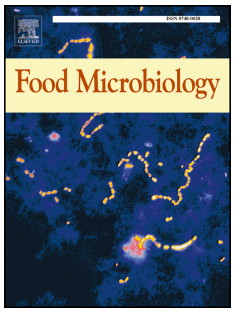

Attribution of Listeria monocytogenes human infections to food and animal sources in Northern Italy

Virginia Filipello, Lapo Mughini-Gras, Silvia Gallina, Nicoletta Vitale, Alessandro Mannelli, Mirella Pontello, Lucia Decastelli, Marc W. Allard, Eric W. Brown, Sara Lomonaco

PII: S0740-0020(20)30022-8

DOI: https://doi.org/10.1016/j.fm.2020.103433

Reference: YFMIC 103433

To appear in: *Food Microbiology*

Received Date: 18 June 2019

Revised Date: 16 December 2019

Accepted Date: 15 January 2020

Please cite this article as: Filipello, V., Mughini-Gras ,l.mughinigras@uu.nl, L., Gallina, S., Vitale, N., Mannelli, A., Pontello, M., Decastelli, L., Allard, M.W, Brown, E.W, Lomonaco, S., Attribution of Listeria monocytogenes human infections to food and animal sources in Northern Italy, *Food Microbiology*, https://doi.org/10.1016/j.fm.2020.103433.

Attribution of *Listeria monocytogenes* human infections to food and animal sources in Northern Italy

Virginia Filipello[a,b], Lapo Mughini-Gras[c,d], Silvia Gallina[e], Nicoletta Vitale[e], Alessandro Mannelli[a], Mirella Pontello[f], Lucia Decastelli[e], Marc W Allard[g], Eric W Brown[g], Sara Lomonaco[a,g]

[a] University of Turin. Largo P. Braccini, 2 – 10095 Grugliasco, Italy. Email: alessandro.mannelli@unito.it;

[b] Isituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna. Via A. Bianchi, 9 – 25124 Brescia, Italy. Email: virginia.filipello@izsler.it

[c]National Institute for Public Health and the Environment (RIVM), Center for Infectious Disease Control, Antonie van Leeuwenhoeklaan, 9 – 3721 MA Bilthoven, Netherlands. Email: lapo.mughini.gras@rivm.nl

[d] Utrecht University, Institute for Risk Assessment Sciences (IRAS), Yalelaan 2, 3584 CM, Utrecht, The Netherlands. Email: l.mughinigras@uu.nl

[e] Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle d'Aosta. Via Bologna, 148 – 10154 Torino, Italy. Email: silvia.gallina@izsto.it, nicoletta.vitale@izsto.it, lucia.decastelli@izsto.it

[f] Universiry of Milan. Via Rudinì, 3 – 20142 Milano, Italy. Email: mirella.pontello@unimi.it

[g] US Food & Drug Administration. 5001 Campus Drive – 20740 College Park, MD, USA. Email: marc.allard@fda.hhs.gov; eric.brown@fda.hhs.gov; sara.lomonaco@fda.hhs.gov

*Corresponding authors: Virginia Filipello, virginia.filipello@izsler.it; for WGS: Sara Lomonaco, sara.lomonaco@fda.hhs.gov

Abstract

Listeriosis is a foodborne illness characterized by a relatively low morbidity, but a large disease burden due to the severity of clinical manifestations and the high case fatality rate. Increased listeriosis notifications have been observed in Europe since the 2000s. However, the reasons for this increase are largely unknown, with the sources of sporadic human listerioris often remaining elusive. Here we inferred the relative contributions of several putative sources of *Listeria monocytogenes* strains from listerioris patients in Northern Italy (Piedmont and Lombardy regions), using two established source attribution models (i.e. 'Dutch' and 'STRUCTURE') in comparative fashion. We compared the Multi-Locus Sequence Typing and Multi-Virulence-Locus Sequence Typing profiles of strains collected from beef, dairy, fish, game, mixed foods, mixed meat, pork, and poultry. Overall, 634 *L. monocytogenes* isolates were collected from 2005 to 2016. In total, 40 clonal complexes and 51 virulence types were identified, with 36% of the isolates belonging to possible epidemic clones (i.e. genetically related strains from unrelated outbreaks). Source attribution analysis showed that 50% of human listerioris cases (95% Confidence Interval 44-55%) could be attributed to dairy products, followed by poultry and pork (15% each), and mixed foods (15%). Since the contamination of dairy, poultry and pork products are closely linked to primary production, expanding actions currently limited to ready-to-eat products to the reservoir level may help reducing the risk of cross-contamination at the consumer level.

Keywords

Listeria monocytogenes, listeriosis, food safety, epidemic clones, source attribution, molecular epidemiology

Introduction

Listeria monocytogenes is a bacterial foodborne pathogen that rarely causes severe disease in healthy individuals. Indeed, clinical listeriosis mainly occurs in at-risk groups: pregnant women, elderly people, immunocompromised people, unborn babies, and neonates (Lomonaco, Nucera, and Filipello 2015). In Europe, the incidence of listeriosis is approximately 0.48 per 100,000 inhabitants, and infections can occur either in a sporadic or epidemic form (EFSA and ECDC 2018). Several wild and domestic animals can also acquire infection with *L. monocytogenes*, particularly mammals and birds, which are also considered potential zoonotic reservoirs of the pathogen (Vivant, Garmyn, and Piveteau 2013). Among mammals, ruminants are the most susceptible to listeriosis, and *L. monocytogenes* subtypes associated with human listeriosis cases have been identified in bovine farms as well (Nightingale et al. 2004; Rocha et al. 2013). In birds, listeriosis mainly occurs sporadically, and it is believed that birds may act as a potential source for the infection in ruminants through the contamination of pastures and feed crops (Dhama et al. 2013; Locatelli et al. 2013). While exposure to infected animals and contaminated agricultural environments rarely appear to be directly linked to human infections, animal-derived food products that are consumed raw or undercooked, refrigerated RTE stored for long periods, as well as manure-contaminated fresh produce, often cause disease in humans (Nightingale et al. 2004; Lopez-Valladares, Danielsson-Tham, and Tham 2018). Moreover, unlike most foodborne pathogens, *L. monocytogenes* can grow in conditions of fairly low moisture, high salt concentration, and most importantly, at refrigeration temperatures, thereby conferring ability to persist and multiply in the food environment (Matthews, Kniel, and Montville 2017).

In case of human infection, the ubiquitous nature of *L. monocytogenes* and ability to survive for long periods outside the host, coupled with a relatively long incubation period, may

hamper the identification of the source (Dhama et al. 2015). Indeed, by the time of listeriosis diagnosis, food leftovers are very seldom available, and recalling the exact food consumption history preceding the infection may also be difficult (Amato et al. 2017; Jackson, Iwamoto, and Swerdlow 2010). Source attribution modelling based on microbial subtyping offers the opportunity to overcome these difficulties. Indeed, source attribution allows for the quantification of the relative contributions of the main animal, food, and environmental sources of foodborne disease, and attributions can be estimated at different points along the food chain, including production, distribution, and consumption (Pires et al. 2009).

Source attribution based on microbial subtyping relies on the characterisation of isolates using different phenotyping or genotyping methods (Andreoletti et al. 2008). Human cases are then probabilistically attributed to sources by comparing the subtype distributions of human source strains through mathematical models (Mughini-Gras and van Pelt 2014). Two main families of source attribution models are available: the so-called 'frequency matching' and 'population genetics' models, each with several advantages and disadvantages, as discussed in a recent opinion paper (Mughini-Gras et al. 2018). Overall, the source attribution approach has proven useful in prioritising and guiding control strategies, allowing for the identification of the most important reservoirs of specific pathogens (Boysen et al. 2014).

Multi-Locus Sequence Typing (MLST) and Multi-Virulence-Locus Sequence Typing (MVLST) are sequence-based methods in which Single Nucleotide Polymorphisms (SNPs) in fragments of a set of genes are used to determine allelic variants. MLST is based on a set of 7 housekeeping genes, while MVLST is based on a set of 6 virulence genes. MLST has been used to study and describe the population structure and phylogeny of *L. monocytogenes*, while MVLST has been used to identify Epidemic Clones (ECs) in outbreak investigations (Ragon et

al. 2008; Amato et al. 2017; Lomonaco et al. 2013; Chen, Zhang, and Knabel 2005; Knabel et al. 2012). An advantage of using allele-based methods is the presence of a shared nomenclature based on reference strains publicly available on dedicated databases (MLST, http://bigsdb.pasteur.fr/Listeria/Listeria.html; MVLST, https://sites.google.com/site/mvlstdatabase).

The aim of this study was to quantify the relative contributions of several putative sources of human listeriosis cases in Northern Italy by using two established source attribution modelling approaches based on MLST and MVLST data for clinical *L. monocytogenes* strains and strains from beef, dairy, fish, game, mixed foods, mixed meat, pork, and poultry. To further describe the strains circulating in the considered area the majority of the isolates were analysed with Whole Genome Sequencing (WGS), and screened for Antimicrobial Resistance (AMR) genes and SNP clustering through the NCBI Pathogen Detection pipeline.

Materials and Methods

Isolates collection

A total of 634 *L. monocytogenes* isolates were available for this study. These included 218 isolates from human listeriosis patients and 416 from various food sources, divided into 8 categories (i.e. beef, dairy, fish, game, mixed food, mixed meat, pork, and poultry). Clinical isolates were collected between 2005 and 2016 through a voluntary network of hospital laboratories in two Northern Italy regions, i.e. Lombardy and Piedmont (Mammina et al. 2013; Filipello et al. 2015). The food isolates were collected between 2004 and 2015 during the routine surveillance carried out by the Regional Animal Health and Food Safety Institutes (IZS) or in previous research projects aimed at studying the epidemiology of *L. monocytogenes* along the food chain carried out by the Department of Veterinary Sciences of the University of Turin.

Molecular typing

The whole genome sequences for 510 isolates, represented by food and environmental (*n*=416) and clinical isolates (*n*=94), were obtained at the Center for Food Safety and Applied Nutrition (CFSAN) of the US Food and Drug Administration (Lomonaco et al. 2018). DNA extraction was performed using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany), following manufacturer's instructions. DNA libraries were generated using the Illumina Nextera XT DNA Library Preparation Kit. WGS was performed on a MiSeq or a NextSeq system using a 2×250 bp or a 2x150 bp paired-end MiSeq/NextSeq Reagent Kit, respectively (Illumina, San Diego, CA, USA). MLST and MVLST data were extracted from the WGS data (Lomonaco et al. 2018). The remaining 124 clinical isolates were typed with MLST and MVLST as previously described (Chen, Zhang, and Knabel 2005; Ragon et al. 2008). Sequence Types (STs) and Virulence Types (VTs) were defined using the allelic sequences of the different loci schemes available in the respective online databases (MLST, https://bigsdb.pasteur.fr/listeria/listeria.html and MVLST, https://sites.google.com/site/mvlstdatabase/) and were used to assign isolates to Clonal Complexes (CCs) (i.e. groups of isolates with at least 6 alleles in common with another member of the same group) and to identify ECs. Both MLST and MVLST data were visualized using Minimum Spanning Trees (MSTs), generated by the PHYLOViZ software (Francisco et al. 2012).

WGS data for the strains described herein is also available on the NCBI Pathogen Detection database (NCBI PD, https://www.ncbi.nlm.nih.gov/pathogens/), a centralized system integrating WGS data for several bacterial pathogens obtained from different sources with the scope of rapidly linking food or environmental isolates to clinical isolates to discover potential sources of contamination and aid traceback/outbreak investigations. Single-linkage clustering (with SNP

distance of 50 SNPs) is used to identify closely related sets of isolates and assign SNP cluster accessions (i.e. PDS#). Individual phylogenetic trees are available for each SNP cluster, based on maximum compatibility (Cherry 2017). Isolates that are not within 50 SNPs of any other isolate are not assigned to a SNP cluster. The NCBI Pathogen Detection pipeline also provides data about the AMR genotype listing the antimicrobial resistance genes that have been identified by the NCBI AMR Finder process. As of April 1st, 2019, the NCBI PD database contains 26,567 *L. monocytogenes* isolates, and the isolates analysed herein can be found under BioProject ID PRJNA304956. Data on the NCBI PD is available for 508 of the 510 *L. monocytogenes* strains typed with WGS under BioProject PRJNA304956 (Lomonaco et al., 2018). Two strains (CFSAN045809 and CFSAN049182) were excluded from NCBI PD because their genome size was considered too small and outside the accepted ranges. Overall, 514 isolates are listed under BioProject PRJNA304956, with 6 strains (CFSAN044745, 044769, 046011, 046039, 046086, 049217) not included in the original publication (Lomonaco et al., 2018), and thus not considered herein.

Source attribution modelling

Human cases were attributed to the putative sources by applying two different models in parallel, the 'Dutch model' (Lapo Mughini-Gras, Franz, and van Pelt 2018) and 'STRUCTURE' (Pritchard, Stephens, and Donnelly 2000). The Dutch model is a simple frequency-matching model that compares the number of human cases caused by a specific subtype (i.e. ST or VT), with the relative occurrence of that subtype in each source. This model was applied separately on MLST and MVLST data, resulting in two model-data type combinations (MLST Dutch and MVLST Dutch). STRUCTURE is a population genetics, Bayesian clustering model that uses multi-locus genotype data to infer population structure and to assign individuals in a sample to

populations. This model was applied separately to MLST, MVLST, and coupled MLST+MVLST data (genotypic profiles defined by the combined 13 alleles), resulting in three model-data type combinations (MLST STRUCTURE, MVLST STRUCTURE, and MVLST+MLST STRUCTURE). For a more detailed description of the source attribution models, we refer to previous papers (Pritchard, Stephens, and Donnelly 2000; Lapo Mughini-Gras, Franz, and van Pelt 2018).

Statistical analysis

To assess differences in attributions over the different model-data type combinations (i.e. MLST Dutch, MVLST Dutch, MLST STRUCTURE, MVLST STRUCTURE, and MLST+MVLST STRUCTURE), the attributable proportions of cases were compared by exact two-tailed binomial test for each model-data type combination. To evaluate the agreement between attributions, a correlation matrix between the 5 model-data type combination was calculated using the Pearson correlation coefficient (rho). For each model-data type combination, the attributable proportions were ordered and ranked in ascending order. A median was calculated for each food category taking into account each value and the median of the ranks was used to provide an overall classification. All analyses were performed by open source software R (R Development Core Team).

Results

MLST typing

MLST results were available for 628 of the 634 isolates. MLST results were not available for six isolates (378, 379, 409, 598, 600, 609; S1). Among the typed isolates, 596 isolates belonged to 40 different CCs, and 32 isolates belonged to 9 singleton STs (not belonging to any

CC). The most significant group of clonal isolates was represented by ST9 (*n*=185 isolates, 29%), corresponding to 3 different VTs. (VT11, VT160, and VT162). In total 14 CCs accounted for 95% of the isolates (Figure 1; S1).

MVLST typing

MVLST results were available for all 634 isolates. In total, 51 different VTs were identified (S1), 17 isolates did not belong to any previously assigned VT and were therefore assigned to new VTs (VT160-VT168). Overall, VT11 represented the most abundant group of isolates (*n*=186, 29%), corresponding to ST9 (*n*=180) and ST204 (*n*=6). Overall, 36% (*n*=228) of the isolates belonged to 9 ECs (Table 1). In particular, ECs represented 22% (*n*=90) of the food chain isolates, and 64% (*n*=138) of the clinical isolates. The population structure of the isolates typed with MVLST and the proportion of the different sources identified for each VT are described in Figure 2.

W***GS analysis: antimicrobial resistance and SNP clusters***

Based on the NCBI Pathogen Detection browser, out of 508 isolates typed with WGS the *tet*(M) gene coding for resistance to tetracycline was found in 5.3% (*n*=27), while one isolate was listed with the *tet* gene. No presence of penicillin resistance genes was observed. Eighty-one isolates (*n*=22 clinical and *n*=59 food/environmental) were not assigned to any SNP clusters, while the remaining 427 isolates belonged to a total of 71 SNP clusters, as of April 1st, 2019 (Tables 2 and 3). About 32% (*n*=23) of the SNP clusters were “local”, comprising only isolates (*n*=73) from this study and not correlating with isolates from different countries/sources (Table 3). Of the 23 local SNP clusters, 16 only comprised food/environmental isolates (grouping from 2 to 8 isolates each), 6 only clinical isolates (grouping 2 or 3 isolates each), and 1 comprised both clinical and food/environmental isolates. The latter (PDS000006278.4) grouped 3 isolates within

11 SNPs, collected from a patient (blood) in 2014 and swabs from dairy plants collected in 2004 and 2014.

The remaining ~68% of SNP clusters (*n*=48) were "global", comprising 354 strains that were similar to other 3,179 isolates in the database (Table 2). Overall, among all SNP clusters detected herein, PDS000025311.40 was the largest, grouping a total of 517 isolates (246 clinical and 271 food/environmental/other). The most predominant cluster observed among our isolates was PDS000024241.19 (*n*=138), comprising ~75% of the 184 WGS-derived VT11 isolates, followed by PDS000001093.24 (*n*=35), PDS000024645.27 (*n*=22), and PDS000025311.40 (*n*=20). Isolates belonging to the most common detected profile (i.e. VT11) were distributed in 5 global SNP clusters: VT9/ST11 isolates (*n*=151, 82%) in PDS000024241.19, PDS000011669.6, PDS000025489.2, and PDS000024263.2; and all VT11/ST204 isolates (*n*=6, 3.2%) in PDS000024900.22. The remaining VT11 isolates were either in 5 local SNP clusters (*n*=20, ~11%) (Table 3) or unclustered (*n*=7, 3.8%). In our study, 10 out of the 24 isolates (~42%) from the production chain of Gorgonzola, a Protected Designation of Origin (PDO) blue cheese, are grouped into SNP cluster PDS000001093.24 (*n*=58), which also contains isolates from Gorgonzola, Taleggio, Blue Stilton and blue-veined and mold-ripened cheese isolates from the US and Italy.

Source attribution

All 5 combinations of models and type of data identified dairy products as the main source of human listeriosis cases (maximum attribution 53%, 95% Confidence Interval [95%CI] 46.96-58.42; Figure 3 and 4; S2). Even if the attributions varied, the different sources ranked similarly across the 5 model-data type combinations, with the exception of pork and poultry (Table 4). Specifically, in the Dutch model, pork appears to be the second most important source

(15% and 14% based on MLST and MVLST, respectively); while poultry appears to be more important in STRUCTURE, especially when using MVLST (18%, 95%CI 15.23-21.51; S2).

We observed high agreement among the 5 model-data type combinations (Table 5), with the lowest rho value (0.702, $p<0.0001$) observed between MVLST Dutch and MVLST STRUCTURE, and the highest rho value (0.997, $p<0.001$) between MLST STRUCTURE and MLST+MVLST STRUCTURE. High rho values were also observed between the STRUCTURE and Dutch models, with a rho value of 0.899 ($p<0.0001$) between MLST+MVLST STRUCTURE and MLST Dutch. The high agreement among the different model-data type combinations suggests a high goodness of fit. Increasing the number of loci in STRUCTURE by including 13 loci for MLST and MVLST together did not influence the source attribution results significantly (Figure 4).

Discussion

We characterized a large collection of *L. monocytogenes* isolates from human cases and different putative food sources in Northern Italy and identified the most likely sources of human listeriosis in that area. These results can support risk managers in prioritizing public health interventions. Source attribution using the microbial subtyping method is particularly important for listeriosis, as not all strains have the same ability to cause disease (Nightingale et al. 2008).

In our study, source attribution was performed using 2 models (Dutch and STRUCTURE) and 2 typing methods (MLST and MVLST), considering 8 different food sources. Moreover, WGS was performed to obtain typing data, AMR data, SNP clusters, and comparison with more than 26,000 isolates already present in the NCBI PD on-line databases. The screening of WGS data for AMR genes showed that ~5% (n=27) of the isolates carried the tetracycline-conferring resistance gene *tet*(M), a higher percentage than the 0.5% reported at the European level (Nielsen

et al. 2017). Among our isolates, ~89% (*n*=24) of *tet*(M) positive isolates belonged to ST9/VT11 isolates, that were overrepresented, possibly explaining the higher proportion. As also reported in other studies, *tet*(M) is the resistance gene most frequelntly detected in *L. monocytogenes* due to the transfer through mobile genetic elements from other resistant Gram-positive bacteria (Haubert et al. 2018). No isolates carried penicillin resistance genes, consistently with findings from the European report (Nielsen et al. 2017).

In total, 40 CCs and 51 VTs were identified, with CC9 being the most prevalent type and accounting for 43% of the food isolates and represented by all food sources (S1; Figure 2). On the *Listeria* MLST Pasteur database, CC9 isolates (*n*=223, 6% of all isolates in the database) originated from a wide variety of sources, including natural environment samples. None of the CC9 isolates with available information on the Pasteur database (*n*=12) carried the *tet*(M) gene. In our samples, CC9 mainly corresponded to VT11 and its Single Locus Variants (SLV – isolates with *n*-1 alleles in common to the linked node; VT160 and VT162 in Figure 2). ST9/VT11 had been previously identified as the most predominant and persistent type also in a study that investigated the presence of *L. monocytogenes* in meat processing plant in Spain (Martín et al. 2014), and in a study carried out in a mushroom processing plant in the US (Murugesan et al. 2015). Despite such a broad diffusion, it seems that ST9/VT11 isolates have a minor role in causing clinical cases, as only 5 human clinical strains belongend to this genotype (2.3% of cases; S1), and thus may be more adapted to survive in the environment. Indeed, CC9 has been observed as significantly associated with food and food environment and with a particularly high prevalence of truncated InlA variants, which are associated with hypovirulence (Moura et al. 2017; Nightingale et al. 2008). The main cluster of clinical cases are instead represented by CC101 (*n*=50, 23%) and CC1 (*n*=31, 14.2%). In particular, CC101 is the major cluster of clinical cases, which

had been previously singled out in a 2014 study, where it stood out among different CCs for being the only one with a clear predominance of human isolates (Haase et al. 2014). A novel EC associated with CC101, i.e. ECXI, was recently recognized as involved in two unrelated outbreaks linked to the consumption of Ricotta salata (USA, 2012) and Taleggio cheese (Italy, 2011), both produced in Italy (Amato et al. 2017).

L. monocytogenes types found in foods and clinical isolates only partially overlap (Figures 1 and 2), strengthening the evidence that not all *L. monocytogenes* strains are equally capable of causing invasive disease. Overall, several studies have shown that lineage I *L. monocytogenes* strains are on average more virulent and more frequently associated with human clinical cases than lineage II strains (Lomonaco, Nucera, Filipello, 2015; Pirone-Davies et al., 2018). Such partial overlap was also observed in the local SNP clusters, with the majority (n=16, 69.5%) only grouping food/environmental isolates, followed by 26% comprising just clinical isolates and only 4.3% currently containing both. Among the 81 isolates not currently included in a SNP cluster, more than a half (n=45, 55.6%) were from food and food production enviroments, while the rest was from clinical cases (n=22, 27%) or associated with agriculture (i.e. stools and feeds, n=14, 17.2%). Additionally, a recent study showed that a significant proportion of *L. monocytogenes* isolated from food production environments have reduced virulence (Van Stelten et al. 2016). In light of these data, considering that current regulations in EU and US are based on the sole detection of *L. monocytogenes*, it could be useful and more sustainable (e.g. given the high economic impact due to recalls) to review a risk assessment process that incorporates strain-specific virulence parameters, meaning the identification of virulence genes and their variants that may be applied as markers either for disease-relevant strains or non-virulent strains (Walland et al. 2015). For instance, internalin A and its truncated variants have often been identified

as possible marker for reduced virulence (Van Stelten et al. 2016). Nevertheless, to date straightforward identification of such markers are still lacking, and inconsistent evidences have been reported (Ferreira da Silva et al. 2017).

The different model-data type combinations used in the source attribution analysis identifed dairy products as the main source of human listeriosis (28% to 53%) (Figures 3 and 4, S2). Indeed, in Europe half of the reported outbreaks have been linked to dairy products (Lundén, Tolvanen, and Korkeala 2004). In the Dutch model, pork appeared to be the second source of listeriosis (Figure 3). This may be explained by the overrepresentation of pork isolates over the other sources among the food isolates (28%; S1). This may influence the output, as the Dutch model is a frequency matching based model. On the other hand, poultry appears to be a more important source when using STRUCTURE, particularly with MVLST data (18%; Figure 4; S2). The poultry category comprises both raw meat and cooked preparations and its impact in the Dutch model may have been overshadowed due to the low number of isolates (n=13; S1). Given this, STRUCTURE seems to be more reliable than the Dutch model in overcoming representativeness issues.

Because *L. monocytogenes* is highly susceptible to thermic treatment (i.e. cooking), source attribution of the listeriosis cases is usually carried out only on ready-to-eat (RTE) products (Little et al. 2010; Nielsen et al. 2017), as opposed to diseases like salmonellosis and campylobacteriosis that are studied also at the reservoir level (Pires et al. 2009; Boysen et al. 2014; Lapo Mughini-Gras et al. 2018). Isolates collected at the reservoir level (i.e. non-RTE) were also included in this study and possible associations were found, in particular with poultry (Figure 4, S2). This finding underlines how controlling contamination at the reservoir level could be useful, in terms of preventing cross-contamination that may occur both at the distribution (e.g. deli counters) and

278 at the household level. Indeed, it is still poorly understood how *L. monocytogenes* circulates be-
279 tween animals, humans, and various environments (Walland et al. 2015). In particular, it has
280 been found that bovine farm environments have high prevalence rates of *L. monocytogenes*, in-
281 cluding subtypes linked to human listeriosis cases and outbreaks, and cattle appear to contribute
282 to the amplification and spread of *L. monocytogenes* in the farm environment (Nightingale et al.
283 2004). In Italy, Rocha et al. found 60% and 10% of *L. monocytogenes* isolated from bovine clini-
284 cal cases belonging to ECI and ECX, respectively (Rocha et al. 2013). Poultry is also a recog-
285 nized reservoir of *L. monocytogenes* and contaminated raw meat poses a concrete risk for the
286 human consumer (Dhama et al. 2013). In the US, several ECs were found in chicken processing
287 plants and listeriosis cases and outbreaks have been associated with consumption of undercooked
288 chicken and RTE poultry products (Lomonaco et al. 2013). Moreover, it is not clear whether
289 only specific *L. monocytogenes* subtypes are able to move from the reservoir to the hosts and
290 cause disease (Walland et al. 2015). Consequently, to improve our understanding of the ecology
291 of *L. monocytogenes*, it is important to study the prevalence of *L. monocytogenes* strains in all
292 different niches, such as the farm environment, livestock, raw materials, transport vehicles and
293 containers, manufacturing facilities (e.g. cheese plants) and humans. A recent study identified
294 eight genes significantly associated with food isolates across *L. monocytogenes* lineage II strains,
295 likely playing an important role in the survival and proliferation of *L. monocytogenes* in the food
296 environment. The authors indicated the need for futher studies on such genes as such knowledge
297 can help understand how *L. monocytogenes* adapts to the host and food environments (Pirone-
298 Davies et al., 2018).

299 Most other published source attribution studies (mainly on *Salmonella* and *Campylobac-*
300 *ter*) tend to have higher numbers of isolates (Kittl et al. 2013; de Knegt et al. 2016; Mughini-

Gras et al. 2014; Boysen et al. 2014), and it has been reported that is preferable to have at least 100 isolates for each source analysed (Smid et al. 2013). Moreover, selection of isolates should include contemporaneous sampling of isolates from sources and humans from a fixed geographic area. In the current study, samples were collected over a fairly broad timeframe (13-year period, 2004-2016). While broad, such a timeframe was necessary to ensure that the strain collection was as representative as possible within the scope of the study, given the low incidence of listeriosis.

Conclusion

Dairy products were identified as the most important source of human listeriosis in the study area, highlighting the need for specific control measures to reduce *L. monocytogenes* contamination in these products. To date, mainly RTE products have been included in source attribution studies of listeriosis. According to our results, implementing actions currently limited to RTE products also at the reservoir level, may help reducing the risk of cross-contamination at the distribution and household levels.

Considering the scarcity of data suited for source attribution of listeriosis, especially in Italy, this study represents a first stepping-stone for future research. Indeed, this is the first source attribution study for listeriosis in Italy, and its routine application may help mitigating the impact of the disease, both at a national and international level, by targeting the main sources. To reach this goal, collaboration between the different competent authorites in a One Health perspective is of paramount importance.

Declarations of conflict of interest: none.

Funding sources: Fondo Ricerca Locale 2014 and World Wide Style second edition (WWS2) from the Università degli Studi di Torino. VF was also supported with a Ph.D. scholarship by the Homo Sapiens Sapiens Scholarships from INPS. This project was also supported in part by an appointment to the Research Participation Program at the CFSAN, U.S. Food and Drug Administration, administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and FDA. The funding sources had no involvement in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

The findings and conclusions in this report are those of the authors and do not necessarily represent the views or official positions of the Food and Drug Administration. The use of trade names is for identification only and does not imply endorsement.

References

Amato, Ettore, Virginia Filipello, Maria Gori, Sara Lomonaco, Marina Nadia Losio, Antonio Parisi, Pol Huedo, Stephen John Knabel, and Mirella Pontello. 2017. "Identification of a Major Listeria Monocytogenes Outbreak Clone Linked to Soft Cheese in Northern Italy - 2009-2011." *BMC Infectious Diseases* 17 (1): 342. https://doi.org/10.1186/s12879-017-2441-6.

Andreoletti, Olivier, Herbert Budka, Sava Buncic, Pierre Colin, John D. Collins, Aline De, John Griffin Koeijer, et al. 2008. "Overview of Methods for Source Attribution for Human Illness from Food-Borne Microbiological Hazards Scientific Opinion of the Panel on Biological Hazards." http://onlinelibrary.wiley.com/doi/10.2903/j.efsa.2008.764/full.

Boysen, L., H. Rosenquist, J. T. Larsson, E. M. Nielsen, G. Sørensen, S. Nordentoft, and T. Hald. 2014. "Source Attribution of Human Campylobacteriosis in Denmark." *Epidemiology and Infection* 142 (8): 1599–1608. https://doi.org/10.1017/S0950268813002719.

Chen, Yi, Wei Zhang, and Stephen J. Knabel. 2005. "Multi-Virulence-Locus Sequence Typing Clarifies Epidemiology of Recent Listeriosis Outbreaks in the United States." *Journal of Clinical Microbiology* 43 (10): 5291–94. https://doi.org/10.1128/JCM.43.10.5291-5294.2005.

Cherry, Joshua L. 2017. "A Practical Exact Maximum Compatibility Algorithm for Reconstruction of Recent Evolutionary History." *BMC Bioinformatics* 18 (1): 127. https://doi.org/10.1186/s12859-017-1520-4.

Dhama, Kuldeep, Kumaragurubaran Karthik, Ruchi Tiwari, Muhammad Zubair Shabbir, Sukhadeo Barbuddhe, Satya Veer Singh Malik, and Raj Kumar Singh. 2015. "Listeriosis in Animals, Its Public Health Significance (Food-Borne Zoonosis) and Advances in Diagnosis and Control: A Comprehensive Review." *Veterinary Quarterly* 35 (4): 211–35. https://doi.org/10.1080/01652176.2015.1063023.

Dhama, Kuldeep, Amit Kumar Verma, S. Rajagunalan, Amit Kumar, Ruchi Tiwari, Sandip Chakraborty, and Rajesh Kumar. 2013. "Listeria Monocytogenes Infection in Poultry and Its Public Health Importance with Special Reference to Food Borne Zoonoses." *Pakistan Journal of Biological Sciences: PJBS* 16 (7): 301–8.

Ferreira da Silva, Margarida, Vânia Ferreira, Rui Magalhães, Gonçalo Almeida, Artur Alves, and Paula Teixeira. 2017. "Detection of Premature Stop Codons Leading to Truncated Internalin A among Food and Clinical Strains of Listeria Monocytogenes." *Food Microbiology* 63 (May): 6–11. https://doi.org/10.1016/j.fm.2016.10.033.

Francisco, Alexandre P., Cátia Vaz, Pedro T. Monteiro, José Melo-Cristino, Mário Ramirez, and João A. Carriço. 2012. "PHYLOViZ: Phylogenetic Inference and Data Visualization for Sequence Based Typing Methods." *BMC Bioinformatics* 13 (1): 87. https://doi.org/10.1186/1471-2105-13-87.

Haase, Jana K., Xavier Didelot, Marc Lecuit, Hannu Korkeala, L. monocytogenes MLST Study Group, and Mark Achtman. 2014. "The Ubiquitous Nature of Listeria Monocytogenes Clones: A Large-Scale Multilocus Sequence Typing Study." *Environmental Microbiology* 16 (2): 405–16. https://doi.org/10.1111/1462-2920.12342.

Haubert, Louise, Carlos Eduardo Pouey da Cunha, Graciela Völz Lopes, and Wladimir Padilha da Silva. 2018. "Food Isolate Listeria Monocytogenes Harboring TetM Gene Plasmid-Mediated Exchangeable to Enterococcus Faecalis on the Surface of Processed Cheese."

Food Research International 107 (May): 503–8. https://doi.org/10.1016/j.foodres.2018.02.062.

Jackson, K. A., M. Iwamoto, and D. Swerdlow. 2010. "Pregnancy-Associated Listeriosis." *Epidemiology and Infection* 138 (10): 1503–9. https://doi.org/10.1017/S0950268810000294.

Kittl, Sonja, Gerald Heckel, Bożena M. Korczak, and Peter Kuhnert. 2013. "Source Attribution of Human Campylobacter Isolates by MLST and Fla-Typing and Association of Genotypes with Quinolone Resistance." *PloS One* 8 (11): e81796. https://doi.org/10.1371/journal.pone.0081796.

Knabel, Stephen J., Aleisha Reimer, Bindhu Verghese, Mei Lok, Jennifer Ziegler, Jeffrey Farber, Franco Pagotto, et al. 2012. "Sequence Typing Confirms That a Predominant Listeria Monocytogenes Clone Caused Human Listeriosis Cases and Outbreaks in Canada from 1988 to 2010." *Journal of Clinical Microbiology* 50 (5): 1748–51. https://doi.org/10.1128/JCM.06185-11.

Knegt, Leonardo V. de, Sara M. Pires, Charlotta Löfström, Gitte Sørensen, Karl Pedersen, Mia Torpdahl, Eva M. Nielsen, and Tine Hald. 2016. "Application of Molecular Typing Results in Source Attribution Models: The Case of Multiple Locus Variable Number Tandem Repeat Analysis (MLVA) of Salmonella Isolates Obtained from Integrated Surveillance in Denmark." *Risk Analysis* 36 (3): 571–88. https://doi.org/10.1111/risa.12483.

Little, Christine L., Sara M. Pires, Iain A. Gillespie, Kathie Grant, and Gordon L. Nichols. 2010. "Attribution of Human Listeria Monocytogenes Infections in England and Wales to Ready-to-Eat Food Sources Placed on the Market: Adaptation of the Hald Salmonella Source Attribution Model." *Foodborne Pathogens and Disease* 7 (7): 749–56. https://doi.org/10.1089/fpd.2009.0439.

Locatelli, Aude, Géraldine Depret, Claudy Jolivet, Sonia Henry, Samuel Dequiedt, Pascal Piveteau, and Alain Hartmann. 2013. "Nation-Wide Study of the Occurrence of Listeria Monocytogenes in French Soils Using Culture-Based and Molecular Detection Methods." *Journal of Microbiological Methods* 93 (3): 242–50. https://doi.org/10.1016/j.mimet.2013.03.017.

Lomonaco, Sara, Silvia Gallina, Virginia Filipello, Maria Sanchez Leon, George John Kastanis, Marc Allard, Eric Brown, Ettore Amato, Mirella Pontello, and Lucia Decastelli. 2018. "Draft Genome Sequences of 510 Listeria Monocytogenes Strains from Food Isolates and Human Listeriosis Cases from Northern Italy." *Genome Announcements* 6 (3). https://doi.org/10.1128/genomeA.01276-17.

Lomonaco, Sara, Daniele Nucera, and Virginia Filipello. 2015. "The Evolution and Epidemiology of Listeria Monocytogenes in Europe and the United States." *Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases* 35 (October): 172–83. https://doi.org/10.1016/j.meegid.2015.08.008.

Lomonaco, Sara, Bindhu Verghese, Peter Gerner-Smidt, Cheryl Tarr, Lori Gladney, Lavin Joseph, Lee Katz, et al. 2013. "Novel Epidemic Clones of Listeria Monocytogenes, United States, 2011." *Emerging Infectious Diseases* 19 (1): 147–50. https://doi.org/10.3201/eid1901.121167.

Lopez-Valladares, Gloria, Marie-Louise Danielsson-Tham, and Wilhelm Tham. 2018. "Implicated Food Products for Listeriosis and Changes in Serovars of Listeria Monocytogenes Affecting Humans in Recent Decades." *Foodborne Pathogens and Disease* 15 (7): 387–97. https://doi.org/10.1089/fpd.2017.2419.

Lundén, J., R. Tolvanen, and H. Korkeala. 2004. “Human Listeriosis Outbreaks Linked to Dairy Products in Europe.” *Journal of Dairy Science*, Electronic Supplement, 87 (July): E6–12. https://doi.org/10.3168/jds.S0022-0302(04)70056-9.

Mammina, Caterina, Antonio Parisi, Anna Guaita, Aurora Aleo, Celestino Bonura, Antonino Nastasi, and Mirella Pontello. 2013. “Enhanced Surveillance of Invasive Listeriosis in the Lombardy Region, Italy, in the Years 2006-2010 Reveals Major Clones and an Increase in Serotype 1/2a.” *BMC Infectious Diseases* 13 (March): 152. https://doi.org/10.1186/1471-2334-13-152.

Martín, Belén, Adriana Perich, Diego Gómez, Javier Yangüela, Alicia Rodríguez, Margarita Garriga, and Teresa Aymerich. 2014. “Diversity and Distribution of Listeria Monocytogenes in Meat Processing Plants.” *Food Microbiology* 44 (December): 119–27. https://doi.org/10.1016/j.fm.2014.05.014.

Matthews, Karl R., Kalmia E. Kniel, and Thomas J. Montville. 2017. *Food Microbiology: An Introduction, Fourth Edition*. American Society of Microbiology. https://doi.org/10.1128/9781555819392.

Moura, Alexandra, Alexis Criscuolo, Hannes Pouseele, Mylène M. Maury, Alexandre Leclercq, Cheryl Tarr, Jonas T. Björkman, et al. 2017. “Whole Genome-Based Population Biology and Epidemiological Surveillance of *Listeria Monocytogenes*.” *Nature Microbiology* 2 (2): 16185. https://doi.org/10.1038/nmicrobiol.2016.185.

Mughini-Gras, L., F. Barrucci, J. H. Smid, C. Graziani, I. Luzzi, A. Ricci, L. Barco, et al. 2014. “Attribution of Human Salmonella Infections to Animal and Food Sources in Italy (2002–2010): Adaptations of the Dutch and Modified Hald Source Attribution Models.” *Epidemiology & Infection* 142 (5): 1070–82. https://doi.org/10.1017/S0950268813001829.

Mughini-Gras, Lapo, Eelco Franz, and Wilfrid van Pelt. 2018. “New Paradigms for Salmonella Source Attribution Based on Microbial Subtyping.” *Food Microbiology*, International Symposium on Salmonella and Salmonellosis 2016, 71 (May): 60–67. https://doi.org/10.1016/j.fm.2017.03.002.

Mughini-Gras, Lapo, Pauline Kooh, Jean-Christophe Augustin, Julie David, Philippe Fravalo, Laurent Guillier, Nathalie Jourdan-Da-Silva, et al. 2018. “Source Attribution of Foodborne Diseases: Potentialities, Hurdles, and Future Expectations.” *Frontiers in Microbiology* 9. https://doi.org/10.3389/fmicb.2018.01983.

Mughini-Gras, Lapo, and Wilfrid van Pelt. 2014. “Salmonella Source Attribution Based on Microbial Subtyping: Does Including Data on Food Consumption Matter?” *International Journal of Food Microbiology* 191 (November): 109–15. https://doi.org/10.1016/j.ijfoodmicro.2014.09.010.

Murugesan, Latha, Zuzana Kucerova, Stephen J. Knabel, and Luke F. LaBorde. 2015. “Predominance and Distribution of a Persistent Listeria Monocytogenes Clone in a Commercial Fresh Mushroom Processing Environment.” *Journal of Food Protection* 78 (11): 1988–98. https://doi.org/10.4315/0362-028X.JFP-15-195.

Nielsen, Eva Møller, Jonas T. Björkman, Kristoffer Kiil, Kathie Grant, Tim Dallman, Anaïs Painset, Corinne Amar, et al. 2017a. “Closing Gaps for Performing a Risk Assessment on Listeria Monocytogenes in Ready-to-Eat (RTE) Foods: Activity 3, the Comparison of Isolates from Different Compartments along the Food Chain, and from Humans Using Whole Genome Sequencing (WGS) Analysis.” *EFSA Supporting Publications* 14 (2): 1151E. https://doi.org/10.2903/sp.efsa.2017.EN-1151.

Nightingale, K. K., R. A. Ivy, A. J. Ho, E. D. Fortes, B. L. Njaa, R. M. Peters, and M. Wiedmann. 2008. "InlA Premature Stop Codons Are Common among Listeria Monocytogenes Isolates from Foods and Yield Virulence-Attenuated Strains That Confer Protection against Fully Virulent Strains." *Applied and Environmental Microbiology* 74 (21): 6570–83. https://doi.org/10.1128/AEM.00997-08.

Nightingale, K. K., Y. H. Schukken, C. R. Nightingale, E. D. Fortes, A. J. Ho, Z. Her, Y. T. Grohn, P. L. McDonough, and M. Wiedmann. 2004. "Ecology and Transmission of Listeria Monocytogenes Infecting Ruminants and in the Farm Environment." *Applied and Environmental Microbiology* 70 (8): 4458–67. https://doi.org/10.1128/AEM.70.8.4458-4467.2004.

Pires, Sara M., Eric G. Evers, Wilfrid van Pelt, Tracy Ayers, Elaine Scallan, Frederick J. Angulo, Arie Havelaar, Tine Hald, and Med-Vet-Net Workpackage 28 Working Group. 2009. "Attributing the Human Disease Burden of Foodborne Infections to Specific Sources." *Foodborne Pathogens and Disease* 6 (4): 417–24. https://doi.org/10.1089/fpd.2008.0208.

Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly. 2000. "Inference of Population Structure Using Multilocus Genotype Data." *Genetics* 155 (2): 945–59.

Ragon, Marie, Thierry Wirth, Florian Hollandt, Rachel Lavenir, Marc Lecuit, Alban Le Monnier, and Sylvain Brisse. 2008. "A New Perspective on Listeria Monocytogenes Evolution." *PLoS Pathogens* 4 (9): e1000146. https://doi.org/10.1371/journal.ppat.1000146.

Rocha, Paulo Ricardo Dell'Armelina, Sara Lomonaco, Maria Teresa Bottero, Alessandra Dalmasso, Alessandro Dondo, Carla Grattarola, Fabio Zuccon, et al. 2013. "Ruminant Rhombencephalitis-Associated Listeria Monocytogenes Strains Constitute a Genetically Homogeneous Group Related to Human Outbreak Strains." *Applied and Environmental Microbiology* 79 (9): 3059–66. https://doi.org/10.1128/AEM.00219-13.

Smid, Joost H., Lapo Mughini Gras, Albert G. de Boer, Nigel P. French, Arie H. Havelaar, Jaap A. Wagenaar, and Wilfrid van Pelt. 2013. "Practicalities of Using Non-Local or Non-Recent Multilocus Sequence Typing Data for Source Attribution in Space and Time of Human Campylobacteriosis." *PLOS ONE* 8 (2): e55029. https://doi.org/10.1371/journal.pone.0055029.

Van Stelten, A., A. R. Roberts, C. S. Manuel, and K. K. Nightingale. 2016. "Listeria Monocytogenes Isolates Carrying Virulence-Attenuating Mutations in Internalin A Are Commonly Isolated from Ready-to-Eat Food Processing Plant and Retail Environments." *Journal of Food Protection* 79 (10): 1733–40. https://doi.org/10.4315/0362-028X.JFP-16-145.

Vivant, Anne-Laure, Dominique Garmyn, and Pascal Piveteau. 2013. "Listeria Monocytogenes, a down-to-Earth Pathogen." *Frontiers in Cellular and Infection Microbiology* 3: 87. https://doi.org/10.3389/fcimb.2013.00087.

Walland, J., J. Lauper, J. Frey, R. Imhof, R. Stephan, T. Seuberlich, and A. Oevermann. 2015. "Listeria Monocytogenes Infection in Ruminants: Is There a Link to the Environment, Food and Human Health? A Review." *Schweizer Archiv Fur Tierheilkunde* 157 (6): 319–28. https://doi.org/10.17236/sat00022.

514 **Figure 1** Minimum spanning tree of the 628 *Listeria monocytogenes* isolates typed with MLST.
515 Each circle represents a single Sequence Type (ST) indicated on the tree by the corresponding
516 number. Yellow nodes are group founders and black lines indicate Single Locus Variants (SLV –
517 isolates with *n*-1 alleles in common to the linked node). For each ST, isolates obtained from dif-
518 ferent sources are represented by the colours in the legend. The number and proportion of iso-
519 lates for each source are listed in brackets in the legend.

520 **Figure 2** Minimum spanning tree of the 634 *Listeria monocytogenes* isolates typed with
521 MVLST. Each circle represents a single Virulence Type (VT) indicated on the tree by the corre-
522 sponding number. Yellow nodes are group founders and black lines indicate Single Locus Vari-
523 ants (SLV – isolates with *n*-1 alleles in common to the linked node). For each VT, the colours
524 listed in the legend represent the proportion of isolates from the different sources. Grey slices
525 indicate isolates not assigned to any of the listed sources. The number and proportion of isolates
526 for each source are listed in brackets in the legend.

527 **Figure 3** Source attributions of listeriosis human cases with MVLST and MLST data using the
528 Dutch model (error bars denote 95% confidence intervals). Unknown bar represents clinical cas-
529 es caused by *Listeria monocytogenes* types not found in any source.

530 **Figure 4** Source attributions of listeriosis human cases with MVLST, MLST and
531 MVLST+MLST data using the STRUCTURE model (error bars denote 95% confidence inter-
532 vals).

Table 1. Number of *L. monocytogenes* isolates belonging to each of the currently identified Epidemic Clones (ECs), among the all the strains collected from clinical cases and 8 different food sources.

	Epidemic Clones (ECs)									
Source	**I**	**II**	**IV**	**V**	**VI**	**VII**	**VIII**	**X**	**XI**	**Total**
Human	30	6	8	17		15	10	2	50	**138**
Beef				1						**1**
Dairy	13	7	4	1	1	4	2	1	8	**41**
Fish				2		1				**3**
Game	2						3			**5**
Mixed food			4	2	2	1	1			**10**
Mixed Meat	1		1	2						**4**
Pork	1		6	7	3		1	1		**19**
Poultry			1	3						**4**
Unknown		1		1			1			**3**
Total	**47**	**14**	**24**	**36**	**6**	**21**	**18**	**4**	**58**	**228**

Table 2. List of the 48 "global" SNP cluster, comprising 354 isolates from this study and correlating with 3,178 isolates from different countries/sources avilable on the NCBI PD database (as of April 1st, 2019). The number of environmental/food/other and clinical isolates, is indicated as those originating from this study over the overall number (i.e. #/#). Bold font was used to highlight SNP clusters grouping only isolates from Italy. SNP clusters are determined by the NCBI Pathogen Isolates pipeline and several information are listed for each: Virulence Type (VT), Epidemic clone (EC), Sequence Type (ST), accession number and analysed version, overall number of isolates and specific from this study, and overall number of environmental/food/other and clinical isolates.

						Number of isolates (from this study/overall)		
Sequence Type (ST)	**Clonal Complex (CC)**	**Virulence Type (VT)**	**Epidemic Clone (EC)**	**SNP Cluster Accession ID and Version (as of April 1st, 2019)**		**Total in SNP cluster**	**Environ./ food/other**	**Clinical**
ST1	CC1	VT20	ECI	PDS000003341	.13	2/4	0/0	2/4
				PDS000003348	.26	1/18	1/6	0/12
				PDS000006160	.21	8/9	4/4	4/5
				PDS000041947	.5	1/105	0/25	1/80
ST2	CC2	VT21	ECIV	PDS000024430	.11	9/107	7/42	2/65
				PDS000024474	.2	1/3	0/0	1/3
				PDS000024705	.8	3/30	3/24	0/6
ST3	CC3	VT14	ECVIII	PDS000006340	.10	3/5	1/3	2/2
				PDS000007098	.4	2/4	0/1	2/3
				PDS000009528	.3	1/2	0/0	1/2
				PDS000009530	.3	1/2	0/1	1/1
ST5	CC5	VT63	ECVI	PDS000032961	.1	1/2	1/2	0/0
ST6	CC6	VT19	ECII	PDS000024682	.26	1/273	0/73	1/200
				PDS000024688	.2	2/4	0/0	2/4
				PDS000043734	.1	1/2	1/2	0/0
				PDS000024930	.2	1/5	1/1	0/4
		VT163		PDS000024684	.9	9/53	5/14	0/39
							3/14	1/39
ST7	CC7	VT56	ECVIII	PDS000024618	.8	4/38	4/16	0/22

ST8	CC8	VT59	ECV	**PDS000003019**	.6	1/3	1/3	0/0
				PDS000025311	.40	20/517	17/271	3/246
ST9	CC9	VT11	*	PDS000024241	.19	138/324	136/297	2/27
				PDS000011669	.6	6/9	6/9	0/0
				PDS000025489	.2	4/6	4/6	0/0
				PDS000024263	.2	3/4	3/3	0/1
ST204	CC204			PDS000024900	.26	6/199	6/172	0/27
ST18	CC18	VT118		PDS000025244	.1	2/4	0/1	2/3
ST19	CC19	VT84		PDS000006154	.4	1/14	1/3	0/11
ST29	CC29	VT74		PDS000024749	.4	6/9	1/2	5/7
				PDS000024751	.2	1/3	1/2	0/1
ST32	CC32	VT93		PDS000037504	.2	1/6	1/1	0/5
ST388	CC388			PDS000025477	.5	1/10	1/2	0/8
ST37	CC37	VT61		PDS000032941	.18	4/174	1/111	3/63
ST38	CC101	VT80	ECXI	PDS000001213	.20	10/31	8/15	2/16
ST101				PDS000024823	.11	1/74	0/55	1/19
ST59	CC59	VT119		PDS000011242	.8	1/15	1/8	0/7
ST121	CC121	VT94		PDS000024645	.27	22/430	22/403	0/27
				PDS000024656	.28	7/457	4/424	1/33
		VT109					2/424	0/33
ST155	CC155	VT45		PDS000005514	.13	9/27	0/5	9/22
				PDS000006382	.27	1/128	1/102	0/26
ST217	CC217	VT62		PDS000024967	.21	2/128	2/20	0/108
ST224	CC224	VT124		PDS000009525	.4	1/3	0/2	1/1
ST296	CC88	VT8		PDS000003204	.81	1/128	1/104	0/24
ST325	CC31	VT113		PDS000001093	.24	35/58	30/53	5/5
ST394	CC415	VT2		PDS000009385	.6	1/10	0/9	1/1
ST398	CC398	VT100		PDS000024700	.1	13/14	12/13	1/1
ST425	CC90	VT151		PDS000042587	.1	1/6	0/0	1/6
ST451	CC451	VT140		PDS000024708	.17	1/69	0/29	1/40
ST562	CC562	VT166		PDS000004800	.42	3/7	3/6	0/1
TOTAL						354/3533	269/2345	58/1188

* includes 21 strains carrying *tet*(M) (overall this SNP cluster includes two more *tet*(M)-carrying strains from Italy, which were not included in Lomonaco et al., 2018)
# includes 3 strains carrying *tet*(M)
$ includes 1 strain carrying *tet*

Table 3. List of the 20 "local" SNP cluster, comprising isolates (*n*=73) correlating only with other Italian isolates originating from the current study (as of April 1st, 2019). SNP clusters are determined by the NCBI Pathogen Isolates pipeline and several information are listed for each: Sequence Type (ST), Clonal Complex (CC), Virulence Type (VT), Epidemic clone (EC), accession number and analysed version, overall number of isolates and specific from this study, and overall number of environmental/food/other and clinical isolates. The SNP clusters are divided into three groups, those only grouping environmental/food/other isolates, those grouping only clinical and those grouping both. Bold font was used to highlight the same VT/ST observed in different groups, while * was used to indicate isolates carrying the *tet*(M) gene.

Type of isolates grouped	Sequence Type (ST)	Clonal Complex (CC)	Virulence Type (VT)	Epidemic Clone (EC)	SNP Cluster Accession ID	Version (as of April 1st, 2019)	# of env./food/other isolates	# of Clinical isolates
Only environmental /food / other isolates	**ST1**	**CC1**	**VT20**	**ECI**	PDS000016512	.1	2	0
					PDS000016511	.1	5	0
					PDS000006159	.3	3	0
	ST2	CC2	VT21	ECIV	PDS000005749	.4	3	0
	ST3	**CC3**	**VT14**	**ECVIII**	PDS000009529	.3	4	0
	ST5	**CC5**	**VT63**	**ECVI**	PDS000016519	.1	3	0
	ST9	CC9	VT11	*	PDS000006163	.4	8	0
			VT11		PDS000024252	.1	5	0
			VT162		PDS000024740	.1	4	0
			VT11		PDS000024741	.1	3	0
			VT11		PDS000025500	.1	2	0
			VT160			.1	1	0
			VT11		PDS000024296	.1	2	0
	ST36	**CC36**	**VT75**		PDS000024703	.1	3	0
	ST427	CC29	VT74		PDS000006155	.5	5	0
	ST663	ST663	VT62		PDS000024699	.1	2	0
					PDS000024702	.1	2	0
Only clinical isolates	**ST1**	**CC1**	**VT20**	**ECI**	PDS000024707	.1	0	2
	ST5	**CC5**	**VT63**	**ECVI**	PDS000016343	.1	0	3
	ST7	CC7	VT56	ECVII	PDS000016346	.1	0	2
	ST14	CC14	VT125		PDS000016335	.1	0	2
	ST54	CC54	VT79		PDS000016380	.1	0	2
	ST398	CC398	VT100		PDS000024922	.1	0	2
Both env./food/other and clinical isolates	**ST3**	**CC3**	**VT14**	**ECVIII**	PDS000006278	.4	2	1
TOTAL							59	14

* isolates carrying the *tet*(M) gene

Table 4. Median of ranks and the ranks (in descending order) for each of the 8 food sources and each of the 5 model-data type combination considered herein.

	Dutch		STRUCTURE			
Source	**MLST**	**MVLST**	**MLST**	**MVLST**	**MLST + MVLST**	**Median**
Dairy	1	1	1	1	1	1
Poultry	5	7	2	2	2	2
Mixed food	3	4	3	3	3	3
Fish	6	6	4	5	4	5
Mixed meat	4	3	5	6	5	5
Game meat	7	5	6	4	6	6
Pork	2	2	7	7	7	7
Beef	8	8	8	8	8	8

Table 5. Pearson correlation coefficient (rho) matrix to calculate the agreement between attributions obtained with the 5 model-data type combination considered herein The lowest and highest rho values are marked in bold.

		Dutch		**STRUCTURE**		
		MLST	**MVLST**	**MLST**	**MVLST**	**MLST + MVLST**
Dutch	**MLST**	1	*	*	*	*
	MVLST	0.979	1	*	*	*
STRUCTURE	**MLST**	0.918	0.85	1	*	*
	MVLST	0.762	**0.702**	0.934	1	*
	MLST + MVLST	0.899	0.828	**0.997**	0.953	1

Highlights

- Up to 53% of listeriosis cases in Northern Italy are attributable to dairy products
- 37% of isolates were Epidemic Clones, strains involved in more than one outbreak
- Poultry accounted for up to 18% listeriosis cases
- Including isolates at the reservoir level may identify cross-contamination events

Conflict of Interest and Authorship Conformation Form

Please check the following as appropriate:

- ✓ All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.
- ✓ This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.
- ✓ The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript
  - o The following authors have affiliations with organizations with direct or indirect financial interest in the subject matter discussed in the manuscript:

Author's name	Affiliation