.

Although fixed allele/SNP thresholds to delineate outbreak clusters of *L. monocytogenes* have been proposed, e.g. 7–10 cgMLST alleles or 5 SNPs (EFSA-ECDC, 2018; Moura et al., 2016; Ruppitsch et al., 2015), their use would lead to a spurious attribution of the epidemiologically unlinked isolates to the outbreaks in the present study if the epidemiological data had not been considered. Several publications reported that the genetic diversity of *L. monocytogenes* isolates collected during epidemiological investigations may exceed the proposed thresholds (Chen et al., 2016; Chen et al., 2017b; Jackson et al., 2016), in particular in the long-term outbreaks and cases in which contamination may originate from different sources and/or sources with a genetically diverse population of *L. monocytogenes*, e.g. the natural environment (Pightling et al., 2018; Reimer et al., 2019). Hence, inherent genetic diversity of the organism in isolation source and its mutation rate should also be considered. Moreover, the epidemiologically unrelated isolates (e.g. sporadic cases) may fall into the outbreak cluster, which in itself does not confirm their association with the outbreak. Therefore, the use of a fixed threshold to delineate the outbreak cluster should only act as a guide for its identification, but the diversity of isolates should always be interpreted in the context of a given WGS phylogeny with a good bootstrap support and epidemiological evidence, as suggested before (Chen et al., 2017a; Chen et al., 2017b; Pightling et al., 2018; Reimer et al., 2019).

Although WGS has been shown to enable reliable end-to-end typing of *L. monocytogenes* and other pathogens, variability might arise at several stages of WGS workflow, including strain maintenance, DNA extraction, library preparation, sequencing and bioinformatic analysis (Portmann et al., 2018; Saltykova et al., 2018). Studies investigating the effect of different WGS analytical approaches and/or technical robustness of WGS in highly clonal *L. monocytogenes* groups are limited (Jagadeesan et al., 2019; Li et al., 2017; Reimer et al., 2018). In this study, a single sequencing platform was applied to a similar depth of coverage. Additionally, an exhaustive quality assessment of WGS data and standardised WGS workflows implemented in the BioNumerics software were employed, all minimising the effect of the aforementioned variables on WGS results. Due to the continuous evolution of WGS workflows, comparisons of different platforms, analytical approaches, bioinformatic tools and their settings remain paramount before establishing harmonised WGS-based typing workflows.

In the present study, all three WGS analytical approaches exerted a higher discriminatory power compared with PFGE. This was particularly evident in outbreaks 1 and 3, which included the ST1 isolates with indistinguishable PFGE profiles. Furthermore, they were epidemiologically related, but showed a pairwise distance of >78 SNPs and were thus considered as two separate outbreaks. The discriminatory power wgSNP and wgMLST was comparable and higher than in cgMLST, which is in accordance with previous findings (Henri et al., 2017; Jagadeesan et al., 2019). In prospective outbreak investigation, the use of multiple analytical approaches may be prohibitively time-consuming and resource-intensive. The cgMLST approach, which favours data portability, requires less bioinformatics expertise and is computationally less exhaustive than wgSNP typing, provides a high-discriminatory method, which is in most cases sufficient

for reliable cluster identification (Henri et al., 2017; Jagadeesan et al., 2019). However, the results presented herein support previous findings that wgMLST and/or wgSNP should be used on clusters identified by cgMLST to increase the discriminatory power and reliably estimate pairwise genetic distances, in particular when analysing highly clonal microbial groups (Jagadeesan et al., 2019; Li et al., 2017). Compared with allele-based approaches, wgSNP enables the construction of statistically supported phylogenetic trees and estimation of divergence times, providing an insight into the evolutionary dynamics of the outbreak strain (Chen et al., 2017c; Jagadeesan et al., 2019).

Outbreak isolates belonged to ST1 (CC1), ST18 (CC18), ST21 (CC21) and singleton ST184. CC1 is a hypervirulent clone, which is over-represented among the human and animal clinical isolates (Dreyer et al., 2016; Maury et al., 2016). CC18 and CC21 have also been frequently associated with human disease (Althaus et al., 2014; Maury et al., 2016). On the other hand, ST184 is a rare clone and has, to the best of our knowledge, not yet been associated with listeriosis outbreaks in animals. At the time of writing, no ST184 isolates were recorded in the Institut Pasteur BIGSdb-*Lm* database (https://bigsdb.pasteur.fr/listeria/listeria.html) and a single human clinical ST184 isolate was identified in the NCBI Pathogen Detection tool (https://www.ncbi.nlm.nih.gov/pathogens/). Altogether, because the outbreak strains have the ability to cause listeriosis and most of the animal outbreak CCs from this study are commonly associated with human listeriosis, this supports the hypothesis that the animal-associated strains may have an increased ability to cause disease in humans.

In the ST184 outbreak (outbreak 7), all clinical and environmental isolates belonged to the same outbreak strain. This outbreak had the highest intra-outbreak genetic diversity in comparison with other outbreaks, which may be explained by a higher number of analysed samples of different origin. The presence of the outbreak strain in the faeces of asymptomatic ruminants underlines the importance of faecal shedding in the transmission of listeriosis, leading to faecal contamination of the farm environment, including the silage and water collected from the trough. The contaminated silage was the most likely source of infection of the diseased animals in outbreak 7, as the association of low-quality silage with ruminant listeriosis has been well established (Garcia et al., 2016; Nightingale et al., 2004; Wagner et al., 2005). Furthermore, the contaminated water may have had contributed to the infection dose. These results suggest that the implementation of targeted preventive measures such as improved control of the animal feed as well as good silage production and utilisation practices could decrease the incidence of listeriosis on ruminant farms (Bernardes and do Rego, 2014).

A single clinical form of listeriosis (abortion or neurolisteriosis) was observed in 6/7 outbreaks, whereas in the ST184 outbreak, neurolisteriosis and septicaemia were both reported. Different clinical forms rarely overlap in a single animal or during the outbreak (Garcia et al., 2016; Oevermann et al., 2010; Wagner et al., 2005), which was also supported by the present study. However, the ST184 outbreak was the only outbreak that also involved young animals and two different animal species (sheep, goats). Neurolisteriosis was observed in three animals (one goat and two lambs), whereas septicaemia was noted in two goat kids. This may partially be explained by additional host and environmental factors, e.g. comorbid conditions, infectious dose, route of infection and host age, in particular the latter. Namely, septicaemia is the most frequent manifestation of listeriosis in neonates and lambs, whereas in adult animals, neurolisteriosis and abortion are most commonly reported (Walland et al., 2015; Wesley, 2007).

L. monocytogenes clones have been shown to differ in their virulence and organ tropism, with hypervirulent clones CC1, CC2, CC4 and CC6 being strongly associated with a clinical origin

(Dreyer et al., 2016). In the present study, CC1 was responsible for 3/7 outbreaks, which all included neurolisteriosis cases. This is in agreement with previous findings, in which CC1 has been shown to be over-represented in bovine neurolisteriosis, suggesting its increased neurotropism (Aguilar-Bultet et al., 2018; Dreyer et al., 2016). Therefore, the findings presented herein support the hypothesis that the farm environment is a reservoir of strains with increased propensity for invasive listeriosis in humans, although additional comparative genomics studies are needed to confirm this hypothesis.

The present study included a small number of outbreak-associated *L. monocytogenes* isolates, as the post-mortem examination is usually performed only for the first few animals involved in a suspected outbreak. Increased taxon sampling within the outbreak clusters (i.e. larger number of analysed isolates) may result in a more accurate phylogenetic inference (Pollock et al., 2002). Nevertheless, in 6/7 outbreaks, isolates showed a well-supported monophyletic grouping, indicating the presence of individual outbreak strains. The outbreak 6 isolates also showed a monophyletic grouping, but a low bootstrap support; however, the epidemiological data and the low pairwise SNP difference (13 SNPs) indicate their outbreak association. As previously described, temporally and geographically unrelated isolates may appear related even at a maximum genome resolution, highlighting the importance of epidemiological evidence (Reimer et al., 2019). In the present study, with the exception of outbreaks 1 and 3, there was no temporal or geographical association between different outbreaks. However, it should be noted that a well-conducted contact tracing, which includes the data on animal movements, farm visitors, feed, vehicles or equipment (Nöremark and Widgren, 2014), might reveal additional associations between the outbreaks.

In this study, prophage distribution was also analysed, but was informative in only 3/7 outbreaks, in which the observed prophage profile was characteristic for the outbreak-associated strains. The analysis of mobile genetic elements such as prophages and plasmids has been previously shown useful in the delineation of *L. monocytogenes* outbreak clusters (Wang et al., 2015). However, caution should be applied when including the mobile genetic elements into phylogenetic analyses as they can obscure the true phylogenetic relationships among isolates obtained during the outbreak investigations, as has been shown for *L. monocytogenes* CC8 isolates (Reimer et al., 2019).

In conclusion, the present study adds to our understanding of the genetic diversity of the animal outbreak-associated *L. monocytogenes* strains, aiming to facilitate the interpretation of WGS data in the investigation of listeriosis outbreaks. All three WGS-based analytical approaches used in this study (cgMLST, wgMLST and wgSNP) were able to confirm the monoclonal pattern of the outbreaks, but wgMLST and wgSNP had a higher discriminatory power compared with cgMLST. Genetically most closely related, but epidemiologically unrelated isolates could not be reliably separated from the outbreak-associated isolates. In addition to the importance of epidemiological data for the outbreak confirmation, the present study highlights the importance of performing extensive sampling and employing different WGS-based analytical approaches to increase confidence in strain associations during outbreak investigation.

Supplementary data to this article can be found online at

Funding

The authors acknowledge this work was financially supported by the Slovenian Research Agency (grant for young researchers for B. Papić).

Abbreviations

CC: clonal complex
cgMLST: core-genome multilocus sequence typing
ST: sequence type
wgMLST: whole-genome multilocus sequence typing
WGS: whole-genome sequencing
wgSNP: whole-genome single nucleotide polymorphism

Declaration of interest

None.

Acknowledgements

The authors thank Aleš Budkovič (Veterinary Faculty, Ljubljana) for his help with the implementation of bioinformatics tools. Prof Dr Polona Juntes (Veterinary Faculty, Ljubljana) is acknowledged for providing the metadata associated with the animal clinical isolates. Maja Kavalič (Veterinary Faculty, Ljubljana) is acknowledged for performing PFGE typing. Maja Lepen and Olga Kosar (both Veterinary Faculty, Ljubljana) are acknowledged for the database maintenance of *L. monocytogenes* strains. The authors warmly acknowledge the support of Benjamin Félix, Federica Palma and Yann Sévellec from ANSES (Maisons-Alfort, France) for fruitful discussions and assistance with the bioinformatics tools employed in this study.

References

Aguilar-Bultet, L., Nicholson, P., Rychener, L., Dreyer, M., Gözel, B., Origgi, F.C., Oevermann, A., Frey, J., Falquet, L., 2018. Genetic separation of *Listeria monocytogenes* causing central nervous system infections in animals. Front Cell Infect Microbiol. 8, 20.

Akpavie, S.O., Ikheloa, J.O., 1992. An outbreak of listeriosis in cattle in Nigeria. Rev Elev Med Vet Pays Trop. 45, 263–264.

Althaus, D., Lehner, A., Brisse, S., Maury, M., Tasara, T., Stephan, R., 2014. Characterization of *Listeria monocytogenes* strains isolated during 2011–2013 from human infections in Switzerland. Foodborne Pathog Dis. 11, 753–758.

Antipov, D., Hartwick, N., Shen, M., Raiko, M., Lapidus, A., Pevzner, P.A., 2016. plasmidSPAdes: assembling plasmids from whole genome sequencing data. Bioinformatics 32, 3380–3387.

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

Bernardes, T.F., do Rego, A.C., 2014. Study on the practices of silage production and utilization on Brazilian dairy farms. J Dairy Sci. 97, 1852–1861.

Borucki, M.K., Reynolds, J., Gay, C.C., McElwain, K.L., Kim, S.H., Knowles, D.P., Hu, J., 2004. Dairy farm reservoir of *Listeria monocytogenes* sporadic and epidemic strains. J Food Prot. 67, 2496–2499.

Bundrant, B.N., Hutchins, T., den Bakker, H.C., Fortes, E., Wiedmann, M., 2011. Listeriosis outbreak in dairy cattle caused by an unusual *Listeria monocytogenes* serotype 4b strain. J Vet Diagn Invest. 23, 155–158.

Castro, H., Jaakkonen, A., Hakkinen, M., Korkeala, H., Lindström, M., 2018. Occurrence, persistence, and contamination routes of *Listeria monocytogenes* genotypes on three Finnish dairy cattle farms: a longitudinal study. Appl Environ Microbiol. 84, e02000-17.

Chen, Y., Gonzalez-Escalona, N., Hammack, T.S., Allard, M.W., Strain, E.A., Brown, E.W., 2016. Core genome multilocus sequence typing for identification of globally distributed clonal groups and differentiation of outbreak strains of *Listeria monocytogenes.* Appl Environ Microbiol. 82, 6258–6272.

Chen, Y., Luo, Y., Carleton, H., Timme, R., Melka, D., Muruvanda, T., Wang, C., Kastanis, G., Katz, L.S., Turner, L., Fritzinger, A., Moore, T., Stones, R., Blankenship, J., Salter, M., Parish, M., Hammack, T.S., Evans, P.S., Tarr, C.L., Allard, M.W., Strain, E.A., Brown, E.W., 2017a. Whole genome and core genome multilocus sequence typing and single nucleotide polymorphism analyses of *Listeria monocytogenes* associated with an outbreak linked to cheese, United States, 2013. Appl Environ Microbiol. 83, e00633-17.

Chen, Y., Luo, Y., Curry, P., Timme, R., Melka, D., Doyle, M., Parish, M., Hammack, T.S., Allard, M.W., Brown, E.W., Strain, E.A., 2017b. Assessing the genome level diversity of *Listeria monocytogenes* from contaminated ice cream and environmental samples linked to a listeriosis outbreak in the United States. PLoS One 12, e0171389.

Chen, Y., Luo, Y., Pettengill, J., Timme R., Melka, D., Doyle, M., Jackson, A., Parish, M, Hammack, T.S., Allard, M.W., Brown, E.W., Strain, E.A., 2017c. Singleton sequence type 382, an emerging clonal group of *Listeria monocytogenes* associated with three multistate outbreaks linked to contaminated stone fruit, caramel apples, and leafy green salad. J Clin Microbiol. 55, 931–941.

Dreyer, M., Aguilar-Bultet, L., Rupp, S., Guldimann, C., Stephan, R., Schock, A., Otter, A., Schupbach, G., Brisse, S., Lecuit, M., Frey, J., Oevermann, A., 2016. *Listeria monocytogenes* sequence type 1 is predominant in ruminant rhombencephalitis. Sci Rep. 6, 36419.

Dreyer, M., Thomann, A., Böttcher, S., Frey, J., Oevermann, A., 2015. Outbreak investigation identifies a single *Listeria monocytogenes* strain in sheep with different clinical manifestations, soil and water. Vet Microbiol. 179, 69–75.

EFSA-ECDC 2018. Multi-country outbreak of *Listeria monocytogenes* serogroup IVb, multi-locus sequence type 6, infections probably linked to frozen corn – first update. EFSA supporting publication 2018:EN-1448. 22 pp. doi: 10.2903/sp.efsa.2018.EN-1448

Esteban, J.I., Oporto, B., Aduriz, G., Juste, R.A., Hurtado, A., 2009. Faecal shedding and strain diversity of *Listeria monocytogenes* in healthy ruminants and swine in Northern Spain. BMC Vet Res. 5, 2.

Félix, B., Feurer, C., Maillet, A., Guillier, L., Boscher, E., Kerouanton, A., Denis, M., Roussel, S., 2018. Population Genetic Structure of. Front Microbiol. 9, 684.

Garcia, J.A., Micheloud, J.F., Campero, C.M., Morrell, E.L., Odriozola, E.R., Moreira, A.R., 2016. Enteric listeriosis in grazing steers supplemented with spoiled silage. J Vet Diagn Invest. 28, 65–69.

Graves, L.M., Swaminathan, B., 2001. PulseNet standardized protocol for subtyping *Listeria monocytogenes* by macrorestriction and pulsed-field gel electrophoresis. Int J Food Microbiol. 65, 55–62.

Haley, B.J., Sonnier, J., Schukken, Y.H., Karns, J.S., Van Kessel, J.A., 2015. Diversity of *Listeria monocytogenes* within a U.S. dairy herd, 2004–2010. Foodborne Pathog Dis. 12, 844–850.

Henri, C., Leekitcharoenphon, P., Carleton, H.A., Radomski, N., Kaas, R.S., Mariet, J.F., Felten, A., Aarestrup, F.M., Gerner Smidt, P., Roussel, S., Guillier, L., Mistou, M.Y.,

Hendriksen, R.S., 2017. An Assessment of different genomic approaches for inferring phylogeny of *Listeria monocytogenes*. Front Microbiol. 8, 2351.

Ho, A.J., Ivanek, R., Grohn, Y.T., Nightingale, K.K., Wiedmann, M., 2007. *Listeria monocytogenes* fecal shedding in dairy cattle shows high levels of day-to-day variation and includes outbreaks and sporadic cases of shedding of specific *L. monocytogenes* subtypes. Prev Vet Med. 80, 287–305.

Jackson, B.R., Tarr, C., Strain, E., Jackson, K.A., Conrad, A., Carleton, H., Katz, L.S., Stroika, S., Gould, L.H., Mody, R.K., Silk, B.J., Beal, J., Chen, Y., Timme, R., Doyle, M., Fields, A., Wise, M., Tillman, G., Defibaugh-Chavez, S., Kucerova, Z., Sabol, A., Roache, K., Trees, E., Simmons, M., Wasilenko, J., Kubota, K., Pouseele, H., Klimke, W., Besser, J., Brown, E., Allard, M., Gerner-Smidt, P., 2016. Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. Clin Infect Dis. 63, 380–386.

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

Li, Z., Pérez-Osorio, A., Wang, Y., Eckmann, K., Glover, W.A., Allard, M.W., Brown, E.W., Chen, Y., 2017. Whole genome sequencing analyses of *Listeria monocytogenes* that persisted in a milkshake machine for a year and caused illnesses in Washington State. BMC Microbiol. 17, 134.

Linke, K., Rückerl, I., Brugger, K., Karpiskova, R., Walland, J., Muri-Klinger, S., Tichy, A., Wagner, M., Stessl, B., 2014. Reservoirs of *Listeria* species in three environmental ecosystems. Appl Environ Microbiol. 80, 5583–5592.

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.

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., Nadon, C., Grant, K., Nielsen, E.M., Pot, B., Gerner-Smidt, P., Lecuit, M., Brisse, S., 2016. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. Nat Microbiol. 2, 16185.

Nightingale, K.K., Fortes, E.D., Ho, A.J., Schukken, Y.H., Grohn, Y.T., Wiedmann, M., 2005. Evaluation of farm management practices as risk factors for clinical listeriosis and fecal shedding of *Listeria monocytogenes* in ruminants. J Am Vet Med Assoc. 227, 1808–1814.

Nightingale, K.K., Schukken, Y.H., Nightingale, C.R., Fortes, E.D., Ho, A.J., Her, Z., Grohn, Y.T., McDonough, P.L., Wiedmann, M., 2004. Ecology and transmission of *Listeria monocytogenes* infecting ruminants and in the farm environment. Appl Environ Microbiol. 70, 4458–4467.

Nöremark, M., Widgren, S., 2014. EpiContactTrace: an R-package for contact tracing during livestock disease outbreaks and for risk-based surveillance. BMC Vet Res. 10, 71.

Oevermann, A., Zurbriggen, A., Vandevelde, M., 2010. Rhombencephalitis caused by *Listeria monocytogenes* in humans and ruminants: a zoonosis on the rise? Interdiscip Perspect Infect Dis. 2010, 632513.

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

Pollock, D.D., Zwickl, D.J., McGuire, J.A., Hillis, D.M., 2002. Increased taxon sampling is advantageous for phylogenetic inference. Syst Biol. 51, 664–671.

Portmann A.C., Fournier C., Gimonet J., Ngom-Bru C., Barretto C., Baert L., 2018. A validation approach of an end-to-end whole genome sequencing workflow for source tracking of *Listeria monocytogenes* and *Salmonella enterica*. Front Microbiol. 9:446.

Reimer, A., Weedmark, K., Petkau, A., Peterson, C.L., Walker, M., Knox, N., Kent, H., Mabon, P., Berry, C., Tyler, S., Tschetter, L., Jerome, M., Allen, V., Hoang, L., Bekal, S., Clark, C., Nadon, C., Van Domselaar, G., Pagotto, F., Graham, M., Farber, J., Gilmour, M., 2019. Shared genome analyses of notable listeriosis outbreaks, highlighting the critical importance of epidemiological evidence, input datasets and interpretation criteria. Microb Genom. 5.

Ribot, E.M., Freeman, M., Kelley B. Hise, K.B., Gerner-Smidt P., 2019. PulseNet: entering the age of next-generation sequencing. Foodborne Pathog Dis. 16: 451–456.

Rocha, P.R., Lomonaco, S., Bottero, M.T., Dalmasso, A., Dondo, A., Grattarola, C., Zuccon, F., Iulini, B., Knabel, S.J., Capucchio, M.T., Casalone, C., 2013. Ruminant rhombencephalitis-associated *Listeria monocytogenes* strains constitute a genetically homogeneous group related to human outbreak strains. Appl Environ Microbiol. 79, 3059–3066.

Roussel, S., Michelon, D., Lombard, B., Lailler, R. 2014. Molecular typing of *Listeria monocytogenes* strains isolated from food, feed and animals: state of play and standard operating procedures for pulsed field gel electrophoresis (PFGE) typing, profile interpretation and curation. EFSA supporting publication 2014:EN-702, 81 pp. doi: 10.2903/sp.efsa.2014.EN-702.

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.

Saltykova, A., Wuyts, V., Mattheus, W., Bertrand, S., Roosens N.H.C., Marchal, K., De Keersmaecker, S.C.J., 2018. Comparison of SNP-based subtyping workflows for bacterial isolates using WGS data, applied to *Salmonella enterica* serotype Typhimurium and serotype 1,4,[5],12:i:-. PLoS ONE 13, e0192504.

Schürch, A.C., Arredondo-Alonso, S., Willems, R.J.L., Goering, R.V., 2018. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. Clin Microbiol Infect. 24, 350–354.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.

Stein, H., Stessl, B., Brunthaler, R., Loncaric, I., Weissenbock, H., Ruczizka, U., Ladinig, A., Schwarz, L., 2018. Listeriosis in fattening pigs caused by poor quality silage - a case report. BMC Vet Res. 14, 362.

Vandegraaff, R., Borland, N.A., Browning, J.W., 1981. An outbreak of listerial meningo-encephalitis in sheep. Aust Vet J. 57, 94–96.

Vela, A.I., Fernandez-Garayzabal, J.F., Vazquez, J.A., Latre, M.V., Blanco, M.M., Moreno, M.A., de La Fuente, L., Marco, J., Franco, C., Cepeda, A., Rodriguez Moure, A.A., Suarez, G., Dominguez, L., 2001. Molecular typing by pulsed-field gel electrophoresis of Spanish animal and human *Listeria monocytogenes* isolates. Appl Environ Microbiol. 67, 5840–5843.

Vivant, A.L., Garmyn, D., Piveteau, P., 2013. *Listeria monocytogenes*, a down-to-earth pathogen. Front Cell Infect Microbiol. 3, 87.

Vázquez-Boland, J.A., Kuhn, M., Berche, P., Chakraborty, T., Domínguez-Bernal, G., Goebel, W., González-Zorn, B., Wehland, J., Kreft, J. 2001. *Listeria* pathogenesis and molecular virulence determinants. Clin Microbiol Rev. 14, 584–640.

Wagner, M., Melzner, D., Bagò, Z., Winter, P., Egerbacher, M., Schilcher, F., Zangana, A., Schoder, D., 2005. Outbreak of clinical listeriosis in sheep: evaluation from possible contamination routes from feed to raw produce and humans. J Vet Med B Infect Dis Vet Public Health 52, 278–283.

Walland, J., Lauper, J., Frey, J., Imhof, R., Stephan, R., Seuberlich, T., Oevermann, A., 2015. *Listeria monocytogenes* infection in ruminants: Is there a link to the environment, food and human health? A review. Schweiz Arch Tierheilkd. 157, 319–328.

Wang, Q., Holmes, N., Martinez, E., Howard, P., Hill-Cawthorne, G., Sintchenko, V., 2015. It is not all about single nucleotide polymorphisms: comparison of mobile genetic elements and deletions in *Listeria monocytogenes* genomes links cases of hospital-acquired listeriosis to the environmental source. J Clin Microbiol. 53, 3492–3500.

Wardrope, D.D., MacLeod, N.S., 1983. Outbreak of *Listeria* meningoencephalitis in young lambs. Vet Rec. 113, 213–214.

Wesley, I. 2007. Listeriosis in animals. In: *Listeria*, Listeriosis, and Food Safety. Ryser, E.T., Marth, M.E. (eds.). Boca Raton: CRC Press, 55–84.

Wiedmann, M., Mobini, S., Cole, J.R., Watson, C.K., Jeffers, G.T., Boor, K.J., 1999. Molecular investigation of a listeriosis outbreak in goats caused by an unusual strain of *Listeria monocytogenes*. J Am Vet Med Assoc. 215, 369–371.

Figure legends

Fig. 1. ST1 outbreaks. (A) Phylogenetic tree based on the wgSNP approach. Maximum-likelihood phylogeny was inferred from the concatenated SNP alignment (279 SNPs). The numbers above and below the branches indicate bootstrap values and pairwise SNP distances. Full squares denote the presence of intact prophages, whereas empty squares denote their absence. Brighter colour denotes a lower similarity (<98 % identity) of the prophage sequences. Bar, number of substitutions per nucleotide position. (B) Dendrogram based on the cgMLST approach. Bar, number of allelic differences. (C) Dendrogram based on the wgMLST approach. Bar, number of allelic differences. The outbreak strains are labelled in red.

Fig. 2. ST18 outbreak. (A) Phylogenetic tree based on the wgSNP approach. Maximum-likelihood phylogeny was inferred from the concatenated SNP alignment (217 SNPs). (B) Dendrogram based on the cgMLST approach. (C) Dendrogram based on the wgMLST approach. For additional information, see Fig. 1.

Fig. 3. ST21 outbreaks. (A) Phylogenetic tree based on the wgSNP approach. Maximum-likelihood phylogeny was inferred from the concatenated SNP alignment (139 SNPs). (B) Dendrogram based on the cgMLST approach. (C) Dendrogram based on the wgMLST approach. For additional information, see Fig. 1.

Fig. 4. ST184 outbreak. (A) Phylogenetic tree based on the wgSNP approach. Maximum-likelihood phylogeny was inferred from the concatenated SNP alignment (111 SNPs). (B) Dendrogram based on the cgMLST approach. (C) Dendrogram based on the wgMLST approach. For additional information, see Fig. 1.

Tables

Table 1: Metadata associated with *Listeria monocytogenes* isolates investigated in this study.

Outbreak	Onset time	Clinical form†	Farm	Affected animal species	No. of affected animals	No. of available isolates	MLST ST	MLST CC
Outbreak 1	January 2018	Neurolisteriosis	Farm A	Goats	ND	2	ST1	CC1
Outbreak 2	March 2008	Neurolisteriosis	Farm B	Sheep	3	2	ST1	CC1
Outbreak 3	April 2018	Neurolisteriosis	Farm A	Goats	5	2	ST1	CC1
Outbreak 4	February 2013	Abortion	Farm C	Goats	ND	3	ST18	CC18
Outbreak 5	May 2014	Neurolisteriosis	Farm D	Goats	3	2	ST21	CC21
Outbreak 6	January– February 2015	Neurolisteriosis	Farm E	Sheep	ND	2	ST21	CC21
Outbreak 7	May–June 2013	Neurolisteriosis, septicaemia	Farm F	Goats, sheep	6	11	ST184	ST184

†Neurolisteriosis was defined as isolation of *L. monocytogenes* from brain tissue and the presence of pathognomonic lesions in brain tissue (as determined by histopathology). Listerial abortion was defined as isolation of *L. monocytogenes* from the aborted foetus or maternal reproductive tissues (placenta, amniotic fluid) immediately after the abortion. Septicaemia was defined here as isolation of *L. monocytogenes* from visceral organs or brain tissue without pathognomonic changes characteristic for neurolisteriosis. All clinical isolates derived from different animals. In the ST184 outbreak, five isolates originated from clinical cases, whereas six isolates originated from the environment. ND, not defined.

Table 2: NCBI accession numbers of the reference genomes used for wgSNP typing.

ST	Strain name	NCBI accession number	Draft/complete genome
ST1	NTSN	NZ_CP009897.1	Complete
ST18	MOD1_LS152	NZ_CP020830.1	Complete
ST21	2KSM	JYOI00000000.1	Draft
ST184	L1223	(this study)	Draft

Table 3: Minimum and maximum pairwise allelic/SNP differences among the outbreak-associated *Listeria monocytogenes* isolates (intra-outbreak diversity).

Outbreak	MLST ST	No. of isolates	cgMLST (min–max)	wgMLST (min–max)	wgSNP (min–max)
Outbreak 1	ST1	2	0	0	1
Outbreak 2	ST1	2	0	2	5
Outbreak 3	ST1	2	0	2	5
Outbreak 4	ST18	3	1–3	3–4	3–4
Outbreak 5	ST21	2	2	5	6
Outbreak 6	ST21	2	7	8	13
Outbreak 7	ST184	11	0–4	0–12	3–6

Table 4: Minimum and maximum pairwise allelic/SNP differences among *Listeria monocytogenes* isolates from different outbreaks of the same sequence type (inter-outbreak diversity).

Outbreak	MLST ST	cgMLST (min–max)	wgMLST (min–max)	wgSNP (min–max)
Outbreaks 1–3	ST1	10–25	21–42	70–103
Outbreaks 5–6	ST21	8–9	9–14	12–19