Molecular surveillance of *Listeria monocytogenes* in Germany to control transmission along food supply chains and to prevent human listeriosis cases

Inaugural-Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry, Pharmacy

of Freie Universität Berlin

by

Stefanie Lüth

2021

"However, it was not until the mid 1950's that *L. monocytogenes* really passed from the awkward and unnoticed adolescent among the pathogenic bacteria into an exciting, sometimes coy, almost flirtatious bacterium – one that enticed and captivated and seemed to be everywhere."

Gray, M. L., & Killinger, A. H. (1966). *Listeria monocytogenes* and listeric infections. *Bacteriological reviews*, 30(2), 309–382.

The following thesis was developed at the German Federal Institute for Risk Assessment (Berlin, DE) in the National Reference Laboratory for *Listeria monocytogenes* from January 2017 to June 2020 under supervision of Prof. Dr. Sascha Al Dahouk and Dr. Sylvia Kleta. The work was supported by a grant of the Federal Ministry of Health (GE 2016 03 26) in the framework of the National Research Platform for Zoonoses and by the German Federal Institute for Risk Assessment (1322-668).

I hereby confirm that I have written the accompanying thesis by myself, without contributions from any sources other than those cited in the text.

First Reviewer: Prof. Dr. Sascha Al Dahouk

German Federal Institute for Risk Assessment

Department of Biological Safety

Diedersdorfer Weg 1

D-12277 Berlin

Second Reviewer: Prof. Dr. Jens Rolff

Freie Universität Berlin

Institute of Biology - Zoology

Königin-Luise-Straße 1-3

D-14195 Berlin

Date of disputation: 11.05.2021

Table of contents

a. List of abbreviations I

b. List of figures III

c. List of tables III

d. List of peer-reviewed publications and conference contributions IV

a. Peer-reviewed publications IV

b. Talks as speaker V

c. Posters as presenter VI

1 Zusammenfassung 1

2 Summary 4

3 About *Listeria monocytogenes* and listeriosis 6

3.1 Listeriosis in livestock 6

3.2 Human listeriosis 7

3.3 *Listeria monocytogenes* in food industry 9

4 Objectives and structure of the work 12

5 Results and discussion covering the major findings published within the thesis 14

5.1 Molecular typing methods for *Listeria monocytogenes* 14

5.1.1 WGS as a typing tool 16

5.1.2 The challenges for WGS data sharing 18

5.1.3 Comparison of WGS methods 20

5.2 Applications of WGS-based typing 24

5.2.1 WGS typing helps to improve process hygiene 24

5.2.2 WGS in outbreak clarifications 26

5.2.3 WGS for risk assessment 29

5.3 Tools for risk communication 31

6 Conclusions and future perspectives 35

7 Danksagung 38

8 References 39

9 Appendix: Publications and supplementary material 48

a. List of abbreviations

AGES	-	Österreichische Agentur für Ernährungssicherheit (English name: Austrian Agency for Health and Food Safety Ltd.)
AMR	-	Antimicrobial Resistance
BBACA	-	Blast-score-ratio-Based Allele Calling Algorithm
BfR	-	Bundesinstitut für Risikobewertung (English name: German Federal Institute for Risk Assessment)
CARD	-	Comprehensive Antibiotic Resistance Database
CC	-	Clonal Complex
CDC	-	Centers for Disease Prevention and Control
CFU	-	Colony Forming Units
cgMLST	-	Core Genome Multi Locus Sequence Typing
EC	-	European Commission
ECDC	-	European Centre for Disease Prevention and Control
ECOFF	-	Epidemiological Cut-Off Value
EEA	-	European Economic Area
EFSA	-	European Food Safety Authority
EHEC		Enterohaemorrhagic *Escherichia coli*
EPIS	-	Epidemic Intelligence Information System
EU	-	European Union
HACCP	-	Hazard Analysis and Critical Control Point
L.	-	*Listeria*
LIPI	-	*Listeria* Pathogenicity Island
MIC	-	Minimum Inhibitory Concentration

MLST	-	Multi Locus Sequence Typing
NCBI	-	National Centre for Biotechnology Information
NGS	-	Next Generation Sequencing
PCR	-	Polymerase Chain Reaction
PFGE	-	Pulsed-Field Gel Electrophoresis
RASFF	-	Rapid Alert System for Food and Feed
RKI	-	Robert Koch Institute
RTE	-	Ready-to-eat
SNP	-	Single Nucleotide Polymorphism
ST	-	Sequence Type
UPGMA	-	Unweighted Pair Group Method with Arithmetic Mean
VFDB	-	Virulence Factor Database
wgMLST	-	Whole Genome Multi Locus Sequence Typing
WGS	-	Whole Genome Sequencing

b. List of figures

Figure 1: Number of reported listeriosis cases in Germany from 2001 to 2019. 9

Figure 2: Overview of the food safety criteria for *L. monocytogenes* in ready-to-eat foods laid down in European Regulation (EC) No 2073/2005. 10

Figure 3: Levels at which *L. monocytogenes* should be controlled. 12

Figure 4: Principle of PFGE. 15

Figure 5: Principle of NGS. 17

Figure 6: Approaches for WGS-based typing. 18

Figure 7: Concepts for WGS data sharing and analysis. 19

Figure 8: Clustering as the basis for genomic epidemiology. 20

Figure 9: Comparison/Benchmarking of WGS-based typing methods. 22

Figure 10: Epidemiological curve of a listeriosis outbreak coloured by the German federal states that notified the cases. 26

Figure 11: Overview of the concerted interaction between the public health and the food sector for listeriosis outbreak investigations. 28

Figure 12: Procedure of a notification via the RASFF. 32

c. List of tables

Table 1: Key figures of bacterial zoonotic diseases in Germany 2019. 7

d. List of peer-reviewed publications and conference contributions

a. Peer-reviewed publications

Lüth S[†,*], Deneke C[†], Kleta S, Al Dahouk S. Translatability of WGS typing results can simplify data exchange for surveillance and control of *Listeria monocytogenes*. Microbial Genomics. 2020; mgen000491.

Halbedel S*, Wilking H, Holzer A, Kleta S, Fischer MA, **Lüth S**, Pietzka A, Huhulescu S, Lachmann R, Krings A, Ruppitsch W, Leclercq A, Kamphausen R, Meincke M, Wagner-Wiening C, Contzen M, Kraemer IB, Al Dahouk S, Allerberger F, Stark K, Flieger A. Large nationwide outbreak of invasive listeriosis associated with blood Sausage, Germany, 2018–2019. Emerging Infectious Diseases. 2020; 26(7):1456-1464.

Lüth S*, Halbedel S, Rosner B, Wilking H, Holzer A, Roedel A, Dieckmann R, Vincze S, Prager R, Flieger A, Al Dahouk S, Kleta S. Backtracking and forward checking of human listeriosis clusters identified a multiclonal outbreak linked to *Listeria monocytogenes* in meat products of a single producer. Emerging Microbes & Infections. 2020; 9(1): 1600-1608.

Gwida M[†], **Lüth S**[†,*], El-Ashker M, Zakaria A, El-Gohary F, Elsayed M, Kleta S, Al Dahouk S. Contamination pathways can be traced along the poultry processing chain by whole genome sequencing of *Listeria innocua*. Microorganisms. 2020; 8(3):414.

Adler M*, **Lüth S**, Kleta S, Al Dahouk S. Draft genome sequence of a *Listeria monocytogenes* isolate of core genome multilocus sequence typing complex type 2521 from ready-to-eat meat sausage related to an outbreak (Sigma1) in Germany. Microbiology Resource Announcement. 2020; 9(18).

Lüth S*, Boone I, Kleta S, Al Dahouk S. Analysis of RASFF notifications on food products contaminated with *Listeria monocytogenes* reveals options for improvement in the rapid alert system for food and feed. Food Control. 2019; 96:479-487.

Lüth S*, Kleta S, Al Dahouk S. Whole genome sequencing as a typing tool for foodborne pathogens like *Listeria monocytogenes* – The way towards global harmonisation and data exchange. Trends in Food Science & Technology. 2018; 73:67-75.

[†] - These authors contributed equally to the work.

* - Corresponding author

b. Talks as speaker

Lüth S: WGS-basierte Typisierung von Listerien am BfR. BfR-Symposium: *Listeria monocytogenes* – Aktuelle Herausforderungen für die Lebensmittelsicherheit und den Verbraucherschutz 16.-17.11.2020, Berlin, DE

Lüth S, Deneke C, Kleta S, Al Dahouk S: Concordance between genomic typing methods can simplify data exchange for surveillance and control of infectious diseases. International Symposium on Zoonoses Research, 16.-18.10.2019, Berlin, DE

Lüth S, Kleta S, Merle R, Alter T, Halbedel S, Flieger A, Al Dahouk S: Molekulare Typisierung von *Listeria monocytogenes* in Lebensmitteln als Grundlage für eine effiziente Risikobewertung und Bekämpfung der Listeriose in Deutschland. Fortbildungsveranstaltung Lebensmittelüberwachung des Landesverbands der Tierärzte im Öffentlichen Dienst Mecklenburg-Vorpommern, 10.4.2019, Klink, DE

Lüth S, Kleta S, Halbedel S, Flieger A, Al Dahouk S: Surveillance of *Listeria monocytogenes* in foodstuffs and food processing plants by WGS in Germany. Joint Conference "Foodborne pathogens and whole genome sequencing", 26.-28.3.2019, Paris, FR

Lüth S, Kleta S, Merle R, Alter T, Halbedel S, Flieger A, Al Dahouk S: Molekulare Typisierung von *Listeria monocytogenes* in Lebensmitteln als Grundlage für eine effiziente Risikobewertung und Bekämpfung der Listeriose in Deutschland. 3. LGL Kongress Lebensmittelsicherheit, 21.-22.11.2018, Erlangen, DE

Lüth S, Kleta S, Merle R, Alter T, Al Dahouk S: MolTypList - Molecular typing of *Listeria monocytogenes* in isolates from foodstuffs and food-processing plants in Germany. 26th International ICFMH Conference FoodMicro, 3.-6.9.2018, Berlin, DE

Lüth S, Kleta S, Al Dahouk S: Molecular typing of *Listeria monocytogenes* in food as basis for an efficient risk assessment and control of listeriosis in Germany. Junior Scientist Zoonoses Meeting, 7.-9.6.2018, Hamburg, DE

Lüth S, Kleta S, Flieger A, Halbedel S, Prager R, Merle R, Alter T, Al Dahouk S: Molecular typing of *Listeria monocytogenes* in foodstuffs to combat human listeriosis in Germany. National Symposium on Zoonoses Research, 12.-13.10.2017, Berlin, DE

c. Posters as presenter

Lüth S, Kleta S, Deneke C, Ferrari D, Romdhane R B, Halbedel S, Alter T, Flieger A, Al Dahouk S: Whole genome sequencing for surveillance and control of *Listeria monocytogenes*: A success story about reducing the burden of listeriosis in Germany. International Symposium on Problems of Listeria and Listeriosis (ISOPOL) 2019, 24.-27.9.2019, Toronto, CA

Lüth S, Kleta S, Merle R, Alter T, Al Dahouk S: Whole genome sequencing of *Listeria monocytogenes* food isolates reveals population structure and facilitates outbreak clarification. National Symposium on Zoonoses Research, 17.-19.10.2018, Berlin, DE

Lüth S, Boone I, Kleta S, Al Dahouk S: Analysis of Rapid Alert System for Food and Feed (RASFF) notifications on products contaminated with *Listeria monocytogenes* affecting Germany, 2001-2015. 26th International ICFMH Conference – FoodMicro, 3.-6.9.2018, Berlin, DE

1 Zusammenfassung

Die humane Listeriose ist eine vergleichsweise seltene, aber schwerwiegende Infektionserkrankung. Die hohe Hospitalisierungs- und Sterblichkeitsrate machen sie weltweit zu einem großen Problem für die öffentliche Gesundheit. Ältere, Schwangere und Immungeschwächte haben ein erhöhtes Infektionsrisiko. Verursacht wird die Listeriose durch den Verzehr von Lebensmitteln, die mit dem Bakterium *Listeria* (*L.*) *monocytogenes* kontaminiert sind. *Listeria monocytogenes* ist weit verbreitet in der Umwelt und Tierwelt. In die Lebensmittelkette gelangt es entweder über kontaminierte rohe tierische oder pflanzliche Produkte oder über Kreuzkontaminationen während der Verarbeitung von Lebensmitteln. Das übergeordnete Ziel der vorliegenden Arbeit war es, mittels modernster molekularepidemiologischer Methoden und genetischem Profiling von Bakterienisolaten Listeriosefälle beim Menschen zu verhindern.

Neben der klassischen Epidemiologie werden molekulare Typisierungsmethoden angewendet, um den Eintrag und die Verbreitung von *L. monocytogenes* entlang der Lebensmittelkette bis hin zum Verbraucher aufzudecken. In den letzten Jahren wurde die molekulare Typisierung durch die Möglichkeit zur Gesamtgenomsequenzierung (WGS, *whole genome sequencing*) bakterieller Isolate revolutioniert. Anhand der in der vorliegenden Promotionsarbeit gewonnenen Erkenntnisse konnte der bisherige Goldstandard zur Feintypisierung, die Pulsed-Field-Gelektrophorese, im Nationalen Referenzlabor (NRL) für *L. monocytogenes* in 2018 komplett durch die WGS abgelöst werden. Der Wert der WGS als hochauflösendes, molekulares Überwachungstool für *L. monocytogenes* wurde im Rahmen des Promotionsprojekts umfassend an die Stakeholder entlang der Lebensmittelkette kommuniziert. Auf der Basis der im Zuge des Promotionsprojekts im NRL etablierten Protokolle steht das Tool den nationalen Überwachungsbehörden nun zur Verfügung.

Innerhalb des Promotionsvorhabens wurde die bioinformatische Analyse von WGS-Daten von *L. monocytogenes* für den Datenaustausch optimiert, beispielhaft ein Listerioseausbruch und die Verbreitung von *Listeria* spp. in einem lebensmittelverarbeitenden Betrieb untersucht, sowie die Bedeutung des europäischen Schnellwarnsystems für lebensmittelbedingte Krankheitsausbrüche evaluiert.

Mittels WGS lassen sich Kontaminations- und Infektionsketten mit beispielloser Genauigkeit nachvollziehen und so unterbrechen. Die Verfahren zur Analyse von WGS-Daten sind jedoch vielfältig und noch nicht umfassend harmonisiert, was den Datenaustausch erschwert. In der vorliegenden Arbeit wurde gezeigt, dass mit verschiedenen WGS-Analysemethoden für *L. monocytogenes* größtenteils vergleichbare Ergebnisse generiert werden können. Ein innerhalb der Promotionsarbeit entwickelter Übersetzungscode ermöglicht es, Informationen zu Clustern auch ohne harmonisierte Methoden zwischen Sektoren (z.B. Lebensmittelsicherheit, Öffentliches Gesundheitswesen) und Ländern

auszutauschen. Dieses Vorgehen bietet einen großen Vorteil, indem es gerade in zeitkritischen Situationen wie Listerioseausbrüchen eine schnelle Kommunikation zwischen den Stakeholdern erlaubt.

Eine wichtige Voraussetzung, um Listerioseinfektionen zu verhindern, ist ein stärkerer Fokus auf die Betriebshygiene in der Lebensmittelverarbeitung. Deutlich wurde dies am Beispiel einer im Rahmen der Promotionsarbeit durchgeführten Studie einer Geflügelverarbeitungskette. Mittels WGS-Typisierung konnte gezeigt werden, dass *Listeria* spp. vom Tier und aus der Umwelt bis in das fertige Lebensmittelprodukt übertragen werden. Zusätzlich zeigte sich eine Ausbreitung auf Oberflächen in der Produktionsumgebung, wodurch gefährliche Kreuzkontaminationen begünstigt werden.

Seit der Einführung der WGS für die Typisierung von *L. monocytogenes* in Deutschland konnte in sektorübergreifender Zusammenarbeit eine Vielzahl von Listerioseausbrüchen aufgeklärt und beendet werden. Die vorliegende Arbeit gibt Einblick in einen großen, über mehrere Jahre anhaltenden, nationalen Listeriseausbruch, der im Rahmen des Promotionsvorhabens federführend untersucht wurde. Mittels WGS-Typisierung konnte der Ausbruch auf verzehrfertige Fleischprodukte und deren Herstellerbetrieb zurückverfolgt werden. Durch eine vorwärts gerichtete Überprüfung (*forward checking*) als Ergänzung zu der in Ausbruchsuntersuchungen üblichen Rückverfolgung (*backtracking*), konnten zwei verschiedene Cluster von Listeriosefällen demselben Betrieb zugeordnet werden. Nur durch solch eine zweiseitige Kontrollstrategie, bei der *backtracking* und *forward checking* kombiniert werden, können Listerioseausbrüche in Zukunft wesentlich minimiert werden. Als Resultat der intensiven Beprobung des Herstellerbetriebs zeigte sich eine sehr diverse *L. monocytogenes*-Population mit teils über mehrere Jahre persistierenden Stämmen. *Listeria monocytogenes* wurde insbesondere in für Reinigung und Desinfektion schwer zu erreichenden Nischen gefunden. Diese Studie deutete auf Schwachpunkte in der Betriebshygiene hin, die die Produktion kontaminierter Lebensmittel nach sich zogen und so Auslöser des Listerioseausbruchs waren.

Wird *L. monocytogenes* in einem Lebensmittel gefunden, ermöglicht das europäische Schnellwarnsystem für Lebensmittel und Futtermittel (RASFF, *Rapid Alert System for Food and Feed*) eine internationale Kommunikation dieses Risikos und beispielsweise umgehende Produktrückrufe. Auf diese Weise trägt es essenziell dazu bei, Verbraucher vor lebensmittelbedingten Infektionen zu schützen. Die Analyse von RASFF-Meldungen bezüglich *L. monocytogenes* im Rahmen der vorliegenden Arbeit verdeutlichte die Komplexität internationaler Verarbeitungs- und Vertriebswege. Dies unterstreicht, wie wichtig auch die internationale Zusammenarbeit im Hinblick auf die Lebensmittelsicherheit ist. Die Tatsache, dass einzelne lebensmittelverarbeitende Betriebe in zahlreiche RASFF-Meldungen in unterschiedlichen Jahren involviert waren, zeigte erneut den Einfluss unzureichender Betriebshygiene und Persistenz auf die Ausbreitung von *L. monocytogenes*. Zudem verdeutlichten die Studienergebnisse den Bedarf einer stärkeren Eigenverantwortung lebensmittelverarbeitender Betriebe.

Schon während der Laufzeit zeigte sich der Erfolg der im Rahmen des Promotionsprojekts etablierten und im Projektverlauf stetig optimierten WGS-basierten Überwachungsstrategie von *L. monocytogenes*. Die Zahl gemeldeter Listeriosefälle in Deutschland war in 2018 erstmals nicht mehr ansteigend, sondern rückläufig, und sank in 2019 weiter. Das Promotionsvorhaben hat damit entscheidend dazu beigetragen, die Ausbreitung von *L. monocytogenes* entlang der Nahrungskette zu kontrollieren und so die Zahl der Listerioseerkrankungen in Deutschland zu reduzieren.

2 Summary

Human listeriosis is a comparatively rare but serious infectious disease. The high hospitalisation and mortality rate make it a major public health concern worldwide. Elderly and immunocompromised people as well as pregnant women are at increased risk of infection. Listeriosis is caused by consumption of food contaminated with the bacterium *Listeria* (*L.*) *monocytogenes*. *Listeria monocytogenes* is widespread in the environment and in animals. It enters the food chain either via contaminated raw animal or plant products or via cross-contamination during food processing. The overall aim of the present dissertation was to prevent cases of listeriosis in humans using state-of-the-art molecular epidemiological methods and genetic profiling of bacterial isolates.

As a support to classical epidemiology, molecular typing methods are used to monitor the entry and spread of *L. monocytogenes* along the food chain up to the consumer. In recent years, the possibility of whole genome sequencing (WGS) has revolutionised the molecular typing of bacterial isolates. Based on the knowledge gained in the present doctoral project, the previous gold standard for fine typing, pulsed-field electrophoresis, was completely replaced by WGS in the National Reference Laboratory (NRL) for *L. monocytogenes* in 2018. As part of the present project, the value of WGS as a high-resolution molecular surveillance tool for *L. monocytogenes* has been widely communicated to stakeholders along the food chain. Based on the protocols established in the NRL during the doctoral project, the tool is now available to the national monitoring authorities.

Within the dissertation project, the bioinformatic analysis of WGS data of *L. monocytogenes* was optimised for data exchange, a listeriosis outbreak and the spread of *Listeria* spp. in a food processing plant were investigated, and the significance of the European Rapid Alert System for Food and Feed (RASFF) for foodborne outbreaks was evaluated.

Using WGS, contamination and infection chains can be traced with unprecedented precision and thus be stopped. However, the methods for analysing WGS data are diverse and not yet fully harmonised, which hinders data sharing. In the present work, it was shown that different WGS analysis methods for *L. monocytogenes* generate largely comparable results. A translation code developed within this doctoral project allows information on clusters to be exchanged between sectors (e.g. food safety, public health) and countries even without harmonised methods. This approach offers a major advantage as it allows rapid communication between stakeholders, especially in time-critical situations such as listeriosis outbreaks.

An important prerequisite for the prevention of listeriosis infections is a stronger focus on industrial hygiene in food processing. This became clear in the example of a study of a poultry processing chain carried out as part of this doctoral project. By means of WGS typing, it could be shown that *Listeria* spp. are transferred from the animal and from the environment to the finished food product. In addition,

bacteria were shown to spread to surfaces in the production environment, favouring dangerous cross-contamination.

Since the introduction of WGS for typing *L. monocytogenes* in Germany, in cross-sectoral collaboration, a large number of listeriosis outbreaks could be clarified and stopped. This work provides insight into a large national listeriosis outbreak lasting several years, which was lead investigated in the present doctoral project. Using WGS typing, the outbreak was traced back to ready-to-eat meat products and their producer. By using forward checking as a supplement to the usual backtracking in outbreak investigations, two different clusters of listeriosis cases could be assigned to the same producer. Only through such a two-sided control strategy, combining backtracking and forward checking, can listeriosis outbreaks be significantly minimised in the future. As a result of the intensive sampling at the producer, a very diverse *L. monocytogenes* population was found, with some strains persisting for several years. *Listeria monocytogenes* was found especially in niches of the food processing environment, difficult to reach for cleaning and disinfection. This study pointed to weaknesses in industrial hygiene that resulted in the production of contaminated food and thus triggered the listeriosis outbreak.

If *L. monocytogenes* is found in a food product, the European RASFF enables international communication of this risk and, for example, immediate product recalls. In this way, it makes an essential contribution to protecting consumers from foodborne infections. An analysis of the RASFF notifications concerning *L. monocytogenes* as part of this doctoral project highlighted the complexity of international processing and distribution channels. This underlines how important international cooperation is with regard to food safety. The fact that individual producers were involved in many RASFF notifications in different years again showed the influence of inadequate industrial hygiene and persistence on the spread of *L. monocytogenes*. In addition, the results of the study highlighted the need for food processing companies to take greater responsibility for their own operations.

The success of the WGS-based surveillance strategy for *L. monocytogenes*, which was established as part of this doctoral project and continuously optimised over its course, was already evident during the project term. For the first time, the number of reported cases of listeriosis in Germany no longer increased in 2018, but decreased, and continued to decrease in 2019. The dissertation project has thus made a decisive contribution to controlling the spread of *L. monocytogenes* along the food chain and thus reducing the number of listeriosis cases in Germany.

3 About *Listeria monocytogenes* and listeriosis

Listeria (*L.*) *monocytogenes* is a gram-positive, motile, non-spore-forming, facultative anaerobe bacterium [1]. It is ubiquitous in nature and can be found in soil, sewage, plants and livestock [2]. As a facultative intracellular pathogen, it can switch between saprophytic and host-associated lifestyle [3]. *Listeria monocytogenes* belongs to the genus *Listeria*, which currently comprises 21 species. Of those, the six species *L. innocua*, *L. ivanovii*, *L. marthii*, *L. monocytogenes*, *L. seeligeri* and *L. welshimeri* are recognised as *Listeria sensu stricto* due to their common genetic and phenotypic characteristics [4]. Among them, *L. ivanovii* and *L. monocytogenes* are considered the only pathogenic species, leading to the disease listeriosis. The main route of infection is the ingestion of contaminated food. While *L. monocytogenes* can infect humans and ruminants, *L. ivanovii* infection is generally restricted to ruminants [5]. However, also for *L. ivanovii*, rare cases of human infection have been described [6] as well as for the primarily non-pathogenic species *L. innocua* [7, 8]. The 15 species described as *Listeria sensu lato* are *L. aquatica*, *L. booriae, L. cornellensis*, *L. costaricensis*, *L. fleischmannii*, *L. floridensis*, *L. goaensis*, *L. grandensis*, *L. grayi*, *L. newyorkensis*, *L. riparia*, *L. rocourtiae*, *L. thailandensis*, *L. valentina* and *L. weihenstephanensis* [4, 9-12].

3.1 Listeriosis in livestock

In 1924, a bacterium named *Bacterium monocytogenes* was first isolated from livers of sick rabbits and guinea pigs [13]. The later renaming of this genus to *Listeria* [14] makes this the first description of *L. monocytogenes*.

Listeria monocytogenes can infect a variety of different animal hosts, but ruminants like cattle, goats and sheep are most affected [15]. Typical clinical manifestations are neurological disorders caused through encephalitis, and spontaneous abortion. However, complete absence of clinical symptoms is also possible despite bacterial shedding [16, 17]. Generally, diseased as well as asymptomatic carrier animals can contribute to the distribution of *L. monocytogenes* in the environment through saliva, faeces, nasal secretions or aborted material [1, 18]. Listeriosis case numbers in livestock often accumulate in winter and early spring, probably due to the ingestion of contaminated silage during these seasons [19]. Infection of pigs with *L. monocytogenes* is rare, possibly because silage is not a common type of feed for them. Nevertheless, colonisation of the intestine in healthy pigs can lead to the distribution of *L. monocytogenes* through their faeces [20, 21]. Birds are also rarely infected with *L. monocytogenes*, but the pathogen has been isolated from a range of domestic and wild avian species like chicken, goose, duck, turkey or pigeon [22]. Also aquatic animals like fish or shellfish can be carriers of *L. monocytogenes*, presumably as a result of contact with contaminated water [23].

Direct contact transmission from animals to humans is seen very rarely and occurs mainly in veterinarians and farmer, who develop a cutaneous listeriosis without systemic involvement [24]. The vast majority (99%) of human infections are foodborne and due to the consumption of contaminated raw animal-derived products (e.g. meat or milk), raw products of plant origin, which are contaminated by the environment (e.g. soil, water, organic fertilisers) or food products that have been cross-contaminated during processing [25, 26].

3.2 Human listeriosis

Although listeriosis is rare compared to other bacterial zoonoses, the hospitalisation rate and the case fatality rate are remarkably high (Table 1), thus rendering the disease a serious public health concern.

Table 1: Key figures of bacterial zoonotic diseases in Germany 2019 [27].

Disease	**Number of cases**	**Number of fatal cases**	**Case fatality rate (in %)**	**Hospitalisation rate (in %)**
Listeriosis	591	40	6.77	95
Salmonellosis	13,693	18	0.13	37
EHEC disease	1,877	1	0.05	28
Campylobacter enteritis	61,526	4	0.01	23

After uptake of contaminated food, *L. monocytogenes* enters the intestinal epithelium of the gastrointestinal tract and crosses it to reach the lamina propria. From there, it goes into lymph and blood and is carried to mesenteric lymph nodes, liver and spleen where it starts to multiply [5]. If the infection cannot be cleared through the immune response at this stage, *L. monocytogenes* reaches the bloodstream and can infect various other tissues [5, 28]. In immunocompromised humans, the pathogen can cross the blood-brain barrier [29] and in pregnant women the placental barrier [30]. Elderly or immunocompromised people and pregnant women therefore have an increased risk of developing invasive listeriosis. Then, clinical presentation is associated with life-threatening symptoms such as septicaemia, encephalitis or meningitis [5]. Pregnant women usually show no or only slight flu-like symptoms, but transmission to the unborn child often leads to sepsis and multiple organ manifestations, which can result in premature birth, miscarriage or stillbirth. New-borns with listeriosis have a high risk of mortality due to multi-organ failure and/or insufficient lung maturity [31]. In healthy people who do not belong to a risk group for listeriosis, the infection usually remains enteric and leads to febrile

gastroenteritis, which is usually mild and self-limiting [32]. The incubation time varies and depends strongly on the clinical manifestation of the disease. It can be 1 to 67 days for invasive listeriosis (longest incubation period in pregnancy-related cases) and six hours to four days for non-invasive listeriosis [33]. The infection dose is unclear, but likely depends on the immune status of the patient, the virulence of the infecting strain and the food matrix, which influences how well the pathogen can survive the stomach passage [34, 35].

In Germany, the recommended antibiotic therapy for listeriosis is based on amoxicillin or ampicillin in high doses, in combination with an aminoglycoside (usually gentamicin) for at least three weeks, unless the patient is pregnant [36]. Cotrimoxazole is recommended in the second place. Although no resistance to those antibiotics has been found in German *L. monocytogenes* isolates [37-39], the antibiotic therapy is not always effective. Reasons can be the intracellular lifestyle of the pathogen, immunosuppression of the patient or a late stage of disease.

Human listeriosis is a notifiable disease since 2001 in accordance with §7 of the German infection protection law. Conditions for reporting are isolation of *L. monocytogenes* from normally sterile sites like blood or liquor or in swabs from new-borns or maternal tissue (e.g. placenta), or clinical-epidemiological confirmation of a case [36]. The binational consultant laboratory for *L. monocytogenes* at the German Robert Koch Institute (RKI) and the Austrian Agency for Health and Food Safety Ltd. (AGES) collects *L. monocytogenes* strains isolated from human infections in Germany and Austria. During the last years, approximately 450 German isolates were collected annually, corresponding to approximately two thirds of all notified German listeriosis cases [40].

In Germany, numbers of listeriosis infections continuously increased from 2011 (337 cases) to 2017 (770 cases), corresponding to an incidence increase from 0.4 to 0.9 cases per 100,000 population (Figure 1) [41, 42]. In 2018, the number of listeriosis cases decreased by 9% to 701 reported cases compared to the previous year (incidence of 0.8 cases per 100,000 population) and further decreased to 591 cases in 2019 (incidence of 0.8 cases per 100,000 population) [27, 43].

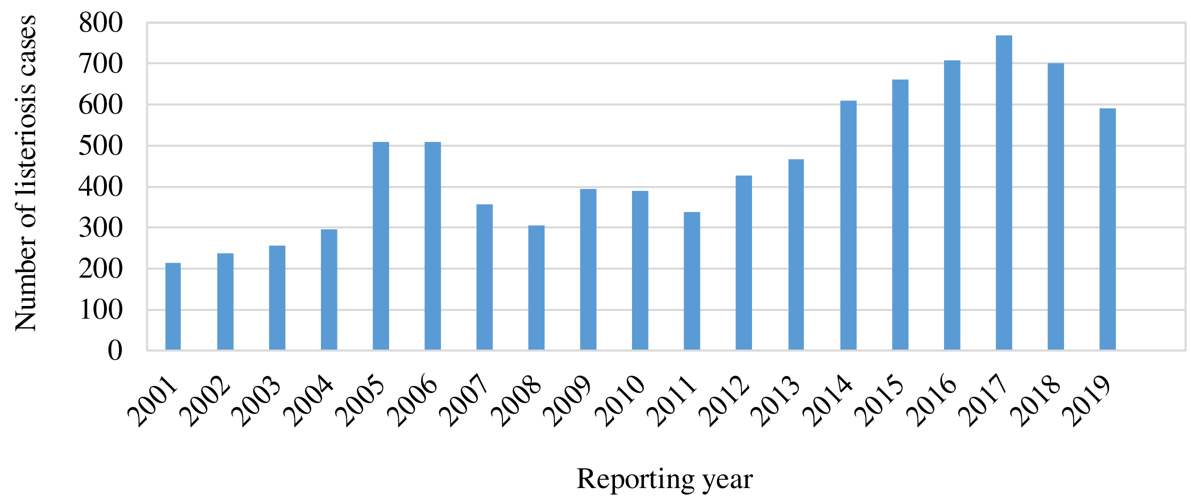


Figure 1: Number of reported listeriosis cases in Germany from 2001 to 2019 [27, 41-58].

A rising trend in listeriosis cases was also observed in the European Union (EU) and the European Economic Area (EEA). More specifically, in the years 2008 to 2015, there was a significant increase of listeriosis cases in the group of 25 to 44 year old women (probably pregnancy-related listeriosis) and in the group of women and men over 75 years of age [59]. In 2018, a total of 2549 cases were reported [60]. The case fatality rate was substantially higher than in Germany in the same year (15.6% versus 4.4%)[43].

3.3 *Listeria monocytogenes* in food industry

Listeria monocytogenes is non-fastidious. It is resilient to salt (up to 10% *w/v* NaCl) and can grow at a broad temperature (0 to 45°C) and pH range (4.1 to 9.6) [61-64]. Even freezing at -20°C does not significantly reduce the number of live bacteria [65]. In this way, *L. monocytogenes* is well adapted to survival and even growth in preserved and chilled food products and food processing environments, properties that make the pathogen a serious problem for food industry [25]. Typical categories of affected foods are milk, meat or fish products and vegetables [25]. Especially contaminated ready-to-eat (RTE) products that are not subjected to bactericidal treatment (e.g. heating) directly prior to eating pose a risk of infection to the consumer. According to European regulation (EC) No 2073/2005 (Figure 2), absence of *L. monocytogenes* in 25 g RTE food during the entire shelf life is required if intended for infants or special medical purposes [66]. Otherwise, contamination must not exceed 100 CFU/g. In the latter case, an additional rule applies when the food matrix supports growth of *L. monocytogenes*. Then, absence in 25 g is required before the product has left the direct control of the food business operator who produced it [66]. An exemption may be granted if the operator can prove (e.g. by shelf life studies)

that the threshold of 100 CFU/g is not exceeded during shelf life. Then the limit of 100 CFU/g until the end of shelf life also applies before leaving direct control.

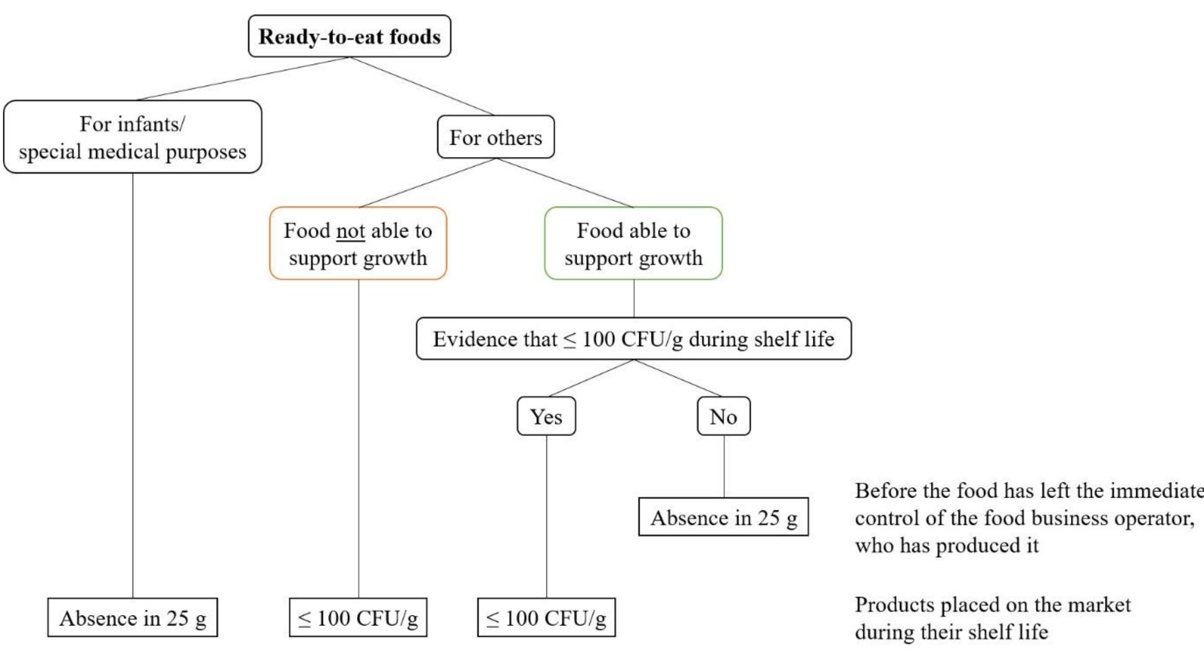


Figure 2: Overview of the food safety criteria for L. monocytogenes in ready-to-eat foods laid down in European Regulation (EC) No 2073/2005.

Either a pH > 4.4 and water activity > 0.92 or a combination of pH > 5.0, water activity > 0.94 and NaCl < 16% are generally considered as intrinsic food conditions that support growth of *L. monocytogenes* [34]. The regulations concerning the maximum permissible levels of the pathogen in foodstuffs are not uniform worldwide [23]. A zero tolerance (absence in 25 g sample) for example applies in all cases in the United States. Whether this approach is of advantage over the EU-wide 100 CFU/g limit is controversially discussed, for instance due to the large economic losses caused through product recalls [67, 68].

Listeria monocytogenes can enter and settle in food processing plants (e.g. in niches that are difficult to reach) via contaminated raw products or contamination from the environment. This creates the risk of cross-contamination during food production. Persistent *L. monocytogenes* strains in food processing plants have been identified as the most common post processing contaminants [68]. Persistence may be caused by a complex interplay of different insufficiently applied hygiene measures, e.g. no strict separation between clean and dirty areas, use of sub-lethal biocide concentrations or incomplete cleaning [69-72]. This may favour establishment of tolerance to biocides, biofilm formation or persistence in niches. *Listeria monocytogenes* itself is mostly a weak biofilm former, but it can take advantage of multispecies biofilms [73-75]. Biofilms are on the one hand hard to remove and on the other hand

impede the accessibility of the bacteria by biocides [76, 77]. Persisting *L. monocytogenes* strains may contaminate food products during processing, explaining the recurring entry of certain strains into the food chain [72]. If no measures are taken, persistence can last for decades [76, 78].

The National Reference Laboratory for *L. monocytogenes* at the German Federal Institute for Risk Assessment (BfR) receives more than 1000 *L. monocytogenes* isolates per year from food and the production environment, which have been tested positive as planned, suspect, trace or zoonoses monitoring samples in the official food monitoring throughout Germany. They allow insight into the exposure of the consumer to *L. monocytogenes* via both RTE and non-RTE foods of animal and plant origin in Germany. In Germany, high prevalence of *L. monocytogenes* is found in cold-smoked and graved fish (7 to 18%), hot-smoked fish (3 to 9%), differently preserved fish (4 to 10%) and differently stabilised meat products (10 to 17%) [79-87]. In all these product groups, quantitative tests have shown that the safety criteria for *L. monocytogenes* laid down in Regulation (EC) No 2073/2005 were regularly not met. High prevalence of *L. monocytogenes* was also found in minced meat (7 to 22%) and minced meat preparations, which are offered to the consumer for raw consumption (15 to 24%). Numerous other RTE food products such as cheese made from raw or pasteurised milk, heat-treated meat products (e.g. sausages), delicacy salads, sliced lettuce and raw vegetables may also be contaminated with *L. monocytogenes.*

4 Objectives and structure of the work

The overall goal of the present work was to reduce and to prevent human listeriosis cases. However, there is no simple solution for how to reach this goal. Instead, it requires an interaction of many different, closely interconnected and interdependent factors. Starting points lie on three different levels within the food chain from stable to table (Figure 3).

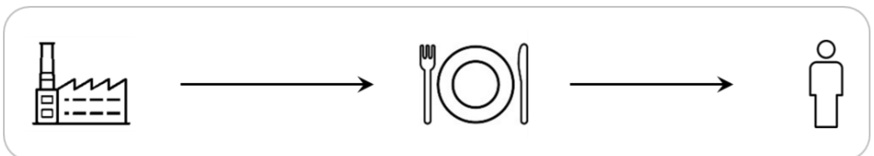


Figure 3: Levels at which L. monocytogenes should be controlled. From left to right: food processing, food product, human patient.

On the first level, contamination has to be prevented during food processing. *Listeria monocytogenes* is ubiquitous in nature and this cannot be stopped. Nevertheless, entry into and consolidation in processing plants as well as transfer during food processing can be avoided. The application of hazard analysis and critical control point (HACCP) principles is mandatory to food business operators according to Regulation (EC) No 852/2004. On the second level, contamination of final food products has to be identified in order to prevent infection of consumers. This should happen at the earliest possible time so that ideally contaminated products are not placed on the market. Otherwise, at least rapid communication through alarm systems can help to quickly take a contaminated product off the market again. On the third level, human infections have to be traced back to eliminate their source and thus prevent further infections.

Above all, molecular typing of bacterial strains is the key to understand dissemination paths inside the levels and across them. The higher the resolution of a typing method, the more details become visible. However, it is not only important to use the most powerful method. It is also essential to ensure that data can be shared across sectors (e.g. food safety, public health) and countries in order to do justice to the different stakeholders in food production along global supply chains.

The following articles published as a result of this work describe various aspects of the multifactorial process needed to control *L. monocytogenes*. Contents of the publications will be presented in context and not in chronological sequence:

(1) Review Article: Whole genome sequencing as a typing tool for foodborne pathogens like *Listeria monocytogenes* – The way towards global harmonisation and data exchange [88]

- Overview on whole genome sequencing (WGS) methods (wet lab and dry lab)
- Status quo and challenges for standardisation and data sharing

(2) Research Article: Translatability of WGS typing results can simplify data exchange for surveillance and control of *Listeria monocytogenes* [89]

- Comparison of different WGS-based typing methods (focus on dry lab)
- Proposal of an alternative approach to data exchange without complete method standardisation

(3) Research Article: Contamination pathways can be traced along the poultry processing chain by whole genome sequencing of *Listeria innocua* [90]

- WGS-based tracking of contamination along the farm-to-fork continuum (using *L. innocua* as a model for *L. monocytogenes*)
- Identification of entry points and transmission routes during food processing

(4) Research Article: Backtracking and forward checking of human listeriosis clusters identified a multiclonal outbreak linked to *Listeria monocytogenes* in meat products of a single producer [91]

- WGS-based outbreak clarification
- Importance of industrial hygiene for sustainable prevention of listeriosis infections
- Assessment of resistance and virulence potential of *L. monocytogenes* strains

(5) Research Article: Analysis of RASFF notifications on food products contaminated with *Listeria monocytogenes* reveals options for improvement in the rapid alert system for food and feed [92]

- Analysis of the EU-wide reporting system for food contamination (independent of typing)
- Derivation of starting points for improving food safety

5 Results and discussion covering the major findings published within the thesis

5.1 Molecular typing methods for *Listeria monocytogenes*

Trace-back of human cases to contaminated food products is one of the key requirements for the control of foodborne pathogens. However, especially in the case of listeriosis, classical epidemiology has its limitations. First, the incubation period of listeriosis infections is highly variable and can be lengthy [33]. This fact together with the severity of disease complicates patient interviews on food consumption history, which would be important to narrow down the list of possible causative food vehicles. Second, especially during the last years, food vehicles apart from the classical hazards have been found as source of listeriosis infections, further broadening the range of potential candidates [93-95]. Third, listeriosis outbreaks are rarely fulminant events, but rather characterised by intermittent cases of disease over years (e.g. [91, 96, 97], also see section 5.2.2). In order to support classical epidemiology, molecular epidemiology using molecular typing methods has been established in *L. monocytogenes* surveillance and outbreak investigations.

One possibility for sub-typing of *L. monocytogenes* strains is phenotyping. Laboratory methods include classical serotyping [98], phage typing [99] or multilocus enzyme electrophoresis [100]. However, these methods are subject to natural variability due to stress factors or growth phase [101]. As a result, genotypic typing methods have been introduced as a more robust approach [102]. Classical serotyping, for instance, has been mostly replaced by PCR that separates the strains into four major serovars [103].

The *L. monocytogenes* genome consists of a single, circular chromosome with a size of 2.9 mega base pairs and an average G+C content of 39% [104]. For a long time, the gold standard for outbreak investigations was pulsed-field gel electrophoresis (PFGE) due to its high discriminatory power compared to serotyping. The bacterial genome is digested using two restriction enzymes, *AscI* and *ApaI*, and the resulting DNA fragments are separated according to size in a periodically changing electric field (Figure 4). The resulting band pattern corresponds to the fingerprint of a specific strain and forms the basis for assessment of the degree of relatedness over several strains.

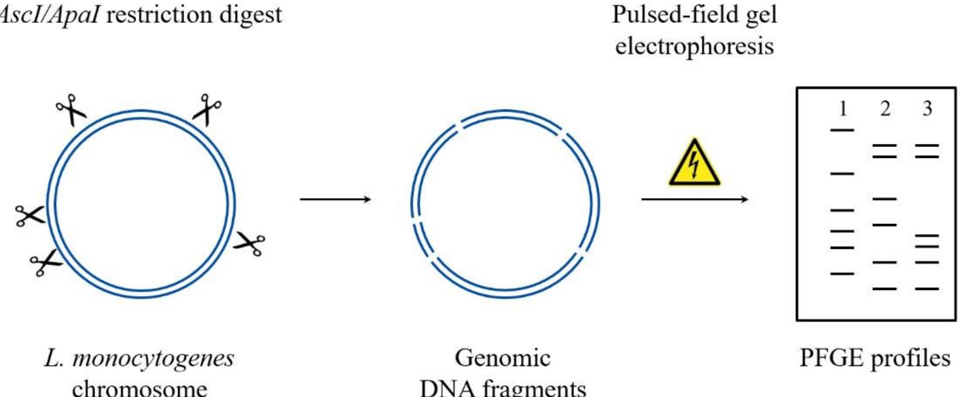


Figure 4: Principle of PFGE.

The advantage of PFGE is the availability of uniform standards for laboratory methods, data analysis and evaluation of results as well as a common database infrastructure on food and human side [105-108]. However, the disadvantages are the very high workload and the limited resolving power. Since genetic changes will be only detected if they fall by chance into the recognition sites of the restriction enzymes used, the genetic relationship of strains is easily over- or underestimated [97]. This makes the results not always reliable.

Besides PFGE, sequence-based typing has also played an important role. One of the most common tools is multi locus sequence typing (MLST) [109], based on sequence variability of seven housekeeping genes. MLST is preferably used to roughly study population structures and was traditionally performed by Sanger sequencing [110]. With the introduction of Next Generation Sequencing (NGS) in 2005 [111], completely new possibilities opened up. WGS is now providing a typing method that allows to dissolve the boundaries of PFGE and MLST and to obtain maximum information about an isolate on the genomic level. The sequencing reads resulting from WGS can be used to reconstruct entire bacterial genomes, to infer phylogenies and to identify genetic markers (e.g. virulence and resistance genes). Thus, WGS has the potential to replace several of its predecessor methods simultaneously. An extensive overview on methods used in the field of WGS data generation and analysis in general discussing progress but also challenges in the worldwide application and harmonisation is provided in the review article published as part of this thesis [88].

Publication: Whole genome sequencing as a typing tool for foodborne pathogens like *Listeria monocytogenes* – The way towards global harmonisation and data exchange

Article type: Review article

Journal: Trends in Food Science & Technology

Impact Factor: 11.08

Authors: Stefanie Lüth, Sylvia Kleta, Sascha Al Dahouk

DOI: https://doi.org/10.1016/j.tifs.2018.01.008

Contributions Stefanie Lüth:

I searched the literature and developed the concept of the article, created the illustrations and wrote the first draft. As first and corresponding author, I submitted the article and considered and responded to the reviewer comments on behalf of all co-authors.

5.1.1 WGS as a typing tool

The introduction of NGS has revolutionised the field of bacterial typing by enabling WGS of strains. Various devices, based on different sequencing technologies (e.g. detection of light or voltage changes), can be used to perform NGS. In principle, the procedure is always the same. Starting from genomic bacterial DNA, a genomic library is prepared and used for sequencing. The result is sequencing reads that contain the genetic information at nucleotide level (Figure 5). The read length depends on the sequencing platform and mode used and ranges between 36 and 300 base pairs.

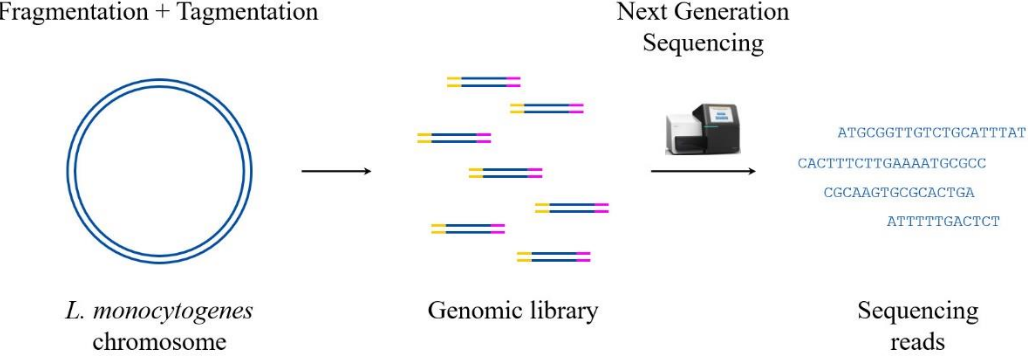


Figure 5: Principle of NGS.

Before the sequencing data is further processed, its quality has to be checked. In NGS, every nucleotide is sequenced multiple times to cover the entire genome and to make the base calls more reliable. The frequency with which a single nucleotide was sequenced on average is described by the coverage value (product of read length and number of reads divided by the haploid genome length). However, not only the value itself, also the evenness of coverage across the sequenced genome is an important quality criterion [112]. The Phred Quality or Q-score gives the quality of reads [113]. It indicates the logarithmic probability of an incorrect base with, for instance, Q30 specifying the probability of one incorrect base in 1000, Q20 one incorrect base in 100 etc. A consistently high coverage together with a high Q-score implies a reliable sequencing result.

The sequencing reads can be further processed in two different ways. They can be compared directly with a genome by aligning/mapping them to it, or they can be used to *de novo* reconstruct a genome by assembly. Due to the shortness of the reads from NGS, it is usually not possible to reconstruct a complete genome through assembly. Instead, shorter, coherent sequences are created, the so-called contigs. Such an assembly of several contigs with gaps in between is called a draft genome. The quality of such a draft genome can be determined by the number of contigs and the N50 value, which is a measure for the weighted median contig size [114]. The lower the number of contigs and the higher the N50 value, the more continuous and better an assembly is.

In order to assess the relationship between bacterial sequences, there are again two different approaches: the allele-based approach, where sequences are compared gene-by-gene, or the Single Nucleotide Polymorphism (SNP)-based approach, where sequences are compared nucleotide-by-nucleotide (Figure 6).

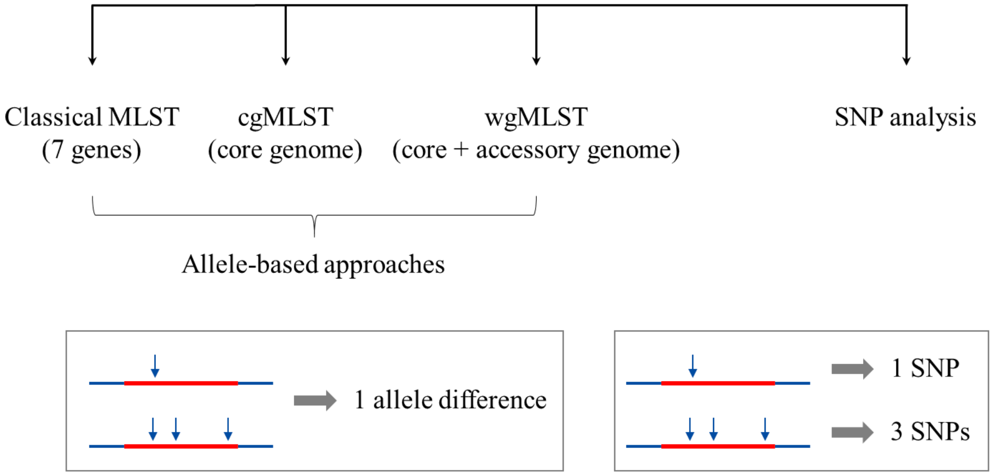


Figure 6: Approaches for WGS-based typing. Red: allele, blue arrow: nucleotide change. Modified from [88].

In the allele-based approach, only the entire allele is considered, no matter how many individual mutations are present inside. The advantage of this method is that the allele type can be easily expressed by a number. In case of a unified nomenclature, simple data exchange is possible through a table containing those numbers. For the classical MLST based on the seven housekeeping genes [109], such a system is already in use (https://bigsdb.pasteur.fr/listeria). By combining allele numbers, an MLST sequence type (ST) can be determined. MLST STs that only differ by one of the seven allele numbers are further combined to MLST clonal complexes (CCs). The classical MLST can be extended to the core genome (cgMLST), defined as the set of genes that is shared among the members of a certain species, and the whole genome (wgMLST), additionally considering genes from the accessory genome (Figure 6). In the SNP-based approach, every single nucleotide change is counted. Although the method has a higher resolution than the allele-based approaches, its data are not as easy to exchange.

5.1.2 The challenges for WGS data sharing

The resolution of WGS clearly exceeds that of its predecessor methods [115, 116], making it the top candidate for gaining detailed insights into the genetic relationship of strains. The challenge in using WGS for typing is that the methods for data generation as well as for their evaluation are extremely diverse. This is problematic particularly in the case of foodborne diseases such as listeriosis, because different sectors (e.g. food safety, public health) have to exchange typing information with each other. In the age of globalised food trade, international communication is also necessary. Therefore, different concepts for data sharing have been proposed (Figure 7).

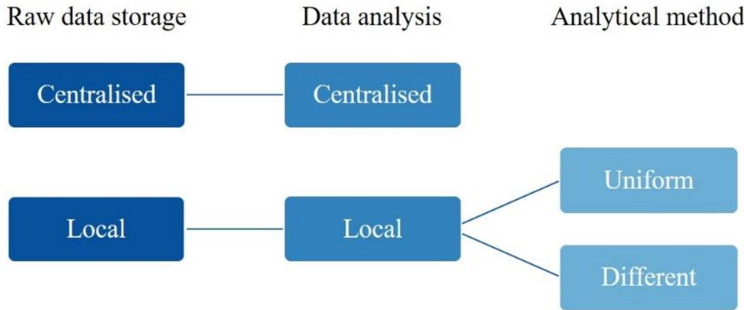


Figure 7: Concepts for WGS data sharing and analysis.

Although the concepts are already used in various forms, none of them has yet been accepted and implemented worldwide. Reservations are to be found at organisational, technical, cultural and scientific level [117]. A centralised concept is in use for example in the United States [118]. This automatically ensures that there are no compatibility problems between typing results from different analytical methods. Only a sufficient quality of the raw data must be guaranteed. The main reservations on this approach concern data protection. In addition, the sequencing laboratory itself cannot influence the analysis, which can particularly become a problem in time-critical situations like outbreaks. Another concept uses local data storage but a uniform analytical method. In this way, intermediate results can be exchanged or uploaded to a central database. The data sovereignty remains with the sequencing laboratory while the analytical results can be easily put into a broader context. However, similar to the centralised approach, it is necessary to agree on a specific method for data analysis in order to be able to exchange the intermediate results seamlessly. This can also be problematic because different laboratories have developed and prefer different procedures, which can lead to uptake challenges of a prescribed method. The widespread status quo at present is therefore that data is stored locally and analysed locally using various methods. Selected raw data is then exchanged on a per-event basis and, if necessary, re-analysed elsewhere. Although the most flexible approach, it is also the only one not supporting real-time monitoring.

Apart from the discussions about sharing WGS-based typing information, it is even more difficult to agree on a concept for sharing accompanying metadata. Genomic data alone is not critical, but publication of accompanying metadata is [119]. Only by combining sequence data with the appropriate metadata can the full data potential be exploited. More than just for the WGS data alone, this has the potential for conflicts with economic or scientific interests. Particularly with regard to cases of illness, the protection of personal data is also an important topic of dispute. Limiting access rights for individual groups of people could help to ensure that sensitive information is only in the hands of those authorised. However, it is difficult to reach agreement on which data are considered sensitive. Further discussions are therefore needed to find widely acceptable solutions in this area.

5.1.3 *Comparison of WGS methods*

One of the most important applications of WGS-based typing data is molecular/genomic epidemiology. The basic idea behind it is: if isolates are genetically closely related, there is a high probability that there is also an epidemiological link (e.g. strains are part of the same outbreak) [120]. In this way, correlations between *L. monocytogenes* isolates from food and/or production environments and from listeriosis patients can be identified. Clustering based on WGS typing (Figure 8) depends on the methods and the threshold value used.

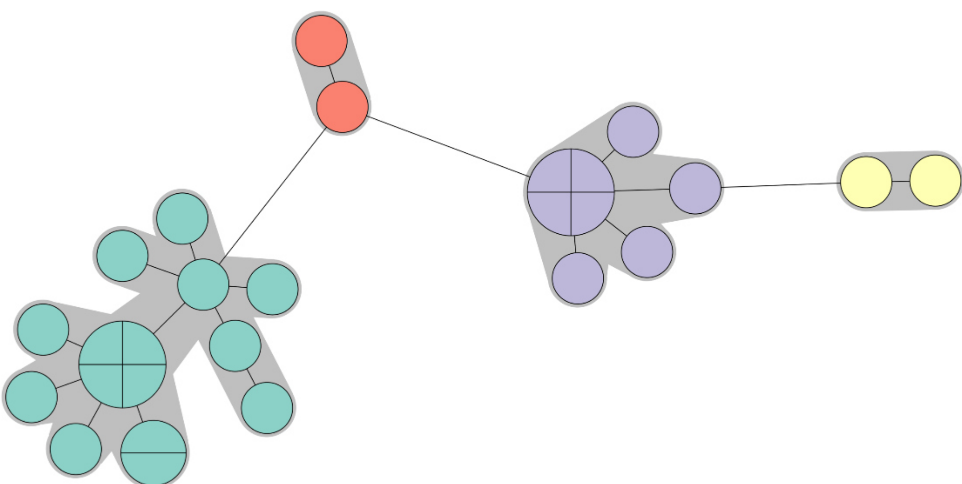


Figure 8: Clustering as the basis for genomic epidemiology. Example of a Minimum Spanning Tree: each circle represents an isolate or a group of identical isolates (individual isolates separated by horizontal and vertical lines), and clusters are marked by a grey background colour.

How can the same results be achieved between laboratories in the food sector and the public health sector or in different countries? As described in 5.1.2, the first option is to use exactly the same analytical method to yield exactly the same clusters. However, the use of such a uniform method does not (yet) fit current reality. The second option is to ensure results are comparable between different analytical methods. A detailed study of this second option together with the proposal of an alternative approach to data exchange without complete method standardisation was published as part of this thesis [89].

Publication: Translatability of WGS typing results can simplify data exchange for surveillance and control of *Listeria monocytogenes*

Article type: Research article

Journal: Microbial Genomics

Impact Factor: 5.52

Authors: Stefanie Lüth†, Carlus Deneke†, Sylvia Kleta, Sascha Al Dahouk

† - Co-First authors

DOI: https://doi.org/10.1099/mgen.0.000491

Contributions Stefanie Lüth:

I have grown and sequenced the vast majority of *L. monocytogenes* strains. Based on literature research, I developed the concept for the study and took care of the involvement of a bioinformatician (C. Deneke) for the implementation. Except for the cgMLST analysis with chewBBACA, I performed all analyses of the sequencing data. I also took care of making the sequence data available in public repositories. I developed the concept for the statistical analysis of the data, applied most of it myself and created the illustrations from it. I wrote the first draft except for a few small sections. As first and corresponding author, I submitted the article and considered and responded to the reviewer comments on behalf of all co-authors.

The basis for assessing the degree of relationship between bacterial strains is the determination of differences between their genomes. In the case of allele-based approaches, the number of allele differences is used; in the case of SNP-based approaches, the number of nucleotide differences is used. If more than two strains are compared with each other, the distance values are combined in a distance matrix. To determine clusters and also to visualise them, different clustering algorithms can be applied. Common algorithms are hierarchical clustering (single-, complete- or average-linkage), UPGMA (Unweighted Pair Group Method with Arithmetic Mean) or Neighbour Joining [121].

To determine the distances between isolates at all, there are commercial and free and open-source software solutions. Commercial tools are designed to be user-friendly and thus equipped with a graphical user interface or pre-set workflows. Even for users without bioinformatics knowledge, such solutions are usually easy to use. Disadvantages are license fees and the black-box character of the analyses. Free and open-source software solutions can be used independently of licence fees and provide full

transparency, but are partly command line based, which requires basic programming skills. Additionally, as the tools are sometimes the output of temporary scientific research projects, support is not always guaranteed in the long term.

Two commonly used commercial tools for *L. monocytogenes* typing are Ridom SeqSphere+ (Ridom GmbH) and BioNumerics (Applied Maths). Both contain their own cgMLST and wgMLST scheme by default. Ridom SeqSphere+ uses a cgMLST scheme containing 1701 genes [122], whereas BioNumerics uses a scheme containing 1748 genes [115]. An open-source alternative is chewBBACA (Blast-score-ratio-Based Allele Calling Algorithm) [123], into which a scheme can be fed at will. Besides the cgMLST analysis, BioNumerics can also be used for SNP analysis. An open-source alternative for SNP analysis, which is considered to be very reliable, is Snippy [124, 125]. There are also approaches for SNP analysis that do not require a reference genome [126], but the use of a reference genome is more common. Sequencing reads are mapped against this reference to identify SNPs. Both closed (complete) and draft genomes (assemblies) can be used as reference, but the degree of relationship of the reference genome to the strains under investigation makes a difference [127, 128]. The more similar two sequences are, the more reads can be mapped and the more nucleotides can be compared with each other. This increases the resolving power. Since strains are genetically very similar within an MLST CC, an MLST CC-specific reference genome can be used as a closely related reference genome for SNP analysis [127].

In our study [89], we compared commercial and non-commercial software solutions for cgMLST and SNP analysis using two different cgMLST schemes and three different kinds of reference genomes (Figure 9).

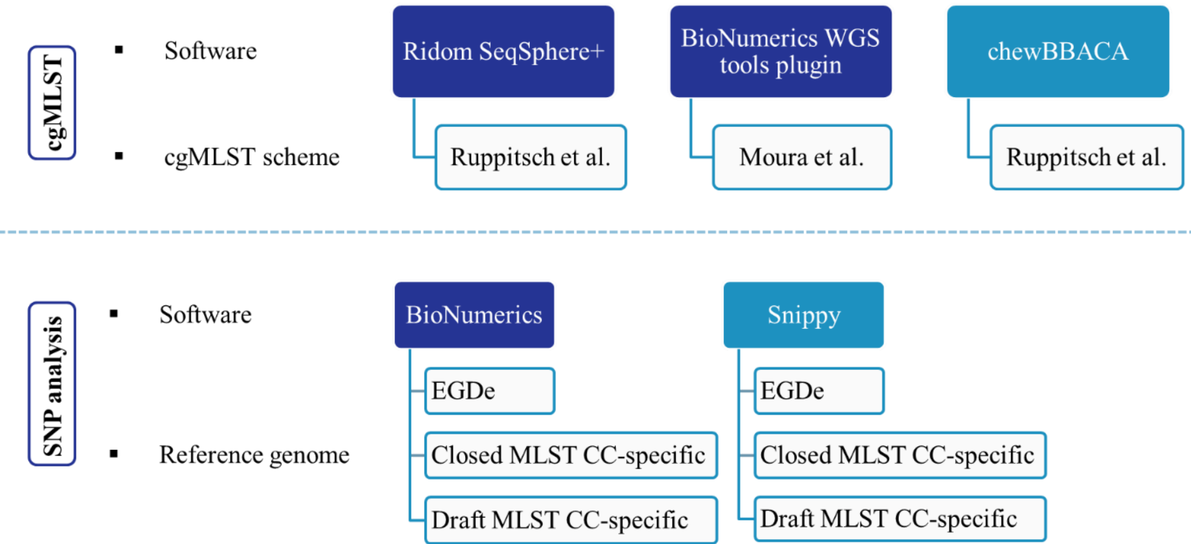


Figure 9: Comparison/Benchmarking of WGS-based typing methods. Dark blue: commercial software, light blue: non-commercial software. EGDe: general reference genome.

The distances between isolates were well matched in the different methods. The exception was the SNP analysis in BioNumerics with the general reference genome of the strain EGDe, whose distance values only showed a correlation of 0.77 to 0.85 to all other methods. Apart from this, the correlations of the distance values among tested methods were between 0.96 and 1 and as such near perfect. While it turned out that closed and draft genomes could be used equally as reference genomes in SNP analysis, it made a big difference whether a general or specific reference genome was used. With a specific reference genome, up to four times more SNPs could be found as with a general one. Additionally, using a specific reference genome, the differences between different software could be minimised. Not only for a higher accuracy, but also for a better method comparability it makes sense to use closely related rather than general reference genomes in SNP analysis.

Based on the good agreement between the distances, we were interested in finding out whether a clustering from the methods would also match well. This could serve as evidence that clustering information from different methods is interchangeable. The two cgMLST schemes were each published with an epidemiologically well-defined threshold value [115, 122]. Isolates falling within a cluster of such pre-defined thresholds are not beyond doubt, but likely linked epidemiologically [129]. The threshold published for the Ruppitsch scheme is ten allele differences between neighbouring isolates, the threshold for the Moura scheme is seven allele differences [115, 122]. In our study, clustering from the cgMLST methods with the different schemes was more consistent when using the same threshold than when using the scheme-specific values.

The different cgMLST approaches shown in Figure 9 at the two thresholds were used to establish a translation code with the other methods. This corresponds to the scenario where a particular laboratory sends a request related to a cluster that has been found with a particular method, in our case a cgMLST method. Now the primary laboratory would like to know if other laboratories using other methods (cgMLST or SNP) find more isolates that belong to the same cluster. This procedure is common practice for example in international listeriosis outbreak situations where one country needs to know whether outbreak-related strains are found in more countries. The other laboratories can then look up, at which threshold value of a comparison method (= their in-house method) the best possible cluster agreement to the reference method at the reference threshold can be achieved. Depending on the percentage of cluster agreement, clusters can then be communicated more or less seamlessly between reference and comparison method. With a threshold of seven allele differences as reference, achievable cluster agreement between the different combinations of our tested methods was between 90.9 and 100%, with threshold of ten allele differences as reference between 77.9 and 100%. Of course, the percentage of cluster match must always be kept in mind to evaluate the probability that two clusters from two methods really do match. Nevertheless, overall high agreement between the different combinations of methods tested opens up the possibility of exchanging cluster information without harmonisation of methods.

5.2 Applications of WGS-based typing

With WGS, a typing method has been found that allows bacterial transmission chains to be traced with unprecedented precision. In principle, there are two approaches how to use this information. One approach is backward looking and aims at tracing reported listeriosis cases to their source of infection. The other approach aims at preventing listeriosis cases before they even occur. Especially for the latter approach, hygiene measures during food processing are of fundamental importance.

5.2.1 WGS typing helps to improve process hygiene

A proof of how much WGS-based typing can help to identify entry points and transmission routes in food processing environments has been published as part of this work [90]. Such knowledge forms the basis for improving hygiene measures to such an extent that *Listeria* spp. cannot contaminate food and reach the consumer.

Publication: Contamination pathways can be traced along the poultry processing chain by whole genome sequencing of *Listeria innocua*

Article type: Research article

Journal: Microorganisms

Impact Factor: 4.15

Authors: Mayada Gwida†, Stefanie Lüth†, Maged El-Ashker, Amira Zakaria, Fatma El-Gohary, Mona Elsayed, Sylvia Kleta, Sascha Al Dahouk

† - Co-First authors

DOI: https://doi.org/10.3390/microorganisms8030414

Contributions Stefanie Lüth:

Based on the generated data and literature research, I created the concept for the article. I evaluated the sequencing data (SNP analysis and screening for virulence genes) and created figures based on these data. I described and discussed all the results concerning sequencing and placed them in the overall context of the study. I also took care of making the sequence data available in public repositories. As first and corresponding author, I submitted the article and considered and responded to the reviewer comments on behalf of all co-authors.

Listeria innocua is highly similar to the pathogenic species *L. monocytogenes*, it only lacks the typical virulence genes [130]. We used the non-pathogenic species *L. innocua* as an indicator organism simulating persistence and transmission of the pathogen *L. monocytogenes*. Due to the high similarity of *Listeria* spp. in terms of distribution and adaptability, the presence of the non-pathogenic *L. innocua* in processing lines and foodstuffs indicates poor hygiene conditions that favour the occurrence of the pathogen *L. monocytogenes*. Using the example of a poultry processing chain, we demonstrated the transmission of *L. innocua* from the farm level to the final food, proven by very small SNP distances between bacterial isolates. Positive cloacal swabs identified chickens as a natural reservoir. *Listeria innocua* was also found in the surrounding environment, especially on the walls of the farm and in soiled litter of the stables. Entry points to the slaughterhouse therefore appeared to be animal raw material and environmental contamination. WGS does not provide information about the direction of the contamination, but the fact that the same strain (less than 10 SNPs) originating from the farm was also found on a table in the slaughterhouse as well as in the finished chicken filet, speaks for cross-contamination during processing. It is a sanitary problem if *Listeria* spp. are found on surfaces in the food processing industry, since cross-contamination during processing is an important risk factor for food safety [131]. To prevent transmission of *Listeria* spp. to a germ-free food, contamination and persistence in the processing environment must be prevented. The other way around, when dealing with contaminated raw material, it is important to ensure that the contamination does not spread within the production environment and is eliminated from the food before consumption.

Listeria monocytogenes contamination can be controlled at two different levels in food industry: in the production facility or in the final food product. Good hygienic practices including regular cleaning and disinfection allow for the production of safe foods. Apart from classical biocides, a variety of different approaches can be applied for this purpose. Biological disinfectants like lytic bacteriophages (for example Listex^{TM}) cannot only be used in final foodstuffs, but also for elimination of biofilms in food processing environments. Their advantage is that, in contrast to classical biocides, which can pose a chemical risk, no food safety concerns are related to their use during processing [132]. However, bacteria may develop resistances, which impede the effectiveness of these measures [133]. Physical disinfection of processing environments and food can be achieved through the application of cold plasma, UV light or ozone fumigation [134-136]. Measures to eliminate contamination in the final food product include in-package thermal pasteurisation, irradiation, food preservatives (either synthetic antimicrobials or biopreservatives like lactic acid bacteria, bacteriocins or plant extracts) and high-pressure processing [137]. However, some of those treatments might negatively affect texture and taste and are therefore not applicable to every food product. Additionally, with the exception of pasteurisation, none of the methods is capable of eliminating *L. monocytogenes*; the pathogen can only be reduced to a lower level. In practice, often a combination of preservative factors is applied to ensure microbial food safety [138].

Viable bacteria may still remain in the food matrix and can multiply during shelf life, exceeding the limits set in Regulation (EC) No 2073/2005.

5.2.2 WGS in outbreak clarifications

When it is too late and listeriosis cases have already been reported, one of the most important actions is to trace these infections to their source. This is the only way to eliminate the cause to prevent further cases.

Listeriosis outbreaks are often protracted, geographically and temporally spread and mostly cases appear without obvious epidemiological link (e.g. [91, 96, 97]). This combination makes them difficult to detect for public health authorities. The typing of *L. monocytogenes* isolates from patients is therefore essential for the detection of outbreaks. In this ways, seemingly sporadic cases that might be widely distributed over time and space can be linked. The genetic relationship between clinical isolates consequently should lead to the search for a common food source. In the example of an epidemiological curve shown in Figure 10, listeriosis cases of a single outbreak were spread over six years and 12 German federal states. Without molecular typing, it would have been difficult to link these cases.

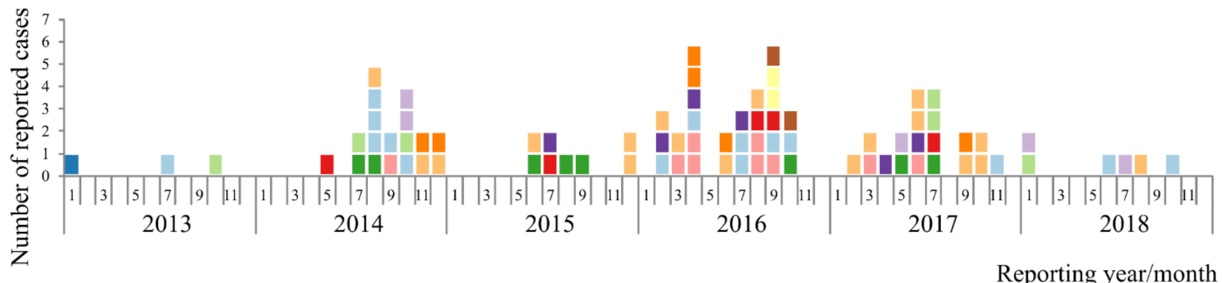


Figure 10: Epidemiological curve of a listeriosis outbreak coloured by the German federal states that notified the cases. Each box represents an individual case. Modified from [91].

Since the introduction of WGS, more and smaller listeriosis outbreaks have been clarified in the United States [139]. In line with this, the numbers of listeriosis cases that could be linked to a food source have risen sharply [118]. In Germany, the use of WGS also enabled the clarification of previously unexplained outbreaks, some of which were very prolonged or particularly large [91, 97, 140, 141].

An example of one of the largest listeriosis outbreaks described since the introduction of WGS-based typing in Germany was published as part of this thesis, showing the course of an outbreak investigation [91]. Only through a finely tuned interaction between molecular typing, classical epidemiology and operational testing together with intersectoral collaboration, the outbreak source could be identified.

Publication: Backtracking and forward checking of human listeriosis clusters identified a multiclonal outbreak linked to *Listeria monocytogenes* in meat products of a single producer

Article type: Research article

Journal: Emerging Microbes & Infections

Impact Factor: 5.78

Authors: Stefanie Lüth, Sven Halbedel, Bettina Rosner, Hendrik Wilking, Alexandra Holzer, Alice Roedel, Ralf Dieckmann, Sziliva Vincze, Rita Prager, Antje Flieger, Sascha Al Dahouk, Sylvia Kleta

DOI: https://doi.org/10.1080/22221751.2020.1784044

Contributions Stefanie Lüth:

I have grown and sequenced some of the non-clinical *L. monocytogenes* strains and monitored the data contributed by the co-authors involved (sequence data of clinical isolates, epidemiological data on disease cases). Based on the aggregated data of all partners, I developed the concept for the study. I evaluated the sequencing data (cgMLST analysis and screening for resistance and virulence genes) and created figures based on these data. I wrote the first draft except for a few small sections and made the sequence data available in public repositories. As first and corresponding author, I submitted the article and considered and responded to the reviewer comments on behalf of all co-authors.

Close collaboration between public health and the food sector is necessary to investigate a foodborne disease outbreak (Figure 11). In Germany, the common procedure for clarifying a listeriosis outbreak starts with the detection of a cluster of clinical cases by the public health authorities [40]. This is followed by a request to the food control authorities to find matching isolates from food and/or food processing environments.

	Listeriosis patients	***L. monocytogenes*-positive food or food processing environment**
Molecular surveillance	• Typing of isolates	• Typing of isolates
Epidemiology	• Analysis of reported data • Patient interviews	• Analysis of production and distribution chains • Sampling at producers

Figure 11: Overview of the concerted interaction between the public health and the food sector for listeriosis outbreak investigations.

The kick-off for the clarification of the exemplary outbreak [91] was a cgMLST-based match between a clinical isolate from a listeriosis cluster and a *L. monocytogenes* food isolate. Another isolate, from the same type of food but sampled in an earlier year, also matched. The fact that both food isolates came from the same manufacturer was the first epidemiological confirmation of the typing results. To obtain further epidemiological evidence, patient interviews about their consumption habits were conducted. The results further confirmed the suspicion that the consumption of contaminated RTE meat products was indeed the cause of listeriosis cases. The fact that one of the matching food isolates could be dated back to 2014 and the other one to 2016 indicated that the outbreak strain had been in the production facility for several years. In order to find out whether the strain was still present there and, above all, where exactly, extensive sampling of the manufacturing plant was carried out. From 2017 to 2018, over 200 *L. monocytogenes* isolates were obtained from the production environment and from food. Based on their typing results, an interesting situation arose. Usually, public health authorities search for causative food based on human disease cases (backtracking). However, when looking for matches of the newly sampled and typed isolates from the producer to previously unsolved human clusters, a match with a second cluster of clinical isolates was found. Forward checking therefore proved to be extremely useful. Only through such a two-sided control strategy combining backtracking and forward checking, outbreaks can not only be clarified, but essentially minimised. The two clusters of clinical isolates were genetically distant from each other and, without the epidemiological evidence, would have probably never been assigned to the same outbreak. This underlines the necessity to combine genomic and classical epidemiology. This example proves that it is not sufficient to selectively remove the source of a specific outbreak. The entire company hygiene has to be rethought to sustainably prevent further listeriosis cases. The fact that the listeriosis cases originating from a single food processing plant were distributed throughout Germany (cases in 12 of 16 federal states) clearly shows the extent to which such measures can have a positive influence on overall national health.

For the first outbreak cluster, the entire transmission chain could be traced from the processing environment over the food product to the diseased consumer. For the second cluster, the food product was missing in the chain. Both transmission paths show that cross-contamination between food and production environment must have played a fundamental role in this outbreak. Indeed, isolates of the same cluster were detected over four years, which proves the persistence of strains at the production site and explains why it is possible that an outbreak of listeriosis can spread over such a long time. The *L. monocytogenes* isolates sampled at the producer appeared susceptible to the biocides that were used to clean and disinfect the processing environment, but were mostly hidden in hard-to-clean niches/areas like conveyor belts, pulleys, freezers, accompanying parts like condensate lines or cable ducts or gullies. If a RTE food product comes into direct or indirect contact with such contaminated sites after the bactericidal processing step (e.g. heating), there is a high probability that this contamination will reach the consumer. It is therefore essential to avoid cross-contamination at this point more than at any other stage of the manufacturing process. In addition to pure industrial hygiene, the design of a processing plant also plays a decisive role in the production of safe food. For example, hard-to-clean spots should be avoided when planning workflows and equipment.

Besides the two outbreak-associated clusters, there were 17 other cgMLST genotypes of isolates in the manufacturing plant. The high diversity shows the extent of the contamination problem. Twelve cgMLST genotypes were found only once and did not fall into a cluster with other isolates. The reason for this could be that a sample is only a snapshot and represents a small section of the entire food processing environment. On the other hand, this could also indicate the sporadic introduction of new strains from outside into the processing plant, for instance via raw materials. This is exactly what we observed in our other study along the farm-to-fork continuum [90]. In fact, outbreak investigation must not end at the production level. The entire food chain needs to be considered so that the problem can be tackled at its roots and its outgrowths. Software has been developed to reconstruct food distribution networks [142]. However, compiling the relevant data for this purpose remains a challenge. The distribution channels are superregional and often international. In addition to economic interests, national data protection regulations can conflict with data aggregation.

5.2.3 WGS for risk assessment

Besides the determination of relationships, WGS is also very useful for the characterisation of bacterial isolates. It is possible to investigate the presence and absence of certain genes but also mutations within them. The most common applications are the detection of antimicrobial resistance (AMR) and virulence genes. The presence of specific AMR genes can be an indication of a reduced susceptibility of a bacterial isolate to antibiotics or disinfectants. This is an important aspect in the selection of appropriate control measures, whether for the treatment of sick patients or for the effective elimination of *L. monocytogenes*

from production environments. However, there may be discrepancies between genotype and phenotype due to expression patterns, functionally similar genes etc. [143]. One example for such a scenario is the fosfomycin resistance gene *fosX,* which all *L. monocytogenes* strains naturally carry [144, 145]. Although the gene is present and leads to resistance against fosfomycin *in vitro*, the situation *in vivo* is different. Inside the human host, the bacterial sugar phosphate permease Hpt is activated through the bacterial regulator PrfA, which leads to uptake and hence susceptibility to fosfomycin regardless of the presence of the *fosX* gene [146, 147]. This example shows how important it is to know regulatory pathways and molecular mechanisms in detail in order to draw conclusions about phenotypic effects based on the presence or absence of individual genes. There are various resistance gene databases available for screening of bacterial sequences, for example CARD [148], ResFinder [149] or the NCBI resistance database [150]. Of course, genes in a query genome can only be found if they are deposited in the database and are available for matching. Therefore, the choice of database alone can influence the result. Additionally, there are no clear rules like the minimum inhibitory concentration (MIC) or the epidemiological cut-off value (ECOFF) for the phenotypic classification as susceptible or resistant based on pure gene presence or absence [151]. As a consequence, a combination of genotypic and phenotypic methods is regularly used to evaluate AMR. This approach was also chosen in the outbreak investigation described in section 5.2.2 [91]. Draft genomes of all study strains were screened for AMR genes. As expected, they all carried the genes *fosX* and *lin*, conferring resistance to fosfomycin and lincomycin, respectively, and naturally present in the species *L. monocytogenes* [145]. No additional resistance genes to antibiotics were found, which implies that the standard antibiotics used to treat listeriosis should be effective. However, in about 20% of the isolates, the benzalkonium chloride (BAC) tolerance genes *bcrB* and *bcrC* [152] were identified. BAC is a broad-range biocide used as disinfectant or preservative [153]. The *bcrB-bcrC*-positive isolates were phenotypically tested [38] and could be confirmed as reduced susceptible to BAC by an increased MIC value. In this case, genotypic and phenotypic data matched. As none of the detergents or disinfectants used in the processing plant were based on BAC, this did not seem to have any practical implications on the persistence of *L. monocytogenes* in the production environment [91]. Nevertheless, such a procedure for AMR screening is important to allow for critical evaluation of existing industrial hygiene measures.

Another important issue is the evaluation of the virulence potential of *Listeria* spp. circulating within the food chain. To date, four different *Listeria* pathogenicity islands (LIPI) have been described, but only LIPI-1, -3 and -4 are found in the species *L. monocytogenes*. LIPI-2 is only present in *L. ivanovii* and thought to be involved into the pathogenic tropism of this species for ruminants [154]. LIPI-1 is composed of the six genes *prfA, plcA, hly, mpl, actA, plcB* that are essential for the intracellular life cycle of *L. monocytogenes* [5]. The *prfA* gene product is the master regulator of virulence-associated gene transcription and assumed to substantially control the switch between saprophytic and pathogenic lifestyle of *L. monocytogenes* [3, 155]. The gene products of *plcA* and *hly* are required for the escape

from vacuolar compartments, *actA* is needed for actin-based intracellular motility and *mpl* and *plcB* are necessary for cell-to-cell spread. Either LIPI-1 or the bacterial internalin gene *inlA* are found in atypical haemolytic *L. innocua* [156], suggesting that these genes fundamentally account for the pathogenicity of a strain or even species. In the *L. innocua* isolates of the above-mentioned study [90], no such virulence genes were found, indicating that they were indeed non-pathogenic. LIPI-3 describes the gene cluster of *llsAGHXBYDP* that encodes the cytolytic peptide listeriolysin S [157]. The release of this bacteriocin leads to modifications of the host microbiota during infection [158]. LlsA is the actual peptide, which is modified through the enzymes LlsB, Y and D, transported through LlsGH and cleaved by the protease LlsP. The function of LlsX is unknown. LIPI-4 contains six genes that encode a sugar transport system involved in neural and placental infection [159]. All three pathogenicity islands have been associated with hypervirulence of *L. monocytogenes* strains [159, 160]. In addition to those pathogenicity islands, a large number of individual other genes involved in the virulence of a strain have been described. An overview can be found in the virulence factor database, the VFDB [161]. An indication of virulence can also be obtained by the affiliation to a particular MLST CC [159]. Based on epidemiological data and demonstrated by infection experiments in a mouse model, MLST CC1, 2, 4 and 6 were classified as hypervirulent and infection-associated, whereas MLST CC9 and 121 were classified as hypovirulent and food-associated. It is to note that all this information provides indications, but does not allow for a yes or no decision on pathogenicity. The absence of pathogenicity islands or virulence factors by no means excludes the possibility that the pathogen causes listeriosis, like we have seen in the outbreak under investigation [91]. Although the outbreak strains were from neither an infection-associated MLST CC nor harboured LIPI-3 or LIPI-4, they had still caused a considerable number of disease cases. Conversely, a probably hypervirulent isolate according to molecular typing results does not necessarily have to be found in humans. Accordingly, strains were found in the production environment that belonged to infection-associated MLST CCs or contained LIPI-3, but could not be linked to a listeriosis case [91]. Therefore, risk assessment based on these characteristics was not possible and all strains of *L. monocytogenes* should be considered as posing a public health hazard.

5.3 Tools for risk communication

Due to the fact that food contamination is not a "one country" problem, communication tools for countries maintaining close trade relations are of crucial importance to allow for a coordinated response to emerging threats and for effective consumer protection. Where there is an interconnectedness of trade, interconnected food safety measures are needed. One valuable system for that purpose is the EU-wide Rapid Alert System for Food and Feed (RASFF). This reporting system is independent of isolate typing and instead focuses on the detection of contamination events in food products and their international distribution channels. The RASFF is based on Regulation (EC) 178/2002 and Regulation (EC) 16/2011 and operates as follows: a national food safety authority reports a food-related, serious, direct or indirect

risk for human health to the European Commission as the manager of the system. After verification, the information is distributed to all the other national contact points via different kinds of notifications (Figure 12). Depending on the type of notification and the distribution of a food product, immediate action (e.g. product recall) may also need to be taken in other countries.

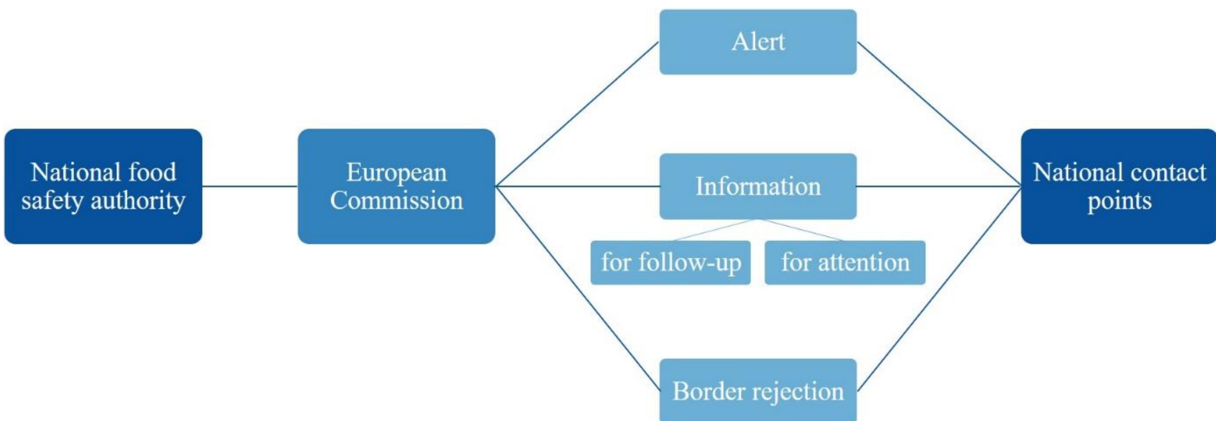


Figure 12: Procedure of a notification via the RASFF (from left to right).

The alert notification and information notification for follow-up are similar, but differ in the speed and kind of reaction that is needed. In the case of an alert notification, an immediate action is required while the information notification for follow-up does not require rapid action. An information notification for attention is released if a product is only present in the notifying country, no longer on the market or has not even been placed on the market. Nevertheless, it informs the other members that there was a risk posed by a contaminated food. A border rejection notification concerns a product that has been rejected at the external borders of the EU and EEA.

The study published as part of this thesis [92] is based on the analysis of RASFF notifications on pathogenic microorganisms in food products with a special focus on *L. monocytogenes*. It assesses trends in the reports available, identifies shortcomings in the current notification system and proposes options for improvement. Although a solid basis for the surveillance of *L. monocytogenes*, optimisations in the RASFF could help to identify food safety risks much faster, improve the quality of food products that enter the market and thus prevent foodborne infections more effectively.

Publication: Analysis of RASFF notifications on food products contaminated with *Listeria monocytogenes* reveals options for improvement in the rapid alert system for food and feed

Article type: Research article

Journal: Food Control

Impact Factor: 4.26

Authors: Stefanie Lüth, Idesbald Boone, Sylvia Kleta, Sascha Al Dahouk

DOI: https://doi.org/10.1016/j.foodcont.2018.09.033

Contributions Stefanie Lüth:

I analysed the extracted data, evaluated them with descriptive statistics, visualised the results and wrote the first draft. As first and corresponding author, I submitted the article and considered and responded to the reviewer comments on behalf of all co-authors.

From 2001 to 2015, 226 RASFF notifications (between 4 and 29 per year, median 16) on food products contaminated with *L. monocytogenes* that affected Germany were released. The products concerned came from 15 different EU member states and five non-European countries. This shows the close international interdependence of food trade and underlines the need for a common alert system like the RASFF. Furthermore, some of the products concerned in the notifications had more than one country of origin, reflecting the complexity of the tracking of commodity chains. The proportions of notifications relating to milk and milk products, fish and fish products and meat and meat products other than poultry were generally highest, although the exact percentages varied from year to year. For half of the notifications concerning milk and milk products, France was the country of origin and most of the time also France reported these products. For just over a third of the notifications concerning fish and fish products, Poland was the country of origin, although Germany reported the majority of them. The country of origin for the meat and meat products was diverse. Notifications concerning milk and milk products were mainly a result of company's own checks and for fish and fish products mainly a result of official controls. Both for the reporting country and for the reporting system, the principle of "The sooner, the better" applies. First, notifying country should match the country of origin of a product to the largest possible extent, since this can prevent contaminated products from entering international distribution in the first place. Second, company's own checks have a clear advantage over official controls because they are more often carried out before the products are placed on the market and potentially come into contact with consumers. In response to several pre-millennial food safety crises,

a White Paper on Food Safety was published [162], which aims at giving more responsibility to food producers to guarantee effective controls right at the beginning of the food chain. An increasing number of RASFF notifications and an increasing trend of company's own checks rather than official controls can be seen as a result of this White Paper, increasing awareness and expanded efforts. In addition to official controls, it is essential that food business operators implement appropriate measures (such as self-monitoring) on their own responsibility to ensure compliance with food safety. In the case of the outbreak described above [91], for instance, a very large number of samples were collected in company's own checks after a first match between a clinical and a food *L. monocytogenes* isolate sampled in an official control. This thorough sampling activity has played a major role to fully clarify the outbreak.

The analysis of RASFF notifications also reflects the importance of industrial hygiene in controlling the spread of *L. monocytogenes*. Some companies were involved in notifications in different years. This is an indication either of persistence within the production environment or recurring re-entries. Especially in the case of fish and fish products, only few companies were involved in many RASFF notifications. This demonstrates how a point source, with large-scale distribution channels, can affect overall international food safety. On a small scale (on a national level), this effect was also observed in the outbreak described in section 5.2.2. Contaminated products from a single manufacturing plant had caused listeriosis cases throughout Germany.

As with the typing/sequence databases, the usefulness of the RASFF could be enhanced by making more detailed metadata available. It is already possible to provide information about the processing state, packaging, slicing category, etc. together with the notification of a product. Such information could feed into improved risk assessment by answering questions such as "Which type of cutting is most commonly reported by RASFF notifications?" or similar. This would enable food safety authorities to identify and eliminate risk factors. Since there is no obligation to register this information, only few reports have been provided with a complete metadata set so far. The major reason for this is that the information is simply not always available and an obligation to complete metadata would hinder reporting activity. However, introduction of an explicit "Unspecified" statement could possibly increase the motivation to enter the data, if available.

6 Conclusions and future perspectives

The overall aim of this research work was to create the basis for an effective surveillance and control of *L. monocytogenes* in Germany together with the identification of starting points for the prevention and sustainable reduction of listeriosis cases.

The molecular typing of *L. monocytogenes* isolates from humans, food and production environments helps to better understand contamination and infection pathways. In 2016, WGS was introduced for typing of *L. monocytogenes* at BfR and since mid-2018, the former method for fine typing, PFGE, has been completely switched to WGS. In combination with the introduction of WGS-based surveillance in the public health sector [40], this was the cornerstone for the start of a success story in the clarification of listeriosis outbreaks in Germany. Previously unexplained outbreaks, some of them very long lasting and with large numbers of cases, have since been clarified and stopped [91, 97, 140, 141]. Within food processing plants, niches and hotspots of *Listeria* could be identified and effectiveness of corrective measures could be evaluated. Thus, WGS also makes a valuable contribution to improving industrial hygiene, which will prevent food contamination and subsequently foodborne diseases. For the first time since 2011, the numbers of listeriosis cases in Germany did not further increase in 2018 and 2019, but actually decreased. This may already be an effect of the improved molecular surveillance and control strategy established within this thesis and now more and more applied by food control authorities in Germany.

A very important aspect for the use of WGS in official monitoring is the existence of validated methods. In addition, there needs to be a way to share data between different laboratories that may use validated, but different methods. Through assessing the degree of comparability between most commonly used WGS approaches, we have found a workable interim solution to facilitate communication between sectors and countries even without having completed method standardisation [89]. Nevertheless, in the long run, harmonisation will be inevitable to increase efficiency and achieve a higher level of automation to deal with the ever-increasing amounts of sequencing data. First, a comprehensive national WGS database for *L. monocytogenes* from humans, food and food processing plants is needed. For this purpose, harmonisation of methods as well as the agreement on uniform evaluation criteria have to be established. This would then enable real-time surveillance instead of event-driven investigations. Matches between clinical and non-clinical isolates would become apparent from the moment the sequence of an isolate is fed into the database, which would automate and speed up the backtracking and forward checking strategy described within this thesis. In addition to the real-time sequencing of isolates required for this purpose, however, retrospective sequencing should not be disregarded in order to do justice to long-lasting outbreaks. Such a strategy might link sporadic cases of listeriosis to a food product before an outbreak occurs and could thus make an important contribution to reducing cases of listeriosis.

Based on WGS typing, a national distribution of *L. monocytogenes* strains became obvious and based on RASFF notifications affecting Germany, an international distribution pattern of contaminated food emerged [91, 92, 140]. This makes it quite clear that *L. monocytogenes* control is not a local problem but requires a national and international strategy. One market requires one database. National databases should finally be combined into an international one. In principle, such a database already exists in the form of the RASFF. However, the existing data do not comprise typing results. On the public health side, there is an analogous typing-independent system, the EU-wide Epidemic Intelligence Information System (EPIS) to exchange information on emerging public health threats. The interaction of all those partial solutions works, but is much more cumbersome than necessary and slows down the tracing of infections to their sources. The ideal case would be to merge the data from these more metadata-focused databases with the typing data. Due to data protection regulations, however, this will probably remain a pipe dream for the time being. Nevertheless, just an international database solution with a minimal set of metadata will be the first step in the right direction. Several initiatives have already started to establish and use such solutions, such as GenomeTrakr [163] or PulseNet [164]. However, there is still a long way to go to speak of comprehensive, global systems [165].

Any database solution assumes that cases of listeriosis have already occurred and must be linked to a food source, so the principle is looking backwards. Although this also plays a crucial role to reduce disease burden, an additional forward-looking strategy would prevent cases of listeriosis before they even occur. The basis for such an approach is a combination of risk assessment and risk management. As has been shown in this work, industrial hygiene plays a crucial role in the spread of *Listeria* spp. [90, 91]. Here, linking typing data and metadata as completely as possible would also be of great advantage. In a combined effort of extensive sampling and typing, *L. monocytogenes* hotspots in the processing chains can be identified and eliminated. A prerequisite for this is that results from company's own checks are also included in molecular surveillance, i.e. that companies also include their data in databases. This would avoid recurring contamination of food and thereby protect consumers. In addition to the identification of risk factors at the process level, attention should also be paid to the identification of risk factors at the molecular level. One means to gain deeper insights into foodborne zoonotic hazards could be genome-wide association studies that identify differences in the gene profiles of different groups of isolates. This may identify further genes that are crucial for causing especially serious human infections. Using genetic data, it is also possible to perform source attribution modelling, where human cases of disease are attributed to different food sources [166]. In this way, it is possible to break down exactly which food categories carry which risk potential. This will help to focus risk management even more effectively. Finally yet importantly, risk communication towards risk groups should also be optimised to increase the awareness in vulnerable population and those who provide food for them (e.g. hospitals, nursing homes, nursing relatives). For example, better education about the disease listeriosis and high-risk foods could help reduce consumption and thus the risk of infection. Only through the

interaction of all these factors, food and public health authorities will not only control but also sustainably combat listeriosis.

Probably the most obvious solution would be to prohibit any *L. monocytogenes* contamination in food. Such a zero-tolerance strategy as used in the United States is, however, controversial. In a risk modelling, Chen and colleagues [67] found that a 100 CFU/g limit, like applied in the EU, could already reduce the number of listeriosis infections in the United States by 99.5%. They therefore question the cost-benefit ratio of a zero-tolerance strategy, which is undoubtedly associated with enormous efforts in the food industry sector. The bottom line is to decide whether *L. monocytogenes* contamination is considered unavoidable (not added) or human added [167]; a question made almost impossible to answer by the ubiquitous nature of *L. monocytogenes*. What is certain is that everything must be done to ensure that contamination of food on the table is as low as possible, in order to protect consumers and thus human health in the best possible way.

7 Danksagung

Zuallererst gilt mein Dank dem BfR und meinen Betreuern Professor Dr. Sascha Al Dahouk und Dr. Sylvia Kleta für die Möglichkeit, an diesem spannenden und so wichtigen Projekt zu arbeiten. Ich danke den beiden für ihr großes Vertrauen und die Freiheiten, die sie mir dabei gelassen haben. Gleichzeitig konnte ich mich immer auf ihren Rückhalt und wertvollen und ehrlichen Rat verlassen. Ich schätze die Möglichkeiten und Chancen, die ich dadurch hatte, sehr. Ich danke Sascha Al Dahouk, dass er mich als Mentor stetig gefördert, aber auch ab und zu herausgefordert hat. Dass er dabei an mich geglaubt hat, hat mich ermutigt und mich mehr als einmal über mich selbst hinauswachsen lassen. Ich bin dankbar dafür, wie viel ich durch ihn lernen konnte. Besonders danken möchte ich Sylvia Kleta dafür, wie sie mich mit ihrer Begeisterung für die Thematik angesteckt und meine Arbeit mit neuen Ideen bereichert und so weitergebracht hat. Die gemeinsam gefeierten Erfolge waren für mich außerdem immer wieder eine große Motivation.

Professor Dr. Jens Rolff danke ich für die Übernahme der universitätsinternen Betreuung meiner Arbeit.

Mein Dank gilt außerdem meinen Kollegen am BfR, die mir mit Rat und Tat zur Seite standen. Ich bedanke mich besonders bei meinem Team aus der 47 für die herzliche Aufnahme, die offene und sehr konstruktive Zusammenarbeit und die Unterstützung im Labor. In guter Erinnerung bleiben mir auch die gemeinsamen Unternehmungen außerhalb der Arbeit, sowie die Momente, in denen das Lachen beim Mittagessen mal wieder aus dem Ruder lief. Außerdem bedanke ich mich bei meinen Kollegen aus 4SZ für die Unterstützung bei der bioinformatischen Datenauswertung und die inspirierenden Gespräche, ohne die die Arbeit nicht zu der geworden wäre, die sie jetzt ist.

Danken möchte ich auch dem Bundesministerium für Gesundheit und der Nationalen Forschungsplattform für Zoonosen für die Förderung des MolTypList-Projekts, aus dem der größte Teil dieser Arbeit entstanden ist, sowie den daran beteiligten Kooperationspartnern vom RKI und der FU Berlin.

Mein größter Dank gilt meiner Familie und besonders meinen Eltern für das immer offene Ohr und die wertvollen Ratschläge. Vor allem bin ich ihnen dankbar für ihre liebevolle und bedingungslose Unterstützung. Ich danke ihnen, dass sie es mir ermöglicht haben, meinen Weg frei zu wählen und dafür, dass sie dabei immer hinter mir stehen. So kann ich alles schaffen.

Ich danke auch meiner „Zweitfamilie“ und den weiteren wichtigen Menschen, die oft schon seit so vielen Jahren an meiner Seite stehen, das Glück mit mir teilen und mich auch bei Schwierigkeiten aufbauen. Sie erden mich, sind mein Gute Laune Pol, wichtige Berater, Motivationscoach und eine große Stütze. Besonders danke ich ihnen dafür, dass sie auch in meiner Promotionszeit mitgefiebert und deren Nebenwirkungen mit so viel Verständnis, Geduld und einer großen Portion Humor begleitet haben.

Danke!

8 References

1. Gray ML, Killinger A. *Listeria monocytogenes* and listeric infections. Bacteriological Reviews. 1966;30(2):309.

2. Farber J, Peterkin P. *Listeria monocytogenes*, a food-borne pathogen. Microbiological Reviews. 1991;55(3):476-511.

3. Freitag NE, Port GC, Miner MD. *Listeria monocytogenes* - from saprophyte to intracellular pathogen. Nature Reviews Microbiology. 2009;7(9):623-8.

4. Orsi RH, Wiedmann M. Characteristics and distribution of *Listeria* spp., including *Listeria* species newly described since 2009. Applied Microbiology and Biotechnology. 2016;100(12):5273-87.

5. Vázquez-Boland JA, Kuhn M, Berche P, Chakraborty T, Domínguez-Bernal G, Goebel W, et al. *Listeria* pathogenesis and molecular virulence determinants. Clinical Microbiology Reviews. 2001;14(3):584-640.

6. Guillet C, Join-Lambert O, Le Monnier A, Leclercq A, Mechaï F, Mamzer-Bruneel MF, et al. Human listeriosis caused by *Listeria ivanovii*. Emerging Infectious Diseases. 2010;16(1):136-8.

7. Perrin M, Bemer M, Delamare C. Fatal case of *Listeria innocua* bacteremia. Journal of Clinical Microbiology. 2003;41(11):5308-9.

8. Favaro M, Sarmati L, Sancesario G, Fontana C. First case of *Listeria innocua* meningitis in a patient on steroids and eternecept. JMM Case Reports. 2014;1(2).

9. Leclercq A, Moura A, Vales G, Tessaud-Rita N, Aguilhon C, Lecuit M. *Listeria thailandensis* sp. nov. International Journal of Systematic and Evolutionary Microbiology. 2019;69(1):74-81.

10. Doijad SP, Poharkar KV, Kale SB, Kerkar S, Kalorey DR, Kurkure NV, et al. *Listeria goaensis* sp. nov. International Journal of Systematic and Evolutionary Microbiology. 2018;68(10):3285-91.

11. Núñez-Montero K, Leclercq A, Moura A, Vales G, Peraza J, Pizarro-Cerdá J, et al. *Listeria costaricensis* sp. nov. International Journal of Systematic and Evolutionary Microbiology. 2018;68(3):844-50.

12. Quereda JJ, Leclercq A, Moura A, Vales G, Gómez-Martín Á, García-Muñoz Á, et al. *Listeria valentina* sp. nov., isolated from a water trough and the faeces of healthy sheep. International Journal of Systematic and Evolutionary Microbiology. 2020;70(11):5868-79.

13. Murray EGD, Webb RA, Swann MBR. A disease of rabbits characterised by a large mononuclear leucocytosis, caused by a hitherto undescribed bacillus *Bacterium monocytogenes* (n.sp.). The Journal of Pathology and Bacteriology. 1926;29(4):407-39.

14. Pirie JHH. *Listeria*: Change of Name for a Genus *Bacteria*. Nature. 1940;145(3668):264.

15. Hoelzer K, Pouillot R, Dennis S. Animal models of listeriosis: a comparative review of the current state of the art and lessons learned. Veterinary Research. 2012;43(1):18.

16. Esteban JI, Oporto B, Aduriz G, Juste RA, Hurtado A. Faecal shedding and strain diversity of *Listeria monocytogenes* in healthy ruminants and swine in Northern Spain. Veterinary Research. 2009;5:2.

17. Ho AJ, Ivanek R, Gröhn YT, Nightingale KK, Wiedmann M. *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. Preventive Veterinary Medicine. 2007;80(4):287-305.

18. Dhama K, Karthik K, Tiwari R, Shabbir MZ, Barbuddhe S, Malik SVS, et al. Listeriosis in animals, its public health significance (food-borne zoonosis) and advances in diagnosis and control: a comprehensive review. Veterinary Quarterly. 2015;35(4):211-35.

19. Sanaa M, Poutrel B, Menard JL, Serieys F. Risk factors associated with contamination of raw milk by *Listeria monocytogenes* in dairy farms. Journal of Dairy Science. 1993;76(10):2891-8.

20. Skovgaard N, Nørrung B. The incidence of *Listeria* spp. in faeces of Danish pigs and in minced pork meat. International Journal of Food Microbiology. 1989;8(1):59-63.

21. Kanuganti SR, Wesley IV, Reddy PG, McKean J, Hurd HS. Detection of *Listeria monocytogenes* in pigs and pork. Journal of Food Protection. 2002;65(9):1470-4.

22. Gray M. Listeriosis in fowls: a review. Avian Diseases. 1958;2(3):296-314.

23. Jami M, Ghanbari M, Zunabovic M, Domig KJ, Kneifel W. *Listeria monocytogenes* in Aquatic Food Products - A Review. Comprehensive Reviews in Food Science and Food Safety. 2014;13(5):798-813.

24. McLauchlin J, Low J. Primary cutaneous listeriosis in adults: an occupational disease of veterinarians and farmers. Veterinary Record. 1994;135(26):615-7.

25. Swaminathan B, Gerner-Smidt P. The epidemiology of human listeriosis. Microbes and Infection. 2007;9(10):1236-43.

26. Al-Ghazali M, Al-Azawi SK. *Listeria monocytogenes* contamination of crops grown on soil treated with sewage sludge cake. Journal of Applied Bacteriology. 1990;69(5):642-7.

27. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2019. 2020.

28. Conlan JW, North RJ. Early pathogenesis of infection in the liver with the facultative intracellular bacteria *Listeria monocytogenes*, *Francisella tularensis*, and *Salmonella typhimurium* involves lysis of infected hepatocytes by leukocytes. Infection and Immunity. 1992;60(12):5164.

29. Disson O, Lecuit M. Targeting of the central nervous system by *Listeria monocytogenes*. Virulence. 2012;3(2):213-21.

30. Bakardjiev AI, Theriot JA, Portnoy DA. *Listeria monocytogenes* traffics from maternal organs to the placenta and back. PLoS Pathogens. 2006;2(6):e66.

31. Lamont RF, Sobel J, Mazaki-Tovi S, Kusanovic JP, Vaisbuch E, Kim SK, et al. Listeriosis in human pregnancy: a systematic review. Journal of Perinatal Medicine. 2011;39(3):227-36.

32. de Noordhout CM, Devleesschauwer B, Angulo FJ, Verbeke G, Haagsma J, Kirk M, et al. The global burden of listeriosis: a systematic review and meta-analysis. Lancet Infectious Diseases. 2014;14(11):1073-82.

33. Goulet V, King LA, Vaillant V, de Valk H. What is the incubation period for listeriosis? BMC Infectious Diseases. 2013;13:11.

34. Buchanan RL, Gorris LGM, Hayman MM, Jackson TC, Whiting RC. A review of *Listeria monocytogenes*: An update on outbreaks, virulence, dose-response, ecology, and risk assessments. Food Control. 2017;75:1-13.

35. Hoelzer K, Chen Y, Dennis S, Evans P, Pouillot R, Silk BJ, et al. New Data, Strategies, and Insights for *Listeria monocytogenes* Dose-Response Models: Summary of an Interagency Workshop, 2011. Risk Analysis. 2013;33(9):1568-81.

36. Robert Koch Institute. Listeriose RKI-Ratgeber 2010 [Available from: https://www.rki.de/DE/Content/Infekt/EpidBull/Merkblaetter/Ratgeber_Listeriose.html.

37. Noll M, Kleta S, Al Dahouk S. Antibiotic susceptibility of 259 *Listeria monocytogenes* strains isolated from food, food-processing plants and human samples in Germany. Journal of Infection and Public Health. 2018;11(4):572-7.

38. Roedel A, Dieckmann R, Brendebach H, Hammerl JA, Kleta S, Noll M, et al. Biocide tolerant *Listeria monocytogenes* isolates from German food production plants do not show cross-resistance to clinically relevant antibiotics. Applied and Environmental Microbiology. 2019:AEM.01253-19.

39. Fischer MA, Wamp S, Fruth A, Allerberger F, Flieger A, Halbedel S. Population structure-guided profiling of antibiotic resistance patterns in clinical *Listeria monocytogenes* isolates from

Germany identifies pbpB3 alleles associated with low levels of cephalosporin resistance. Emerging Microbes & Infections. 2020;9(1):1804-13.

40. Halbedel S, Prager R, Fuchs S, Trost E, Werner G, Flieger A. Whole genome sequencing of recent *Listeria monocytogenes* isolates from Germany reveals population structure and disease clusters. Journal of Clinical Microbiology. 2018;56(6).

41. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2017. 2018.

42. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2011. 2012.

43. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2018. 2019.

44. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2001. 2002.

45. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2002. 2003.

46. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2003. 2004.

47. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2004. 2005.

48. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2005. 2006.

49. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2006. 2007.

50. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2007. 2008.

51. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2008. 2009.

52. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2009. 2010.

53. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2010. 2011.

54. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2012. 2013.

55. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2013. 2014.

56. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2014. 2015.

57. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2015. 2016.

58. Robert Koch Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2016. 2017.

59. EFSA BIOHAZ Panel, Ricci A, Allende A, Bolton D, Chemaly M, Davies R, et al. *Listeria monocytogenes* contamination of ready-to-eat foods and the risk for human health in the EU. EFSA Journal. 2018;16(1):e05134.

60. EFSA, ECDC. The European Union One Health 2018 Zoonoses Report. EFSA Journal. 2019;17(12).

61. Lungu B, Ricke S, Johnson M. Growth, survival, proliferation and pathogenesis of *Listeria monocytogenes* under low oxygen or anaerobic conditions: a review. Anaerobe. 2009;15(1):7-17.

62. Milillo SR, Friedly EC, Saldivar JC, Muthaiyan A, O'Bryan C, Crandall PG, et al. A Review of the Ecology, Genomics, and Stress Response of *Listeria innocua* and *Listeria monocytogenes*. Critical Reviews in Food Science and Nutrition. 2012;52(8):712-25.

63. Thévenot D, Dernburg A, Vernozy-Rozand C. An updated review of *Listeria monocytogenes* in the pork meat industry and its products. Journal of Applied Microbiology. 2006;101(1):7-17.

64. Vogel BF, Hansen LT, Mordhorst H, Gram L. The survival of *Listeria monocytogenes* during long term desiccation is facilitated by sodium chloride and organic material. International Journal of Food Microbiology. 2010;140(2):192-200.

65. Ryser ET, Marth EH. *Listeria*, listeriosis, and food safety. CRC Press. 2007.

66. COMMISSION REGULATION (EC) No 2073/2005 of 15 November 2005 on microbiological criteria for foodstuffs, (2005).

67. Chen Y, Ross WH, Scott VN, Gombas DE. *Listeria monocytogenes*: low levels equal low risk. Journal of Food Protection. 2003;66(4):570-7.

68. Tompkin R. Control of *Listeria monocytogenes* in the food-processing environment. Journal of Food Protection. 2002;65(4):709-25.

69. Rothrock MJ, Jr., Davis ML, Locatelli A, Bodie A, McIntosh TG, Donaldson JR, et al. *Listeria* Occurrence in Poultry Flocks: Detection and Potential Implications. Frontiers in Veterinary Science. 2017;4:125.

70. Gómez D, Iguácel LP, Rota MC, Carramiñana JJ, Ariño A, Yangüela J. Occurrence of *Listeria monocytogenes* in Ready-to-Eat Meat Products and Meat Processing Plants in Spain. Foods. 2015;4(3):271-82.

71. Rodríguez-López P, Rodríguez-Herrera JJ, Vázquez-Sánchez D, López Cabo M. Current Knowledge on *Listeria monocytogenes* Biofilms in Food-Related Environments: Incidence, Resistance to Biocides, Ecology and Biocontrol. Foods. 2018;7(6):85.

72. Carpentier B, Cerf O. Persistence of *Listeria monocytogenes* in food industry equipment and premises. International Journal of Food Microbiology. 2011;145(1):1-8.

73. Renier S, Hébraud M, Desvaux M. Molecular biology of surface colonization by *Listeria monocytogenes*: an additional facet of an opportunistic Gram-positive foodborne pathogen. Environmental Microbiology. 2011;13(4):835-50.

74. Rieu A, Lemaître J-P, Guzzo J, Piveteau P. Interactions in dual species biofilms between *Listeria monocytogenes* EGD-e and several strains of *Staphylococcus aureus*. International Journal of Food Microbiology. 2008;126(1):76-82.

75. Doijad SP, Barbuddhe SB, Garg S, Poharkar KV, Kalorey DR, Kurkure NV, et al. Biofilm-Forming Abilities of *Listeria monocytogenes* Serotypes Isolated from Different Sources. PLOS ONE. 2015;10(9):e0137046-e.

76. Ferreira V, Wiedmann M, Teixeira P, Stasiewicz MJ. *Listeria monocytogenes* persistence in food-associated environments: epidemiology, strain characteristics, and implications for public health. Journal of Food Protection. 2014;77(1):150-70.

77. Gandhi M, Chikindas ML. *Listeria*: a foodborne pathogen that knows how to survive. International Journal of Food Microbiology. 2007;113(1):1-15.

78. Orsi RH, Borowsky ML, Lauer P, Young SK, Nusbaum C, Galagan JE, et al. Short-term genome evolution of *Listeria monocytogenes* in a non-controlled environment. BMC Genomics. 2008;9(1):539.

79. Hartung M. Erreger von Zoonosen in Deutschland im Jahr 2008 – Mitteilungen der Länder zu Lebensmitteln, Tieren, Futtermitteln und Umweltproben. BfR Wissenschaft. 2010.

80. Hartung M, Käsbohrer A. Erreger von Zoonosen in Deutschland im Jahr 2009. BfR Wissenschaft. 2011.

81. Hartung M, Käsbohrer A. Erreger von Zoonosen in Deutschland im Jahr 2010. BfR Wissenschaft. 2012.

82. Hartung M, Käsbohrer A. Erreger von Zoonosen in Deutschland im Jahr 2011. BfR Wissenschaft. 2013.

83. Hartung M, Tenhagen B-A, Käsbohrer A. Erreger von Zoonosen in Deutschland im Jahr 2012. BfR Wissenschaft. 2014.

84. Hartung M, Tenhagen B-A, Käsbohrer A. Erreger von Zoonosen in Deutschland im Jahr 2013. BfR Wissenschaft. 2015.

85. Hartung M, Tenhagen B-A, Alt K, Käsbohrer A. Erreger von Zoonosen in Deutschland im Jahr 2014. BfR Wissenschaft. 2016.

86. Hartung M, Tenhagen B-A, Alt K, Käsbohrer A. Erreger von Zoonosen in Deutschland im Jahr 2015. BfR Wissenschaft. 2018.

87. Hartung M, Alt K, Käsbohrer A, Tenhagen B-A. Erreger von Zoonosen in Deutschland im Jahr 2016. BfR Wissenschaft. 2019.

88. Lüth S, Kleta S, Al Dahouk S. Whole genome sequencing as a typing tool for foodborne pathogens like *Listeria monocytogenes* - The way towards global harmonisation and data exchange. Trends in Food Science & Technology. 2018;73:67-75.

89. Lüth S, Deneke C, Kleta S, Al Dahouk S. Translatability of WGS typing results can simplify data exchange for surveillance and control of *Listeria monocytogenes*. Microbial Genomics. 2020; mgen000491.

90. Gwida M, Lüth S, El-Ashker M, Zakaria A, El-Gohary F, Elsayed M, et al. Contamination Pathways can Be Traced along the Poultry Processing Chain by Whole Genome Sequencing of *Listeria innocua*. Microorganisms. 2020;8(3):414.

91. Lüth S, Halbedel S, Rosner B, Wilking H, Holzer A, Roedel A, et al. Backtracking and forward checking of human listeriosis clusters identified a multiclonal outbreak linked to *Listeria monocytogenes* in meat products of a single producer. Emerging Microbes & Infections. 2020;9(1):1600-8.

92. Lüth S, Boone I, Kleta S, Al Dahouk S. Analysis of RASFF notifications on food products contaminated with *Listeria monocytogenes* reveals options for improvement in the rapid alert system for food and feed. Food Control. 2019;96:479-87.

93. Angelo KM, Conrad AR, Saupe A, Dragoo H, West N, Sorenson A, et al. Multistate outbreak of *Listeria monocytogenes* infections linked to whole apples used in commercially produced, prepackaged caramel apples: United States, 2014-2015. Epidemiology and Infection. 2017;145(5):848-56.

94. Chen Y, Luo Y, Curry P, Timme R, Melka D, Doyle M, et al. 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. 2017;12(2):e0171389.

95. EFSA, ECDC. Multi-country outbreak of *Listeria monocytogenes* serogroup IV b, multi-locus sequence type 6, infections linked to frozen corn and possibly to other frozen vegetables–first update. EFSA Supporting Publications. 2018;15(7):1448E.

96. Ruppitsch W, Prager R, Halbedel S, Hyden P, Pietzka A, Huhulescu S, et al. Ongoing outbreak of invasive listeriosis, Germany, 2012 to 2015. Eurosurveillance. 2015;20(50).

97. Kleta S, Hammerl JA, Dieckmann R, Malorny B, Borowiak M, Halbedel S, et al. Molecular Tracing to Find Source of Protracted Invasive Listeriosis Outbreak, Southern Germany, 2012–2016. Emerging Infectious Diseases. 2017;23(10):1680.

98. Seeliger HPR, Höhne K. Chapter II Serotyping of *Listeria monocytogenes* and Related Species. In: Bergan T, Norris JR, editors. Methods in Microbiology. Volume 13: Academic Press; 1979. p. 31-49.

99. Loessner MJ, Busse M. Bacteriophage typing of *Listeria* species. Applied and Environmental Microbiology. 1990;56(6):1912-8.

100. Bibb WF, Gellin BG, Weaver R, Schwartz B, Plikaytis BD, Reeves MW, et al. Analysis of clinical and food-borne isolates of *Listeria monocytogenes* in the United States by multilocus enzyme electrophoresis and application of the method to epidemiologic investigations. Applied and Environmental Microbiology. 1990;56(7):2133-41.

101. Liu D. Identification, subtyping and virulence determination of *Listeria monocytogenes*, an important foodborne pathogen. Journal of Medical Microbiology. 2006;55(6):645-59.

102. Brosch R, Buchrieser C, Rocourt J. Subtyping of *Listeria monocytogenes* serovar 4b by use of low-frequency-cleavage restriction endonucleases and pulsed-field gel electrophoresis. Research in Microbiology. 1991;142(6):667-75.

103. Doumith M, Buchrieser C, Glaser P, Jacquet C, Martin P. Differentiation of the major *Listeria monocytogenes* serovars by multiplex PCR. Journal of Clinical Microbiology. 2004;42(8):3819-22.

104. Glaser P, Frangeul L, Buchrieser C, Rusniok C, Amend A, Baquero F, et al. Comparative genomics of *Listeria* species. Science. 2001;294(5543):849-52.

105. Félix B, Danan C, Van Walle I, Lailler R, Texier T, Lombard B, et al. Building a molecular *Listeria monocytogenes* database to centralize and share PFGE typing data from food, environmental and animal strains throughout Europe. Journal of Microbiological Methods. 2014;104:1-8.

106. Félix B, Niskanen T, Vingadassalon N, Dao TT, Assere A, Lombard B, et al. Pulsed-field gel electrophoresis proficiency testing trials: toward European harmonization of the typing of food and clinical strains of *Listeria monocytogenes*. Foodborne Pathogens and Disease. 2013;10(10):873-81.

107. Michelon D, Felix B, Vingadassalon N, Mariet JF, Larsson JT, Moller-Nielsen E, et al. PFGE standard operating procedures for *Listeria monocytogenes*: harmonizing the typing of food and clinical strains in Europe. Foodborne Pathogens and Disease. 2015;12(3):244-52.

108. ECDC, EFSA, van Walle I, Guerra B, Borges V, André Carriço J, et al. EFSA and ECDC technical report on the collection and analysis of whole genome sequencing data from food-borne pathogens and other relevant microorganisms isolated from human, animal, food, feed and food/feed environmental samples in the joint ECDC-EFSA molecular typing database. EFSA Supporting Publications. 2019;16(5):1337E.

109. Salcedo C, Arreaza L, Alcala B, De La Fuente L, Vazquez J. Development of a multilocus sequence typing method for analysis of *Listeria monocytogenes* clones. Journal of Clinical Microbiology. 2003;41(2):757-62.

110. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences. 1977;74(12):5463-7.

111. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature. 2005;437(7057):376-80.

112. Loman NJ, Misra RV, Dallman TJ, Constantinidou C, Gharbia SE, Wain J, et al. Performance comparison of benchtop high-throughput sequencing platforms. Nature Biotechnology. 2012;30(5):434-9.

113. Richterich P. Estimation of errors in "raw" DNA sequences: a validation study. Genome Research. 1998;8(3):251-9.

114. Nagarajan N, Pop M. Sequence assembly demystified. Nature Reviews Genetics. 2013;14(3):157-67.

115. Moura A, Criscuolo A, Pouseele H, Maury MM, Leclercq A, Tarr C, et al. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. Nature Microbiology. 2016;2:16185.

116. Stasiewicz MJ, Oliver HF, Wiedmann M, den Bakker HC. Whole-genome sequencing allows for improved identification of persistent *Listeria monocytogenes* in food-associated environments. Applied and Environmental Microbiology. 2015;81(17):6024-37.

117. World Health Organization. Whole genome sequencing for foodborne disease surveillance: landscape paper. 2018.

118. Jackson BR, Tarr C, Strain E, Jackson KA, Conrad A, Carleton H, et al. Implementation of Nationwide Real-time Whole-genome Sequencing to Enhance Listeriosis Outbreak Detection and Investigation. Clinical Infectious Diseases. 2016;63(3):380-6.

119. Aarestrup FM. Integrating Genome-based Informatics to Modernize Global Disease Monitoring, Information Sharing, and Response. Emerging Infectious Diseases. 2012;8(11).

120. Wang Y, Pettengill JB, Pightling AW, Timme R, Allard M, Strain E, et al. Genetic Diversity of *Salmonella* and *Listeria* Isolates from Food Facilities. Journal of Food Protection. 2018;81(12):2082-9.

121. Pavlopoulos GA, Soldatos TG, Barbosa-Silva A, Schneider R. A reference guide for tree analysis and visualization. BioData Mining. 2010;3:1.

122. Ruppitsch W, Pietzka A, Prior K, Bletz S, Fernandez HL, Allerberger F, et al. Defining and Evaluating a Core Genome Multilocus Sequence Typing Scheme for Whole-Genome Sequence-Based Typing of *Listeria monocytogenes*. Journal of Clinical Microbiology. 2015;53(9):2869-76.

123. Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J, Santos S, et al. chewBBACA: A complete suite for gene-by-gene schema creation and strain identification. Microbial Genomics. 2018;4(3).

124. Bush SJ, Foster D, Eyre DW, Clark EL, De Maio N, Shaw LP, et al. Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism–calling pipelines. GigaScience. 2020;9(2).

125. Seemann T. Snippy - Rapid haploid variant calling and core genome alignment 2015 [Available from: https://github.com/tseemann/snippy.

126. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics. 2015;31(17):2877-8.

127. Møller Nielsen E, Björkman JT, Kiil K, Grant K, Dallman T, Painset A, et al. 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. 2017;14(2).

128. Kwong JC, Mercoulia K, Tomita T, Easton M, Li HY, Bulach DM, et al. Prospective Whole-Genome Sequencing Enhances National Surveillance of *Listeria monocytogenes*. Journal of Clinical Microbiology. 2016;54(2):333-42.

129. Pightling AW, Pettengill JB, Luo Y, Baugher JD, Rand H, Strain E. Interpreting Whole-Genome Sequence Analyses of Foodborne Bacteria for Regulatory Applications and Outbreak Investigations. Frontiers in Microbiology. 2018;9:1482.

130. Buchrieser C, Rusniok C, Consortium L, Kunst F, Cossart P, Glaser P. Comparison of the genome sequences of *Listeria monocytogenes* and *Listeria innocua*: clues for evolution and pathogenicity. FEMS Immunology & Medical Microbiology. 2003;35(3):207-13.

131. Hoelzer K, Pouillot R, Gallagher D, Silverman MB, Kause J, Dennis S. Estimation of *Listeria monocytogenes* transfer coefficients and efficacy of bacterial removal through cleaning and sanitation. International Journal of Food Microbiology. 2012;157(2):267-77.

132. EFSA Panel on Biological Hazards. Evaluation of the safety and efficacy of Listex™ P100 for reduction of pathogens on different ready-to-eat (RTE) food products. EFSA Journal. 2016;14(8):e04565.

133. Moye ZD, Woolston J, Sulakvelidze A. Bacteriophage Applications for Food Production and Processing. Viruses. 2018;10(4):205.

134. Timmons C, Pai K, Jacob J, Zhang G, Ma LM. Inactivation of *Salmonella enterica*, Shiga toxin-producing *Escherichia coli*, and *Listeria monocytogenes* by a novel surface discharge cold plasma design. Food Control. 2018;84:455-62.

135. Montgomery NL, Banerjee P. Inactivation of *Escherichia coli* O157:H7 and *Listeria monocytogenes* in biofilms by pulsed ultraviolet light. BMC Research Notes. 2015;8(1):235.

136. Nicholas R, Dunton P, Tatham A, Fielding L. The effect of ozone and open air factor on surface-attached and biofilm environmental *Listeria monocytogenes*. Journal of Applied Microbiology. 2013;115(2):555-64.

137. Zhu M, Du M, Cordray J, Ahn DU. Control of *Listeria monocytogenes* Contamination in Ready-to-Eat Meat Products. Comprehensive Reviews in Food Science and Food Safety. 2005;4(2):34-42.

138. Leistner L. Basic aspects of food preservation by hurdle technology. International Journal of Food Microbiology. 2000;55(1):181-6.

139. CDC. The Listeria Whole Genome Sequencing Project 2016 [Available from: https://www.cdc.gov/listeria/surveillance/whole-genome-sequencing.html.

140. Halbedel S, Wilking H, Holzer A, Kleta S, Fischer M, Lüth S, et al. Large Nationwide Outbreak of Invasive Listeriosis Associated with Blood Sausage, Germany, 2018–2019. Emerging Infectious Diseases. 2020;26(7):1456.

141. Lachmann R, Halbedel S, Adler M, Becker N, Allerberger F, Holzer A, et al. Nationwide outbreak of invasive listeriosis associated with consumption of meat products in health care facilities, Germany, 2014-2019. Clinical Microbiology and Infection. 2020;0(0).

142. Weiser AA, Thoens C, Falenski A, Appel B, Filter M, Kaesbohrer A. FoodChain-lab: Tracing Software Supporting Foodborne Disease Outbreak Investigations. Procedia Food Science. 2016;7:101-4.

143. Painset A, Björkman JT, Kiil K, Guillier L, Mariet J-F, Félix B, et al. LiSEQ - whole-genome sequencing of a cross-sectional survey of *Listeria monocytogenes* in ready-to-eat foods and human clinical cases in Europe. Microbial Genomics. 2019;5(2).

144. Fillgrove KL, Pakhomova S, Schaab MR, Newcomer ME, Armstrong RN. Structure and mechanism of the genomically encoded fosfomycin resistance protein, FosX, from *Listeria monocytogenes*. Biochemistry. 2007;46(27):8110-20.

145. Troxler R, von Graevenitz A, Funke G, Wiedemann B, Stock I. Natural antibiotic susceptibility of *Listeria* species: *L. grayi*, *L. innocua*, *L. ivanovii*, *L. monocytogenes*, *L. seeligeri* and *L. welshimeri* strains. Clinical Microbiology and Infection. 2000;6(10):525-35.

146. Scortti M, Lacharme-Lora L, Wagner M, Chico-Calero I, Losito P, Vázquez-Boland JA. Coexpression of virulence and fosfomycin susceptibility in *Listeria*: molecular basis of an antimicrobial in vitro-in vivo paradox. Nature Medicine. 2006;12(5):515-7.

147. Scortti M, Han L, Alvarez S, Leclercq A, Moura A, Lecuit M, et al. Epistatic control of intrinsic resistance by virulence genes in *Listeria*. PLoS Genetics. 2018;14(9):e1007525.

148. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Research. 2020;48(D1):D517-25.

149. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. Journal of Antimicrobial Chemotherapy. 2020;75(12):3491-500.

150. Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, et al. RefSeq: an update on prokaryotic genome annotation and curation. Nucleic Acids Research. 2018;46(D1):D851-60.

151. Ellington MJ, Ekelund O, Aarestrup FM, Canton R, Doumith M, Giske C, et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. Clinical Microbiology and Infection. 2017;23(1):2-22.

152. Elhanafi D, Dutta V, Kathariou S. Genetic Characterization of Plasmid-Associated Benzalkonium Chloride Resistance Determinants in a *Listeria monocytogenes* Strain from the 1998-1999 Outbreak. Applied and Environmental Microbiology. 2010;76(24):8231-8.

153. Merchel Piovesan Pereira B, Tagkopoulos I. Benzalkonium Chlorides: Uses, Regulatory Status, and Microbial Resistance. Applied and Environmental Microbiology. 2019;85(13):e00377-19.

154. Domínguez-Bernal G, Müller-Altrock S, González-Zorn B, Scortti M, Herrmann P, Monzó HJ, et al. A spontaneous genomic deletion in *Listeria ivanovii* identifies LIPI-2, a species-specific pathogenicity island encoding sphingomyelinase and numerous internalins. Molecular Microbiology. 2006;59(2):415-32.

155. de las Heras A, Cain RJ, Bielecka MK, Vazquez-Boland JA. Regulation of *Listeria* virulence: PrfA master and commander. Current Opinion in Microbiology. 2011;14(2):118-27.

156. Moura A, Disson O, Lavina M, Thouvenot P, Huang L, Leclercq A, et al. Atypical hemolytic *Listeria innocua* are virulent, albeit less than *Listeria monocytogenes*. Infection and Immunity. 2019:IAI.00758-18.

157. Cotter PD, Draper LA, Lawton EM, Daly KM, Groeger DS, Casey PG, et al. Listeriolysin S, a Novel Peptide Haemolysin Associated with a Subset of Lineage I *Listeria monocytogenes*. PLoS Pathogens. 2008;4(9):e1000144.

158. Quereda JJ, Meza-Torres J, Cossart P, Pizarro-Cerdá J. Listeriolysin S: A bacteriocin from epidemic *Listeria monocytogenes* strains that targets the gut microbiota. Gut Microbes. 2017;8(4):384-91.

159. Maury MM, Tsai YH, Charlier C, Touchon M, Chenal-Francisque V, Leclercq A, et al. Uncovering *Listeria monocytogenes* hypervirulence by harnessing its biodiversity. Nature Genetics. 2016;48(3):308-13.

160. Vilchis-Rangel RE, Espinoza-Mellado MDR, Salinas-Jaramillo IJ, Martinez-Pena MD, Rodas-Suarez OR. Association of *Listeria monocytogenes* LIPI-1 and LIPI-3 marker llsX with invasiveness. Current Microbiology. 2019;76(5):637-43.

161. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Research. 2019;47(D1):D687-92.

162. European Commission. White Paper on food safety 2000.

163. Timme RE, Sanchez Leon M, Allard MW. Utilizing the Public GenomeTrakr Database for Foodborne Pathogen Traceback. Foodborne Bacterial Pathogens. Humana, New York, NY, 2019; 201-12.

164. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, et al. PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Eurosurveillance. 2017;22(23):30544.

165. Brown E, Dessai U, McGarry S, Gerner-Smidt P. Use of Whole-Genome Sequencing for Food Safety and Public Health in the United States. Foodborne Pathogens and Disease. 2019;16(7):441-50.

166. Pires SM, Evers EG, van Pelt W, Ayers T, Scallan E, Angulo FJ, et al. Attributing the human disease burden of foodborne infections to specific sources. Foodborne Pathogens and Disease. 2009;6(4):417-24.

167. Archer DL. The evolution of FDA's policy on *Listeria monocytogenes* in ready-to-eat foods in the United States. Current Opinion in Food Science. 2018;20:64-8.

9 Appendix: Publications and supplementary material

The publications are listed in the order in which they were presented in this work.

(1) Lüth S, Kleta S, Al Dahouk S. Whole genome sequencing as a typing tool for foodborne pathogens like *Listeria monocytogenes* – The way towards global harmonisation and data exchange. Trends in Food Science & Technology. 2018;73:67-75.

https://doi.org/10.1016/j.tifs.2018.01.008

(2) Lüth S, Deneke C, Kleta S, Al Dahouk S. Translatability of WGS typing results can simplify data exchange for surveillance and control of *Listeria monocytogenes*. Microbial Genomics. 2020; mgen000491.

https://doi.org/10.1099/mgen.0.000491

- Supplement 1, 2 and 3
- Supplementary Table 1, 2 and 3

(3) Gwida M, Lüth S, El-Ashker M, Zakaria A, El-Gohary F, Elsayed M, et al. Contamination Pathways can Be Traced along the Poultry Processing Chain by Whole Genome Sequencing of *Listeria innocua*. Microorganisms. 2020;8(3):414.

https://doi.org/10.3390/microorganisms8030414

(4) Lüth S, Halbedel S, Rosner B, Wilking H, Holzer A, Roedel A, et al. Backtracking and forward checking of human listeriosis clusters identified a multiclonal outbreak linked to *Listeria monocytogenes* in meat products of a single producer. Emerging Microbes & Infections. 2020;9(1):1600-8.

https://doi.org/10.1080/22221751.2020.1784044

- Supplementary Table 1

(5) Lüth S, Boone I, Kleta S, Al Dahouk S. Analysis of RASFF notifications on food products contaminated with Listeria monocytogenes reveals options for improvement in the rapid alert system for food and feed. Food Control. 2019;96:479-87.

https://doi.org/10.1016/j.foodcont.2018.09.033

Trends in Food Science & Technology 73 (2018) 67–75

Contents lists available at ScienceDirect

Trends in Food Science & Technology

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

Review

Whole genome sequencing as a typing tool for foodborne pathogens like *Listeria monocytogenes* – The way towards global harmonisation and data exchange

Stefanie Lüth[a,*], Sylvia Kleta[a], Sascha Al Dahouk[a,b]

[a] German Federal Institute for Risk Assessment (BfR), National Reference Laboratory for Listeria monocytogenes, Department of Biological Safety, Diedersdorfer Weg 1, 12277 Berlin, Germany

[b] RWTH Aachen University Hospital, Pauwelsstraße 30, 52074 Aachen, Germany

A B S T R A C T

Background: Various molecular typing methods are used for the surveillance of foodborne pathogens and outbreak investigations, differing widely in information content and discriminatory power. Presently, not least because of the rapid technological development, the focus is shifting to whole genome sequencing (WGS) as an analytical tool. As a result of globalisation of food trade, a comprehensive understanding of the association between the occurrence of human infections and causative pathogens has to be established to monitor and to prevent their spread. In this respect, the accuracy of WGS clearly supersedes that of previous tools.

Scope and approach: Our review describes the status quo of WGS in surveillance and outbreak investigations of foodborne pathogens through the example of *Listeria monocytogenes*. It highlights the value of WGS in trace-back of infections to food sources and provides an overview of methods used for data generation (wet lab) and analysis (dry lab). Altogether, progress but also challenges in the worldwide practical application of WGS for bacterial typing are described.

Key findings and conclusions: The current status of WGS differs widely between countries and even laboratory sites. A consensus has to be found concerning methods, quality measures, thresholds for data generation and analysis as well as rules for data sharing. International harmonisation is going to be indispensable on the way to data exchangeability which will finally support global control of foodborne pathogens.

1. Introduction

The gram positive bacterium *Listeria monocytogenes* (*L. monocytogenes*) is the causative agent of the infectious disease listeriosis in humans. Although it is widely distributed in the environment, transmission of *L. monocytogenes* to humans mainly occurs via consumption of contaminated food, especially pre-packaged ready-to-eat products. Its ability to form biofilms paired with its resilience to salt, low temperatures and acidic environments enables survival or growth even in preserved and chilled food products rendering *L. monocytogenes* a serious foodborne pathogen (Swaminathan & Gerner-Smidt, 2007).

In 2015, a total of 2,206 human listeriosis cases were reported in the EU, corresponding to an incidence rate of 0.46 per 100,000 population (EFSA & ECDC, 2016). In spite of the low incidence of listeriosis, hospitalisation rates above 90% and mortalities of 20–30% make the disease a serious public health concern. Infection of otherwise healthy adults is rare, mostly leading to non-invasive, mild listeriosis (febrile gastroenteritis) or even absence of symptoms (Aureli et al., 2000). However, cases may accumulate in risk groups, including elderly, pregnant or immunocompromised patients. Then listeriosis can be an invasive disease associated with septicaemia or focal complications such as encephalitis and meningitis (Vázquez-Boland et al., 2001).

Globalisation of food trade, changing consumption habits towards highly processed foods and demographic changes with increase of susceptible populations have augmented risk of foodborne illnesses (Wang et al., 2016). As a result from the high case fatality rate, listeriosis is a notifiable disease in the vast majority of EU member states and associated countries. Occurrence of disease can either be sporadic or outbreak-related. Since foodborne outbreaks are of public health relevance and also cause tremendous economic losses e.g. due to product recall, internationally cross-linked surveillance of *L. monocytogenes* in humans and food is of crucial importance to identify clusters, trace the sources of infections and control outbreaks.

In order to identify epidemiologically linked isolates and thus be

* Corresponding author.
E-mail address: stefanie.lueth@bfr.bund.de (S. Lüth).

https://doi.org/10.1016/j.tifs.2018.01.008
Received 22 August 2017; Received in revised form 4 December 2017; Accepted 18 January 2018
Available online 06 February 2018
0924-2244/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

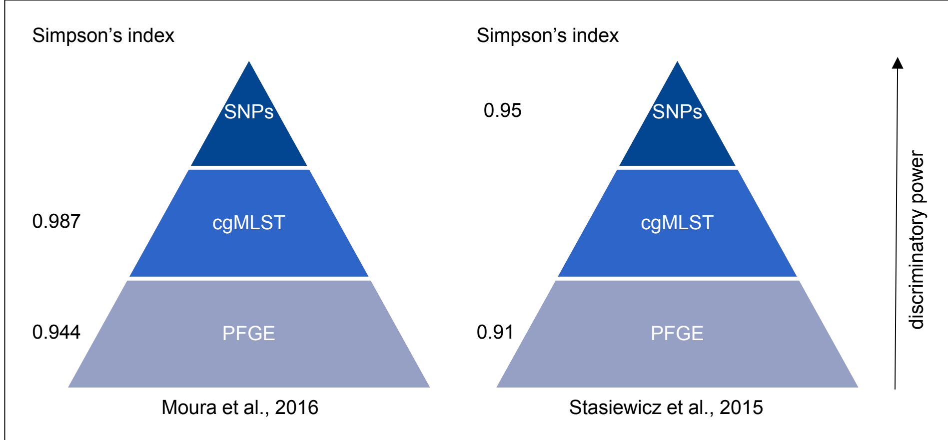


Fig. 1. Comparison of Simpson's indices of WGS-based cgMLST and SNP analysis, and PFGE.
Discriminatory power of WGS-based typing methods exceeds that of PFGE typing as shown by comparison of Simpson's indices. Studies were conducted analysing 100 *L. monocytogenes* isolates (Moura et al., 2016) and 188 *L. monocytogenes* isolates (Stasiewicz et al., 2015), respectively.

able to determine outbreak strains, differentiation of *L. monocytogenes* beyond the species level is indispensable. Consequently, several molecular typing methods have been established that are able to characterise, discriminate and index subtypes of microorganisms (ECDC, 2013). The first subtyping method applied to *L. monocytogenes* was serotyping (Seeliger & Höhne, 1979). Allocation to serotypes is still the initial step to roughly classify *L. monocytogenes* strains, although nowadays implemented by identification of serotype-specific genes through multiplex PCR assay rather than by agglutination (Doumith, Buchrieser, Glaser, Jacquet, & Martin, 2004). In the meantime, a great variety of typing methods have been developed, differing widely according to their discriminatory power, reproducibility, hands-on time and cost. In general, two different typing approaches exist, based either on phenotypic or on genotypic characteristics. Typing and identification methods based on the phenotype include e.g. classical serotyping with antisera, phage typing, multi-locus enzyme electrophoresis or esterase typing. A major weakness of these phenotypic methods is that some captured properties may vary in response to external stress or in dependence of the bacterial growth phase (Liu, 2006). To overcome these difficulties, several DNA-based methods have been established (Brosch, Buchrieser, & Rocourt, 1991). These genotypic methods aim to investigate DNA fragment length polymorphisms of restriction fragments (e.g. by ribotyping, pulsed-field gel electrophoresis), amplified DNA fragments (e.g. by random amplification of polymorphic DNA, repetitive element-PCR) or both (e.g. by amplified fragment length polymorphism), or polymorphisms that are directly found in the nucleotide sequence (e.g. by multi-locus sequence typing, multiple-locus variable number tandem repeat analysis, DNA microarray, whole genome sequencing). Among these techniques, due to its high discriminatory power, a good database and high level of standardisation, pulsed-field gel electrophoresis (PFGE) is often considered as gold standard for subtyping of *L. monocytogenes*. PFGE relies on the analysis of fragments generated by DNA digest with rarely cutting restriction enzymes, in the case of *L. monocytogenes* with *Asc*I and *Apa*I, and their separation on an agarose gel applying a periodically changing electrical field. Comparison of restriction patterns provides information on the relatedness of different isolates. Harmonisation of experimental procedures and data analysis through the European Union Reference Laboratory for L. *monocytogenes* (EURL *Lm*) and the PulseNet International Network USA as well as set-up of centralised databases like the EURL *Lm* database emphasise the fundamental importance of PFGE for routine surveillance of *L. monocytogenes* (Félix, Danan, Van Walle, et al., 2014; Félix et al., 2013; Graves & Swaminathan, 2001). However, PFGE is a rather time-consuming and labour-intensive technique. Furthermore, discriminatory power of PFGE profiles is limited as only nucleotide changes in the restriction enzyme recognition sites are detected. Consequently, relatedness of strains may be over- or underestimated and epidemiologically unrelated isolates may be assigned to one 'pseudo'-cluster whereas even highly related strains fall into distinct clusters.

Thanks to recent developments in next generation sequencing technologies (NGS), whole-genome sequencing (WGS) as a typing tool for *L. monocytogenes* and other foodborne pathogens is gaining in importance (Deng, den Bakker & Hendriksen, 2016; Gilchrist, Turner, Riley, Petri, & Hewlett, 2015). Sequencing of entire bacterial genomes provides an unparalleled depth of information. Base-by-base comparisons of entire genomes are possible as well as retrieval of additional information such as virulence or antimicrobial resistance markers. As opposed to traditional molecular methods like PFGE, WGS provides comprehensive insight into evolution of bacterial strains.

Comparing Simpson's Indices of molecular typing methods, discriminatory power of WGS-based typing clearly exceeds that of PFGE. The Simpson's Index is used to quantify the probability that two unrelated strains are assigned to different typing groups (Hunter & Gaston, 1988). Although the Simpson's Index only yields study-specific values as it depends on the number of identified types and of the isolates that fall into these types, it is a valuable tool for a relative, quantitative evaluation of the discriminatory power of typing methods (Fig. 1). In a study including 188 *L. monocytogenes* isolates from 30 retail delis in three U.S. states over 2 years, Simpson's Index for WGS-based single nucleotide polymorphism (SNP)-based subtyping was 0.95, compared to 0.91 for PFGE (Stasiewicz, Oliver, Wiedmann, & den Bakker, 2015). In a comparison between WGS-based core genome multi-locus sequence typing (cgMLST) and PFGE, Simpson's Indices of 0.987 and 0.944 were calculated based on the analysis of 100 isolates (Moura et al., 2016). In both cases unprecedented precision of outbreak investigations became possible using WGS.

Our reviews aims to demonstrate the value of WGS especially compared to PFGE in the field of foodborne pathogen surveillance through introduction into methodical aspects and presentation of application examples. We intend to provide an application-oriented overview on different approaches towards global data exchangeability and challenges involved, considering differences between the EU, the USA but also international initiatives.

2. Outbreak investigations using WGS for typing

In the USA, WGS was set up for *L. monocytogenes* outbreak investigations by the Centers for Disease Control and Prevention (CDC) in cooperation with the Food and Drug Administration (FDA), the National Institutes of Health (NIH) and the United States Department of Agriculture (USDA) in September 2013 with the long-term aim to completely supersede the pool of other typing techniques including PFGE (Jackson et al., 2016). Several independent studies have provided evidence of the usefulness and value of WGS in outbreak investigations compared to other molecular typing methods (see below).

In an outbreak related to contaminated ice cream, listeriosis cases first accumulated in a single hospital in Kansas from 2014 to 2015 (CDC., 2015a). Although isolates from two of the affected patients shared a common PFGE pattern, patterns for three other patients did not, suggesting independence of cases. However, this assumption was rejected when cgMLST based on WGS data identified four of the five isolates as highly related and thus allowed their link to the outbreak. SNP-based analysis of the WGS data confirmed this attribution by proving a difference of 1–19 SNPs between clinical isolates (Chen et al., 2017). In another outbreak in 2015, WGS analysis did not only allow cluster identification that was not possible via PFGE typing, but also enabled retrospective inclusion of a previously unsolved cluster from 2013 into the outbreak and trace-back to contaminated soft cheese (CDC., 2015c). Overall, routine use of WGS for typing of *L. monocytogenes* isolates from clinical and food samples in the USA has proven undoubtedly successful. The resolution of WGS exceeds the discriminatory power of PFGE and provides more precise and reliable data. As a result, smaller outbreaks can be recognised that would have otherwise been considered as sporadic cases. Furthermore, retrospective analysis allows grouping of individual sporadic cases over a longer period of time to one single outbreak and enables to link outbreak strains to a common source. Consecutive regulatory steps like product recalls or controlled sanitation of production plants can then prevent further listeriosis cases. Since its implementation for routine surveillance in the USA in 2013, WGS typing helped solving a variety of food-related listeriosis outbreaks and also to identify uncommon sources in a listeriosis outbreak, 2014–2015, linked to pre-packaged caramel apples (CDC., 2015b). Overall, identification of more outbreaks with fewer cases per outbreak becomes possible using WGS-based typing (CDC., 2016).

In the EU and associated countries, as opposed to the USA, WGS has not yet entered the status of a comprehensive routine typing method for *L. monocytogenes*. Between 2013 and 2015, twelve to fifteen outbreaks of listeriosis per year have been reported in the EU (EFSA & ECDC, 2016). The vast majority of them were resolved using traditional typing techniques like PFGE in combination with epidemiological evidence. However, to date, several exemplary studies investigating foodborne outbreaks by NGS techniques have been performed and published (Gillesberg Lassen et al., 2016; Kleta et al., 2017; Kvistholm Jensen et al., 2016; Ruppitsch, Prager, et al., 2015; Schmid et al., 2014).

In 2014, within an international collaboration of public health institutes and food authorities, a cluster of seven human listeriosis cases in Germany and Austria that emerged between April 2011 and July 2013 could be identified (Schmid et al., 2014). Initially, the respective outbreak strains were typed with PFGE and fluorescent amplified fragment length polymorphism (fAFLP) where they appeared indistinguishable and were assigned to one cluster. CgMLST based on WGS data, however, was capable to distinguish a cluster of four outbreak strains (≤6 allelic differences) isolated in 2012–2013 from the other three strains isolated in 2011 (≥48 allelic differences) that could subsequently be excluded from the outbreak. In addition, the four confirmed outbreak cases could be traced back to two different Austrian food products, an unaged soft cheese (food isolates differing ≤19 alleles from the human cluster) and a deli-meat (food isolates differing ≤8 alleles from the human cluster). However, no final attribution could be made because thresholds for strain differentiation have not yet been defined. Recently, a genetic distance of ≤10 alleles between human and food isolates has been proposed for unambiguous source attribution which would exclude the soft cheese as a possible source (Ruppitsch, Pietzka, et al., 2015).

During a long-term outbreak of listeriosis in Southern Germany from 2012 to 2015, WGS and cgMLST were used to confirm clustering of human isolates with an unusual PFGE pattern into one outbreak group (Ruppitsch, Prager, et al., 2015). Although six food-related isolates from Austria and Germany showed PFGE patterns identical to the human isolates, WGS revealed their belonging to independent cluster types. Thus, faulty source attribution could be averted. Later, the human cases could be traced back to a contaminated batch of smoked pork belly (Kleta et al., 2017). This observation underlines the importance of WGS for successful and reliable trace-back of listeriosis cases to a food source. Nonetheless, although PFGE cannot keep pace with the discriminatory power of WGS data based analysis, it might still be a suitable alternative in countries or regions where NGS is not established because of economic reasons.

Among European countries, Denmark has officially initiated nationwide real-time WGS typing of human *L. monocytogenes* isolates for routine surveillance in September 2013 (Kvistholm Jensen et al., 2016). In addition, interviews exploring consumption habits of listeriosis patients have been implemented and added to the typing data since June 2014. So far, these combined databases have proven successful in two different outbreak investigations. In 2014, 41 listeriosis cases in Denmark were assigned to one outbreak cluster through WGS-based SNP analysis with genetic differences between the isolates not exceeding 3 SNPs (Kvistholm Jensen et al., 2016). In cooperation with the Danish Veterinary and Food Administration (DVFA) who provided data on routinely collected food samples, human strains could be traced back to a common source, a ready-to-eat delicatessen meat which was subsequently recalled from the national market. In this way, the outbreak could be terminated. Similarly, a total of twenty listeriosis cases notified between 2013 and 2015 in Denmark could be assigned to two distinct outbreaks, each comprising ten cases (Gillesberg Lassen et al., 2016). Both clusters could be traced back to smoked salmon or smoked halibut and trout. Again, WGS typing of human clinical isolates and routinely collected food isolates allowed reliable source attribution, enabled to impose legal measures and thereby saved lives.

3. Wet lab standardisation of NGS methods for foodborne pathogen typing

International collaboration on the control of foodborne pathogens like *L. monocytogenes* has more than ever become indispensable to guarantee food safety. For that purpose, harmonisation and standardisation of WGS-based typing methods across countries and sectors (human, animal, food) need to be established to ensure comparability of typing results and to allow data exchange. To date, a variety of protocols have been developed for typing of *L. monocytogenes* and other foodborne pathogens using NGS technologies. The current challenge lies in identification of differences and definition of generally valid quality metrics to produce consistent results.

Four main factors allow a statement on the quality of sequencing results: coverage or sequencing depth, evenness of coverage, read length and read quality (Loman, Misra, et al., 2012). Coverage describes the average number of times a genome has been sequenced. It is equal to the product of read length and number of reads divided by the haploid genome length (Lander & Waterman, 1988). Bases are usually sequenced multiple times to increase the probability that all genomic regions are covered and to compensate for possible sequencing errors in order to increase confidence in sequencing results. High coverage but also evenness of coverage is essential to be able to consider a sequencing run successful. Read quality can be assessed through Phred Quality or Q-scores. This score gives the logarithmic probability of an incorrect base (Richterich, 1998). For example Q30 represents the probability of one incorrect base in 1000. Its determination is based on the comparison of measurement parameters during base detection (e.g intensity profile, signal-to-noise ratio) with empirically determined reference parameters that are linked to known quality scores. It is to note that Q-scores are hence specific for a platform and even for new hardware, software or chemistry within a platform and are dependent on algorithms used to predict them. One major step towards global harmonisation would be the definition of general quality metrics.

Sequencing technologies have massively evolved during the last decades and are still in a process of continuous development and

Table 1
Overview of sequencers most frequently used for WGS of bacterial pathogens.

Sequencer	Provider	Scale	Technology	Data collection	Read length	Run time[a]
Second Generation Sequencers						
MiSeq	Illumina	benchtop	sequencing by synthesis	optical signal	1 × 36 bp - 2 × 300 bp	4 h - 56 h[b]
Ion Torrent PGM	Life Technologies	benchtop	semiconductor sequencing	pH change	200 - 400 bp	2 h - 7 h
HiSeq 2500	Illumina	production-scale	sequencing by synthesis	optical signal	1 × 36 bp - 2 × 250 bp	7 h (rapid run mode) - 11d [b]
Third Generation Sequencer						
PacBio RS	Pacific Biosciences	benchtop	single molecule real-time sequencing	fluorescence pulse	> 20 kb	0.5 h - 10 h

[a] Manufacturer specifications; depending on run mode, kit and read length.
[b] Includes time for cluster generation.

improvement, accompanied by a substantial cost reduction. Following the 'First Generation' Sanger sequencing (Sanger, Nicklen, & Coulson, 1977), introduction of massively parallel NGS or 'Second Generation' sequencing (SGS) in 2005 (Margulies et al., 2005) revolutionised genomics. Independent of the platform, SGS methods share three main steps to obtain raw sequence data: isolation of DNA, preparation of a sequencing library and sequencing. Methods used for the individual steps, however, differ a lot between platforms and laboratories. Preparation of the library involves fragmentation of DNA and tagging with specific adaptor sequences. These templates are then amplified during sequencing. Two different technologies are used, emulsion PCR and enrichment like in the 454 GS Junior (Roche) or the Ion Torrent PGM (Life Technologies) or solid-phase bridge amplification like in Illumina's MiSeq (Loman, Constantinidou, et al., 2012). One advantage of the MiSeq is that no pre-amplification step is needed which shortens the hands-on time. Although these three benchtop sequencing platforms all rely on the principle of sequencing-by-synthesis, differences lie in the details of sequencing chemistry and sequence reading (Table 1). SGS technologies in general are characterised by a high accuracy and throughput but short read lengths. As a result, single reads can often not be assembled to an entirely closed genome but rather yield a 'draft' genome with unfilled gaps between the reads (Loman, Misra, et al., 2012). Still, this is often sufficient for the purpose of comparative genomics of different highly related strains by mapping the reads to a reference genome (Loman, Constantinidou, et al., 2012; Ronholm, Nasheri, Petronella, & Pagotto, 2016). This so-called reference guided assembly is a valuable tool in phylogenetic and epidemiological investigations.

As an alternative to SGS, in 2011, the first single-molecule, real-time, long-read sequencer, PacBio RS II (Pacific Biosciences) has been put on the market. In this 'Third Generation' sequencer (TGS), the amplification step is omitted and sequencing is based directly on a single DNA molecule. It thereby yields much longer reads than SGS for which fragmented DNA is used (Table 1). High error rates, lower throughput and higher costs per base are disadvantages of this platform (Rhoads & Au, 2015). Nevertheless, this approach is useful in *de novo* assembly of genomes as the long reads help to close gaps between shorter reads. Advantages of second and third generation sequencing can be combined in a complementary approach called 'hybrid sequencing'. Through combination of the high accuracy of SGS and the long reads produced by TGS, a reliable, closed reference genome can be created which can subsequently be used for example for reference guided assembly.

Currently for sequencing of bacterial genomes almost exclusively SGS is used, with a main focus on Illumina sequencers (Schürch & Schaik, 2017). Although general accuracy of SGS systems is high through redundancy of reads, different sequencing technologies exhibit different error characteristics (Junemann et al., 2013; Loman, Misra, et al., 2012). Among benchtop sequencers, the Illumina MiSeq revealed the lowest error rate (rate of < 0.001 indels per 100 bases). It also had the highest throughput per run (1.6 Giga bases of data per run and 60 Mb per hour) and the shortest hands-on time as the amplification step is performed on the sequencer (Junemann et al., 2013; Loman, Misra, et al., 2012). However, selection of the most suitable sequencer heavily depends on the application and specific needs. In clinical context and outbreak investigations, especially high throughput and user convenience are needed at a reasonable price. Besides the technical facts, also subjective preferences play an important role. Exemplary studies comparing different sequencers have shown that results from a single laboratory are neither significantly affected by the sequencing machine nor by the sequencing chemistry (Harris et al., 2013; Kaas, Leekitcharoenphon, Aarestrup, & Lund, 2014). However, detailed and extensive analysis and inter-laboratory evaluation of sequencing practices in use remains to be performed in order to assess minor differences and to establish robustness of results and global comparability.

Global Microbial Identifier (GMI) is an international initiative with the aim of real-time aggregating, sharing, mining and using microbial WGS data (Rindom, 2013). Currently, more than 200 experts from 43 countries are involved. Inclusion of intergovernmental organisations like the World Health Organization (WHO) and the World Organization for Animal Health (OIE) as well as a collaboration with the EU project COMPARE, a multidisciplinary research network establishing a globally linked data information sharing platform system for the control of emerging infectious diseases and foodborne outbreaks (Skiby, 2015), are expected to support international crosstalk and to strengthen the impact of the initiative. One main objective of GMI is the development and realisation of inter-laboratory proficiency testing (PT) to identify steps where quality assurance, control measures or methodological unification are essential to produce standardised high quality sequencing results. In 2014, a pilot PT with only six participants was performed to gain first experience in documentation and practical procedures for this kind of study (Hendriksen et al., 2016). Furthermore, for an optimal adjustment of testing conditions and focus areas prior to a large-scale study, requirements for a general PT among GMI members were interrogated by a survey. Of the 42 respondents, 31% were from the USA, 8.9% and 2.2% from Canada and Australia, respectively, and 51.2% from EU and associated countries (Moran-Gilad et al., 2015). The three most accessible sequencing platforms were MiSeq (23.7%), Ion Torrent PGM (15%) and HiSeq (10.5%), two benchtop and one production-scale sequencer (Table 1). While the benchtop solutions were mostly internally accessible, accessibility to HiSeq was predominantly external. Enquiry of sequencing priorities revealed that foodborne pathogens were the most frequently sequenced pathogens (75%) with high resolution outbreak analysis being the leading application. Among the priority pathogens, *L. monocytogenes* was on the fourth place behind *Escherichia coli, Salmonella* and *Campylobacter* spp. Although the majority of survey respondents agreed that quality filtering and criteria would be important, values specified varied that much that no conclusion could be drawn. For example especially coverage was mentioned as an important quality criterion by 90.9% of

Table 2
Overview of assembly algorithms most frequently used for WGS data.

Assembler	Algorithm	Availability
De novo assembly		
Velvet	De Bruijn Graph	free
Newbler	Overlap/Layout/Consensus	commercial
CLC Genomic	De Bruijn Graph	commercial
SOAPdenovo	De Bruijn Graph	free
Reference-based mapping		
BWA	FM-index	free
Bowtie 2	FM-index	free

respondents, but values ranged from 11-30x (21.6%) over 31-60x (51.3%) to over 60x (18.9%). Apart from the quality criteria, also laboratory methods reported to be used for sample preparation were highly diverse e.g. for DNA or library preparation.

PT based on survey results was performed in 2016 by GMI supported by the U.S. FDA. *Campylobacter coli* and *C. jejuni*, *L. monocytogenes* and *Klebsiella pneumoniae* were selected for analysis. The PT was designed to address three topics: DNA preparation and sequencing procedures, sequencing output, and variant calling of WGS data and cluster analysis. With submission deadline being the 13 January 2017, results remain to be published. Identification of differences and their impact on the sequencing and also analysis results represent an important step towards global harmonisation which would then open the door for international exchange of standardised WGS data.

In general, there are two possible approaches towards global harmonisation of WGS for bacterial typing. The first one is to validate whether different ways are able to yield equivalent results. GMI, for instance, focuses on comparable results among different laboratory practices, sequencing platforms and quality criteria, thereby considering already established local standards. However, this is an organisational and analytical challenge. The second approach is hence the setup of a standard protocol. An exemplary multi-center ring trial was successful in showing that accuracy and reproducibility of NGS based bacterial typing (in this case of *Staphylococcus aureus*) is very high if prescribed methods are applied (Mellmann et al., 2017).

PulseNet USA has developed a Standard Operating Procedure (SOP), PNL32, as a standardised laboratory protocol for WGS of bacterial organisms on the Illumina MiSeq benchtop sequencer (PulseNet, 2015, 2016). PulseNet has been established as a collaboration of CDC, state and local health departments in the USA for real-time comparison of human bacterial pathogens in order to define disease clusters. First initiated for comparison of PFGE profiles, its transition to WGS data is in full progress. With respect to isolates from food and the environment, FDA has launched the GenomeTrakr network. It collaborates with CDC allowing public health authorities to share data from patient and food isolates while investigating foodborne outbreaks and thus ameliorates food safety in the USA. The PulseNet protocol provides standardised and highly detailed methods for DNA isolation and quality control, library preparation and run setup for the sequencer. For example, it stipulates a quality check prior to library preparation where DNA-concentration should at least be 10 ng/μl and meet a 260/280 ratio between 1.75 and 2.05 measured by a Thermo Scientific™ NanoDrop™ spectrophotometer. Furthermore, quality benchmarks for sequence raw data have been specified, more precisely Q-scores and coverage. Q-score has to be Q30 for > 75% of the bases when using a 500 cycle kit and > 85% of bases for a 300 cycle kit and coverage for *L. monocytogenes* needs to be ≥ 20x before upload to PulseNet Central.

Although feasible for laboratories that newly establish WGS, for laboratories that already use WGS, implementation of a standardised protocol could suffer from low compliance as transition would need to be accompanied by investment and change of workflows. Furthermore, as a result of the constant evolution of NGS, continuous adaptation of the standardised procedure is necessary. Newly evolving techniques have to be validated before their inclusion into the SOP which could lead to a delay in their use. For an approach leaving methodical details to individual laboratories and in return defining quality parameters and general thresholds, the use of new techniques would not be a problem as long as final sequence data meet the set criteria. However, such universal quality criteria are very hard to define and are going to require further in-depth analyses and validations until applicable to WGS typing of pathogens in the field of public health and food safety.

4. Dry lab standardisation of NGS methods for foodborne pathogen typing

Through extensive development in NGS technologies, massive amounts of sequence information can be produced within a relatively short period of time. However, bioinformatics tools for the analysis and interpretation of big data struggle to keep pace. In the present time, the bottleneck for integration of NGS-based genome analysis into routine use in disease surveillance is shifting from sequencing to the bio-computational analysis and data storage (Wyres et al., 2014). Currently, no stand-alone tool is able to meet all requirements for a reliable, straightforward and automated analysis of the sequence reads.

For sequencing of bacterial genomes mainly SGS is applied, producing overlapping, short reads. As a result from the high coverage, a high accuracy is achieved. This makes the technology more feasible for variant analyses like SNP detection than the more inaccurate TGS technologies. Still, reads generated by SGS are significantly shorter than those produced by TGS. There are two strategies to deal with this problem: *de novo* assembly of sequence reads to reconstruct a genome or reference-based mapping where single reads are aligned to an already existing, closely related reference genome. For both methods, a variety of different bioinformatics solutions and programs exists (for examples see Table 2).

Algorithms for *de novo* assembly can be grouped into three main categories, all based on graphs: Overlap/Layout/Consensus (OLC)-methods using overlap graphs, de Bruijn Graph (DBG)-methods using k-mer graphs or greedy graph algorithms. Graphs are abstract structures of nodes connected by edges which are used to present relationships. In an overlap graph, the graph represents the sequencing reads (nodes) and their overlaps (edges) of varying length whereas k-mer graphs use subsequences and overlaps of fixed length of k nucleotides (Miller, Koren, & Sutton, 2010). Greedy graphs make use of either the one or the other. Differences lie in the details of graph construction and resulting definition of contiguous sequences of concatenated reads named contigs. Choice of assembler depends on the properties of the sequence data to be used; some assemblers are even specific for a certain sequencing platform. When using DBG assemblers for example, k-mer length has to be adjusted to the read-length while finding a trade-off between sensitivity of smaller and the specificity of longer k-mers (Compeau, Pevzner, & Tesler, 2011). Commercial as well as open-source solutions are available for *de novo* assembly. In a survey among 42 GMI members in 2014, 75% declared using Velvet, freely available software based on DBG and one of the most popular assemblers (Moran-Gilad et al., 2015). Other common assembly software according to the survey was Newbler (46.9%), CLC Genomics (46.9%) and SOAPdenovo (25%). The commercial software Newbler was implemented specifically for 454 GS sequencing platforms and uses OLC whereas the other two programs CLC Genomics (commercial) and SOAPdenovo (freeware), are based on DBG algorithms. It is to note that this survey represents only a relatively small number of software solutions. In practice, a variety of other tools is used as well. Because of the speed of development of software and algorithms, it is difficult to provide a comprehensive and up-to-date ranking. Also changes in the prevalence of use of different sequencing systems influence the popularity of assembly software. For most researchers, commercial software is not a feasible solution. Instead, open-

source software is far more popular. Besides Velvet, SPAdes is a frequently used freeware assembler. In a comparison of seven different assemblers using the quality assessment tool QUAST, it was able to reach the highest amount of mapped genes, the largest N50 value (a measure for the weighted median contig size) and the highest number of complete genes (Gurevich, Saveliev, Vyahhi, & Tesler, 2013). Another example for a freeware assembler is IDBA which in the same study showed its strength in the longest and lowest number of contigs.

Generally, it is difficult to define quality metrics for an assembly as usually the correct answer is not known and assembly errors are very hard to differentiate from biologically relevant SNPs or other genetic changes (Nagarajan & Pop, 2013). Commonly used parameters are hence limited to the assessment of contiguity of the assembly unless an already closed reference genome exists for comparison. Contiguity measures include total size and number of contigs and the N50 value. As a part of the Genome Assembly Gold-standard Evaluations (GAGE) study, eight popular assemblers were compared (Salzberg et al., 2012). One key finding of the study was that the degree of contiguity of an assembly heavily depended on the assembler and the genome to be assembled. Furthermore, quality of the raw data had a considerable impact on the overall quality of the assembly. In general, the correctness of an assembly was found to vary, albeit independently of statistics on contiguity. Consistent with these findings, in another extensive comparison of assemblers, the Assemblathon, significant differences between assemblies from different assembly strategies were revealed (Bradnam et al., 2013; Earl et al., 2011). As a result, it becomes clear that no generic answer can be given to the question of which assembler is the most appropriate for a given dataset. It is rather a case-by-case decision, highly depending on specific requirements.

For detection of variants, in most cases, reference-based mapping as a less computationally intensive method is used. First of all, single sequence reads need to be aligned to a closely related reference genome. There are three categories of alignment algorithms for that purpose: algorithms based on hash tables, based on suffix/prefix tries (i.e. suffix tree, enhanced suffix array or FM-index) or based on merge sorting (Li & Homer, 2010). The latter one is very rarely used, though. According to a survey among GMI members conducted in 2014, most used software solutions were the Burrow Wheeler Aligner (BWA) and Bowtie 2 with 66.7 and 53.3% respectively (Moran-Gilad et al., 2015). Both of them use the FM-index for alignment, the most memory saving and thus most common implementation (Li & Homer, 2010). Other software like e.g. Novoalign, SMALT, MAQ or SHRiMP was only used by 10% of users (Moran-Gilad et al., 2015).

De novo assembly as well as reference-based alignment is only the first step in a series of analytical steps needed for variant detection and in the end clustering of sequences into phylogenetic groups. In general, variants can be called on the basis of SNP detection or gene-by-gene comparison (i.e. cgMLST or wgMLST). Computing efforts in this context are mainly influenced by the decision whether read mapping (less computationally expensive) or a *de novo* assembly (more computationally expensive) is used as starting point for analysis. Although SNP detection provides the highest accuracy, it is often more complex to evaluate. WgMLST and cgMLST serve as a less burdensome alternative providing similar discriminatory power by putting the focus on allelic changes regardless of the number of SNPs involved (Fig. 2). As an extension of the classical MLST which is limited to the analysis of only few housekeeping genes (e.g. 7 for *L. monocytogenes*) (Maiden et al., 1998; Salcedo, Arreaza, Alcala, De La Fuente, & Vazquez, 2003), in cgMLST most genes of the core genome and in wgMLST even the entirety of genes are taken into account. The core genome is defined as the set of genes present in all strains of the same bacterial species whereas the whole genome also comprises accessory genes. One key advantage over SNP detection is that the nomenclature scheme from classical MLST can simply be extended, facilitating consistent classification according to a standardised subtype nomenclature. For L. *monocytogenes*, a core genome scheme has been proposed by Ruppitsch and colleagues in 2015 (Ruppitsch, Pietzka, et al., 2015). Alternatively, a bioinformatics pipeline was designed for cgMLST of *L. monocytogenes*, taking raw sequence reads as input and calculating a core genome profile by comparing it to an expandable database to compile a phylogeny (Pightling, Petronella, & Pagotto, 2015b).

For both, SNP- and allele-based variation detection, different steps and programs have to be combined to receive an informative result. Furthermore, there might be the need for repetitive or consecutive steps within one piece of software. For example a reference-based assembly has to be followed by analysis through a SNP caller to identify mutations. Then, for creation of a phylogenetic tree, another program is usually needed. Likewise, the general analytical process for cgMLST and wgMLST includes assembly, annotation of genes and comparison to a reference. Instead of a user-specific, stepwise analysis, often involving combinations of available programs and custom scripts, the establishment of a generalised and standardised analysis pipeline would be helpful in order to generate a universally valid output.

For isolates of *L. monocytogenes*, different combinations of assembly tools and SNP callers were tested with the result that they varied heavily in the number of true and false positively called SNPs and in accuracy (Pightling, Petronella, & Pagotto, 2015a). Altogether, no general statement on the influence of different parameters could be made as different combinations behaved differently and sometimes even opposed to one another. For example for some combinations, higher coverage led to more true positive identifications of variant sites but in some cases also produced more false positive hits. Besides, read quality trimming and filtering impacted the quality of results either positively or even negatively depending on the software combinations used. This underlines the drastic effect of the variety of methodical implementations on the analysis of NGS data.

Nevertheless, in a proficiency testing for the dry lab part of WGS in 2015 by GMI with more than 40 participants, > 93% of the samples clustered correctly using various analytical approaches (Pettengill et al., 2015). However, number of variants and branch lengths differed considerably indicating that thresholds that led to clustering varied markedly. This shows how difficult it would be to standardise thresholds for different methods. Another study comparing typing capabilities of five different laboratories showed that if methods are prescribed in great detail, a unified output can be reached, underlining the usefulness of a standardised analytical approach (Mellmann et al., 2017). The easiest way for standardisation would probably be the implementation of an analytical pipeline as an aggregation of individually operating segments. This could simplify the transfer of intermediate results between single analytical steps to the level of a single input and output. Current pipelines link individual software pieces for an overarching analysis. Interoperability of components needs to be assured by compatible data formats of program in- and outputs. Furthermore, recordkeeping should be included for transparency and reproducibility of

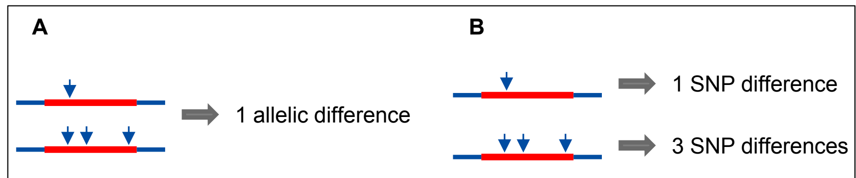


Fig. 2. Detection of differences in cgMLST (A) and SNP analysis (B). In cgMLST analysis, only allelic changes are considered, no matter how many nucleotide changes are found within one allele. In contrast to that, in SNP analysis, every single difference between nucleotide sequences is taken into account.

analysis which is not only crucial for standardisation but ultimately also to secure legal evidence of data (Wyres et al., 2014). In the ideal case, a pipeline should be oriented to the needs of non-specialist, non-bioinformatician users enabling an intuitive and automated analysis.

The American Center for Food Safety and Applied Nutrition (CFSAN), a branch of the FDA, has developed a pipeline for construction of SNP matrices from NGS data (Davis et al., 2015). It combines the following steps: mapping of reads to a reference genome with Bowtie 2, processing of the mapping files with SAMtools, identification of variant sites using VarScan and finally production of a SNP matrix using a custom Python script. The steps are run automatically, converting the input FASTQ data from sequencing reads into a SNP matrix in FASTA format. Still, some points are left to the hands of the user. For example no quality filtering is included. The reason is most likely the number and complexity of possible steps for that purpose and the considerable variations between platforms that make it almost impossible to find a general solution (Edwards & Holt, 2013). Furthermore, the created SNP matrix only serves as an input for the construction of a phylogeny. As a result, despite the robustness and accurateness of the pipeline within the scope of its implementation, it is still no complete solution. Besides the CFSAN pipeline, numerous other, often custom and facility-specific analytical approaches are used and their detailed description would go beyond the scope of this article.

Standardisation for the dry lab analytical part of WGS appears to be even more difficult to achieve than for the wet lab. The easiest way would probably be the establishment of a universal analytical pipeline, in the ideal case directly linked to the sequencer itself and outputting desired results. Although this would hamper flexibility, it would likely be the least complicated and least labour-intensive solution to produce consistent and globally interchangeable data indispensable for molecular surveillance of bacterial infections.

5. Metadata and databases for global sharing of WGS data

On the way to a global sharing of bacterial WGS data, several obstacles remain to be overcome. Besides standardisation of data generation and analysis, international structures and standards for data sharing need to be established.

There are three main databases for the storage of WGS data which together form the International Sequence Database Collaboration (INSDC): the National Center for Biotechnology (NCBI), the European Nucleotide Archive (ENA) and the DNA Data Bank of Japan (DDBJ). GMI as well as the FDA network GenomeTrakr's database use the publicly accessible data layer of parts of the INSDC. GenomeTrakr for example submits data to the NCBI under a single BioProject (Jackson et al., 2016). In Europe, EFSA and ECDC have established a joint database for foodborne pathogens of human and non-human origin. Although not yet used for collection of WGS data but for management of PFGE typing data, it is built to be extended. The ECDC-EFSA database is independent of such public data layers as from the INSDC. It is physically hosted at the TESSy database, the ECDC's database for human strains.

One fundamental issue is the question on the metadata that can be made available together with the WGS data in an open-access database. Metadata is described as "information that is held as a description of stored data" (Dictionary.com), such as isolate specific details like isolation date or source. It is indisputable that such additional information about a bacterial strain provided together with the sequence data greatly increases its utility (Allard et al., 2016). However, national legislation and data protection acts may restrict data sharing and thus limit free data exchange. As a result, there is no consensus about the level of metadata that should be made publicly available. Although different initiatives started collecting bacterial typing data, different concepts of metadata-linkage are proposed. GenomeTrakr includes a minimum set of metadata fields that need to be filled when submitting sequence data to the database. They comprise the collector of the isolate (i.e. the submitting lab), its taxonomic name, sample date and site, the isolation source and sequencing parameters (Allard et al., 2016). GMI expands this list by information on pathogen-associated attributes like specific host or host disease (GMI., 2013). Compared to the GenomeTrakr or the GMI database, for the ECDC-EFSA database, a more restrictive approach is pursued according to EU legislation on data protection. Additional information is restricted to the source of the sample (food, animal, feed, human), typing data and date of sampling whereas for example information related to the origin (country) is considered as potentially sensitive (Rizzi et al., 2017). Identification of the submitting laboratory is also not considered admissible. Furthermore, differentiated access rights for different user groups and stakeholders are incorporated for sensitive data as a compromise between data accessibility and protection. It is to note that the current EU wide Data Protection Basic Regulation including the Data Protection Directive 95/46/EG is about to be changed. However, key principles will remain valid only some aspects have been changed or added. The amendment will become active in May 2018.

Apart from strict legal regulations, there are further reservations regarding free publication of genomic data and metadata (Aarestrup, 2012). As a result of concerns about the ultimate use and possibly the fear of unauthorised application, researchers may not be willing to share their data in public databases before publication. For governments and institutions, competing interests in trade or tourism could be a problem as especially data on foodborne pathogens, the detection of food contamination or even related outbreaks can have far-reaching consequences. Also patenting and intellectual property issues might arise from a free information exchange. Another major challenge is to reconcile protection of confidential patient information and the patients' privacy rights and provision of information needed for epidemiological investigations. These reservations have to be considered when developing a legal framework for public information accessibility. For the EURL *Lm* DB as part of the joint EFSA-ECDC database, compliance with a memorandum of understanding is a prerequisite for participation (Félix, Danan, Makela, et al., 2014). Among others, it regulates data ownership and publication.

An example how well a thought-out and user-oriented database system can work, as among others shown by an almost exponential growth of available genomes, is the Pathosystems Resource Integration Center (PATRIC) (Gillespie et al., 2011; Wattam et al., 2017). PATRIC represents a database coupled with an analysis resource center. Initially designed for the integration of research data and metadata for various pathogens, it now aims to also adapt to the needs of clinical application. Genomic information is linked to metadata including information on organism, isolate, host, sequence, phenotype and project. However, not all fields need to be filled. Although all entries of the database and also information on metadata are publicly available, privacy of data can be maintained by analysing own sequences in a private space without disclosing the information. Still, comparison with public database entries remains possible. On the one hand, this one-way data exchange guarantees protection of possibly sensitive data while allowing reconciliation with already existing data. On the other hand, if all new sequences remain private, progress and timeliness will probably be obstructed. Hence, although technical basis is provided, again a compromise and consensus on the side of the user needs to be found in order to allow for efficient global data sharing.

6. Conclusion

During the last years, WGS has proven its value in the surveillance of *L. monocytogenes* and related outbreak investigations, enabling fast and precise identification of coherent clusters of infection cases, their trace back to food-sources and ultimately elimination of the infection root. So far, several initiatives have been launched to promote WGS based subtyping of *L. monocytogenes* and other foodborne pathogens. However, transition to an international standard remains to be

established.

The One Health strategy makes information exchange between public health, food safety and veterinary authorities indispensable. Hence, a globally accessible sequence database of foodborne pathogens linked with a minimum set of metadata would bring major benefits. The separation of laboratory data and epidemiological or clinical data as well as restricted access to these databases could be an approach to meet data protection criteria. However, more detailed information should be available on demand to effectively protect public health. If an agreement on data format and quality parameters of raw sequence data could be made, direct upload into a centralised analysis pipeline linked to the central database might help to yield standardised and thus comparable sequence information. Still, necessary IT infrastructure has to be established to cope with the problem of data transfer.

The benefits of global data sharing are clear. It helps to provide a comprehensive picture of the appearance and spread of pathogens associated with public health concerns and economic losses around the world. Global data accessibility and exchange is resource-saving as financial burden and workload can be reduced by preventing unnecessary duplication. In addition, it gives the opportunity for a global view and thus improved scientific quality and effective risk management.

Funding

This work was conducted within the project MolTypList which is supported by a grant of the Federal Ministry of Health (GE 2016 03 26), a project in the framework of the German Research Platform for Zoonoses.

Conflicts of interest

None.

References

Aarestrup, F. M. (November 2012). Integrating genome-based informatics to modernize global disease monitoring, information sharing, and response. *Emerging Infectious Disease Journal-CDC, 18*(11).

Allard, M. W., Strain, E., Melka, D., Bunning, K., Musser, S. M., Brown, E. W., et al. (2016). Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *Journal of Clinical Microbiology, 54*(8), 1975–1983. http://dx.doi.org/10.1128/JCM.00081-16.

Aureli, P., Fiorucci, G. C., Caroli, D., Marchiaro, G., Novara, O., Leone, L., et al. (2000). An outbreak of febrile gastroenteritis associated with corn contaminated by Listeria monocytogenes. *New England Journal of Medicine, 342*(17), 1236–1241. http://dx.doi.org/10.1056/NEJM200004273421702.

Bradnam, K. R., Fass, J. N., Alexandrov, A., Baranay, P., Bechner, M., Birol, I., ... Korf, I. F. (2013). Assemblathon 2: Evaluating de novo methods of genome assembly in three vertebrate species. *GigaScience, 2*(1), 10. http://dx.doi.org/10.1186/2047-217x-2-10.

Brosch, R., Buchrieser, C., & Rocourt, J. (1991). Subtyping of Listeria monocytogenes serovar 4b by use of low-frequency-cleavage restriction endonucleases and pulsed-field gel electrophoresis. *Research in Microbiology, 142*(6), 667–675. https://doi.org/10.1016/0923-2508(91)90080-T.

CDC (2015a). *Multistate outbreak of listeriosis linked to blue bell creameries products (final update)*. Retrieved from https://www.cdc.gov/listeria/outbreaks/ice-cream-03-15/.

CDC (2015b). *Multistate outbreak of listeriosis linked to commercially produced, prepackaged caramel apples made from bidart bros. Apples (final update)*. Retrieved from https://www.cdc.gov/listeria/outbreaks/caramel-apples-12-14/.

CDC (2015c). *Multistate outbreak of listeriosis linked to soft cheeses distributed by Karoun Dairies, Inc. (final update)*. Retrieved from https://www.cdc.gov/listeria/outbreaks/soft-cheeses-09-15/.

CDC (2016). *The Listeria whole genome sequencing project.* Retrieved from https://www.cdc.gov/listeria/surveillance/whole-genome-sequencing.html.

Chen, Y., Luo, Y., Curry, P., Timme, R., Melka, D., Doyle, M., ... Strain, E. A. (2017). 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*(2), e0171389. http://dx.doi.org/10.1371/journal.pone.0171389.

Compeau, P. E. C., Pevzner, P. A., & Tesler, G. (2011). How to apply de Bruijn graphs to genome assembly. *Nature Biotechnology, 29*(11), 987–991. http://www.nature.com/nbt/journal/v29/n11/abs/nbt.2023.html#supplementary-information.

Davis, S., Pettengill, J. B., Luo, Y., Payne, J., Shpuntoff, A., Rand, H., et al. (2015). CFSAN SNP Pipeline: An automated method for constructing SNP matrices from next-generation sequence data. *The Perl Journal, 2015*(1), http://dx.doi.org/10.7717/peerj-cs.20.

Deng, X., den Bakker, H. C., & Hendriksen, R. S. (2016). Genomic epidemiology: Whole-genome-sequencing-powered surveillance and outbreak investigation of foodborne bacterial pathogens. *Annual Review of Food Technology, 7*, 353–374. http://dx.doi.org/10.1146/annurev-food-041715-033259.

Dictionary.com. Collins English Dictionary - Complete & Unabridged 10th Edition. Retrieved from http://www.dictionary.com/browse/metadata.

Doumith, M., Buchrieser, C., Glaser, P., Jacquet, C., & Martin, P. (2004). Differentiation of the major Listeria monocytogenes serovars by multiplex PCR. *Journal of Clinical Microbiology, 42*(8), 3819–3822.

Earl, D., Bradnam, K., John, J. S., Darling, A., Lin, D., Fass, J., ... Diekhans, M. (2011). Assemblathon 1: A competitive assessment of de novo short read assembly methods. *Genome Research, 21*(12), 2224–2241.

ECDC (2013). *ECDC roadmap for integration of molecular typing into European level surveillance and epidemic preparedness, Version 1.2.* Retrieved from http://ecdc.europa.eu/en/publications/Publications/molecular-typing-EU-surveillance-epidemic-preparedness-2013.pdf.

Edwards, D. J., & Holt, K. E. (2013). Beginner's guide to comparative bacterial genome analysis using next-generation sequence data. *Microbial Informatics and Experimentation, 3*. http://dx.doi.org/10.1186/2042-5783-3-2 2–2.

EFSA, & ECDC (2016). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2015. *EFSA Journal, 14*(12), e04634. http://dx.doi.org/10.2903/j.efsa.2016.4634 n/a.

Félix, B., Danan, C., Makela, P., Van Walle, I., Lailler, R., Texier, T., ... Roussel, S. (2014a). *Development of a European molecular typing database for food, environmental and veterinary Listeria monocytogenes strains* (euroreference).

Félix, B., Danan, C., Van Walle, I., Lailler, R., Texier, T., Lombard, B., ... Roussel, S. (2014b). Building a molecular Listeria monocytogenes database to centralize and share PFGE typing data from food, environmental and animal strains throughout Europe. *Journal of Microbiological Methods, 104*, 1–8. http://dx.doi.org/10.1016/j.mimet.2014.06.001.

Félix, B., Niskanen, T., Vingadassalon, N., Dao, T. T., Assere, A., Lombard, B., ... Roussel, S. (2013). Pulsed-field gel electrophoresis proficiency testing trials: Toward european harmonization of the typing of food and clinical strains of Listeria monocytogenes. *Foodbourne Pathogens & Disease, 10*(10), 873–881. http://dx.doi.org/10.1089/fpd.2013.1494.

Gilchrist, C. A., Turner, S. D., Riley, M. F., Petri, W. A., & Hewlett, E. L. (2015). Whole-genome sequencing in outbreak analysis. *Clinical Microbiology Reviews, 28*(3), 541–563.

Gillesberg Lassen, S., Ethelberg, S., Bjorkman, J. T., Jensen, T., Sorensen, G., Kvistholm Jensen, A., ... Molbak, K. (2016). Two listeria outbreaks caused by smoked fish consumption-using whole-genome sequencing for outbreak investigations. *Clinical Microbiology and Infections, 22*(7), 620–624. http://dx.doi.org/10.1016/j.cmi.2016.04.017.

Gillespie, J. J., Wattam, A. R., Cammer, S. A., Gabbard, J. L., Shukla, M. P., Dalay, O., ... Mao, C. (2011). PATRIC: The comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. *Infection and Immunity, 79*(11), 4286–4298.

GMI (2013). *Report on the 6th meeting 11-12 september 2013, UC davis*. Sacramento, California. Retrieved from www.globalmicrobialidentifier.org/-/media/Sites/gmi/News-and-events/2013/6th-meeting-2013-report.ashx?la = da.

Graves, L. M., & Swaminathan, B. (2001). PulseNet standardized protocol for subtyping Listeria monocytogenes by macrorestriction and pulsed-field gel electrophoresis. *International Journal of Food Microbiology, 65*(1–2), 55–62. https://doi.org/10.1016/S0168-1605(00)00501-8.

Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: Quality assessment tool for genome assemblies. *Bioinformatics, 29*(8), 1072–1075.

Harris, S. R., Török, M., Cartwright, E. J., Quail, M. A., Peacock, S. J., & Parkhill, J. (2013). Read and assembly metrics inconsequential for clinical utility of whole-genome sequencing in mapping outbreaks. *Nature Biotechnology, 31*(7), 592–594.

Hendriksen, R. S., Pedersen, S. K., Larsen, M. V., Pedersen, J. N., Lukjancenko, O., Kaas, R. S., ... Sintchenko, V. (2016). *The proficiency test (pilot) report of the global microbial identifier (GMI) initiative, year 2014. Report*.

Hunter, P. R., & Gaston, M. A. (1988). Numerical index of the discriminatory ability of typing systems: An application of Simpson's index of diversity. *Journal of Clinical Microbiology, 26*(11), 2465–2466.

Jackson, B. R., Tarr, C., Strain, E., Jackson, K. A., Conrad, A., Carleton, H., ... Gerner-Smidt, P. (2016). Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. *Clinical Infectious Diseases, 63*(3), 380–386. http://dx.doi.org/10.1093/cid/ciw242.

Junemann, S., Sedlazeck, F. J., Prior, K., Albersmeier, A., John, U., Kalinowski, J., ... Harmsen, D. (2013). Updating benchtop sequencing performance comparison. *Nature Biotechnology, 31*(4), 294–296. http://dx.doi.org/10.1038/nbt.2522.

Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M., & Lund, O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One, 9*(8), e104984.

Kleta, S., Hammerl, J., Dieckmann, R., Malorny, B., Borowiak, M., Halbedel, S., ... Al Dahouk, S. (2017). Molecular tracing to find source of protracted invasive listeriosis outbreak, Southern Germany, 2012–2016. *Emerging Infectious Disease Journal, 23*(10), 1680. http://dx.doi.org/10.3201/eid2310.161623.

Kvistholm Jensen, A., Nielsen, E. M., Bjorkman, J. T., Jensen, T., Muller, L., Persson, S., ... Ethelberg, S. (2016). Whole-genome sequencing used to investigate a nationwide outbreak of listeriosis caused by ready-to-eat delicatessen meat, Denmark, 2014. *Clinical Infectious Diseases, 63*(1), 64–70. http://dx.doi.org/10.1093/cid/ciw192.

Lander, E. S., & Waterman, M. S. (1988). Genomic mapping by fingerprinting random clones: A mathematical analysis. *Genomics, 2*(3), 231–239.

Li, H., & Homer, N. (2010). A survey of sequence alignment algorithms for next-

generation sequencing. *Briefings in Bioinformatics, 11*(5), 473–483.
Liu, D. (2006). Identification, subtyping and virulence determination of Listeria monocytogenes, an important foodborne pathogen. *Journal of Medical Microbiology, 55*(6), 645–659.
Loman, N. J., Constantinidou, C., Chan, J. Z., Halachev, M., Sergeant, M., Penn, C. W., ... Pallen, M. J. (2012a). High-throughput bacterial genome sequencing: An embarrassment of choice, a world of opportunity. *Nature Reviews Microbiology, 10*(9), 599–606.
Loman, N. J., Misra, R. V., Dallman, T. J., Constantinidou, C., Gharbia, S. E., Wain, J., et al. (2012b). Performance comparison of benchtop high-throughput sequencing platforms. *Nature Biotechnology, 30*(5), 434–439. http://dx.doi.org/10.1038/nbt.2198.
Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., ... Spratt, B. G. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences, 95*(6), 3140–3145.
Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., ... Chen, Z. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature, 437*(7057), 376–380.
Mellmann, A., Andersen, P. S., Bletz, S., Friedrich, A. W., Kohl, T. A., Lilje, B., ... Harmsen, D. (2017). High interlaboratory reproducibility and accuracy of next-generation sequencing-based bacterial genotyping in a ring-trial. *Journal of Clinical Microbiology, JCM* 02242–02216.
Miller, J. R., Koren, S., & Sutton, G. (2010). Assembly algorithms for next-generation sequencing data. *Genomics, 95*(6), 315–327. http://dx.doi.org/10.1016/j.ygeno.2010.03.001.
Moran-Gilad, J., Sintchenko, V., Pedersen, S. K., Wolfgang, W. J., Pettengill, J., Strain, E., et al. (2015). Proficiency testing for bacterial whole genome sequencing: An end-user survey of current capabilities, requirements and priorities. *BMC Infectious Diseases, 15*(1), http://dx.doi.org/10.1186/s12879-015-0902-3.
Moura, A., Criscuolo, A., Pouseele, H., Maury, M. M., Leclercq, A., Tarr, C., ... Enouf, V. (2016). Whole genome-based population biology and epidemiological surveillance of Listeria monocytogenes. *Nature Microbiology, 2*, 16185.
Nagarajan, N., & Pop, M. (2013). Sequence assembly demystified. *Nature Reviews Genetics, 14*(3), 157–167.
Pettengill, J., Underwood, A., Lukjancenko, O., Strain, E., Karlsmose Pedersen, S., & Hendriksen, R. S. (2015). *2015 GMI PT dry-lab analyses and report.* Retrieved from http://www.globalmicrobialidentifier.org/Workgroups/GMI-Proficiency-Test-Reports.
Pightling, A. W., Petronella, N., & Pagotto, F. (2015a). Choice of reference-guided sequence assembler and SNP caller for analysis of Listeria monocytogenes short-read sequence data greatly influences rates of error. *BMC Research Notes, 8*, 748. http://dx.doi.org/10.1186/s13104-015-1689-4.
Pightling, A. W., Petronella, N., & Pagotto, F. (2015b). The Listeria monocytogenes core-genome sequence typer (LmCGST): A bioinformatic pipeline for molecular characterization with next-generation sequence data. *BMC Microbiology, 15*, 224. http://dx.doi.org/10.1186/s12866-015-0526-1.
PulseNet, C. (2015). *Pulsenet standard operating procedure for illumina miseq data quality control.* Retrieved from https://www.cdc.gov/pulsenet/pdf/pnq07_illumina-miseq-data-qc-508-v1.pdf.
PulseNet, C. (2016). *Laboratory standard operating procedure for pulsenet nextera xt library prep and run setup for the illumina miseq.* Retrieved from https://www.cdc.gov/pulsenet/pdf/pnl32-miseq-nextera-xt.pdf.
Rhoads, A., & Au, K. F. (2015). PacBio sequencing and its applications. *Genomics, Proteomics & Bioinformatics, 13*(5), 278–289. https://doi.org/10.1016/j.gpb.2015.08.002.
Richterich, P. (1998). Estimation of errors in “raw” DNA sequences: A validation study. *Genome Research, 8*(3), 251–259.
Rindom, S. (2013). About GMI - vision and objectives. Retrieved from http://www.globalmicrobialidentifier.org/about-gmi/vision-and-objectives.
Rizzi, V., Felicio, T. D. S., Felix, B., Gossner, C. M., Jacobs, W., Johansson, K., ... Mooijman, K. (2017). *The ECDC-EFSA molecular typing database for European Union public health protection* (Euroreference).
Ronholm, J., Nasheri, N., Petronella, N., & Pagotto, F. (2016). Navigating microbiological food safety in the era of whole-genome sequencing. *Clinical Microbiology Reviews, 29*(4), 837–857. http://dx.doi.org/10.1128/CMR.00056-16.
Ruppitsch, W., Pietzka, A., Prior, K., Bletz, S., Fernandez, H. L., Allerberger, F., ... Mellmann, A. (2015a). Defining and evaluating a core genome multilocus sequence typing scheme for whole-genome sequence-based typing of Listeria monocytogenes. *Journal of Clinical Microbiology, 53*(9), 2869–2876. http://dx.doi.org/10.1128/JCM.01193-15.
Ruppitsch, W., Prager, R., Halbedel, S., Hyden, P., Pietzka, A., Huhulescu, S., ... Wilking, H. (2015b). Ongoing outbreak of invasive listeriosis, Germany, 2012 to 2015. *Euro Surveillance, 20*(50), http://dx.doi.org/10.2807/1560-7917.es.2015.20.50.30094.
Salcedo, C., Arreaza, L., Alcala, B., De La Fuente, L., & Vazquez, J. (2003). Development of a multilocus sequence typing method for analysis of Listeria monocytogenes clones. *Journal of Clinical Microbiology, 41*(2), 757–762.
Salzberg, S. L., Phillippy, A. M., Zimin, A., Puiu, D., Magoc, T., Koren, S., ... Yorke, J. A. (2012). GAGE: A critical evaluation of genome assemblies and assembly algorithms. *Genome Research, 22*(3), 557–567. http://dx.doi.org/10.1101/gr.131383.111.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, 74*(12), 5463–5467.
Schmid, D., Allerberger, F., Huhulescu, S., Pietzka, A., Amar, C., Kleta, S., ... Mellmann, A. (2014). Whole genome sequencing as a tool to investigate a cluster of seven cases of listeriosis in Austria and Germany, 2011-2013. *Clinical Microbiology and Infections, 20*(5), 431–436. http://dx.doi.org/10.1111/1469-0691.12638.
Schürch, A. C., & Schaik, W. (2017). Challenges and opportunities for whole-genome sequencing–based surveillance of antibiotic resistance. *Annals of the New York Academy of Sciences, 1388*(1), 108–120.
Seeliger, H. P. R., & Höhne, K. (1979). Chapter II serotyping of Listeria monocytogenes and related species. In T. Bergan, & J. R. Norris (Vol. Eds.), *Methods in microbiology: Vol. 13*, (pp. 31–49). Academic Press.
Skiby, J. E. (2015). About COMPARE. Retrieved from http://www.compare-europe.eu/about.
Stasiewicz, M. J., Oliver, H. F., Wiedmann, M., & den Bakker, H. C. (2015). Whole-genome sequencing allows for improved identification of persistent Listeria monocytogenes in food-associated environments. *Applied and Environmental Microbiology, 81*(17), 6024–6037. http://dx.doi.org/10.1128/AEM.01049-15.
Swaminathan, B., & Gerner-Smidt, P. (2007). The epidemiology of human listeriosis. *Microbes and Infection, 9*(10), 1236–1243. https://doi.org/10.1016/j.micinf.2007.05.011.
Vázquez-Boland, J. A., Kuhn, M., Berche, P., Chakraborty, T., Domínguez-Bernal, G., Goebel, W., ... Kreft, J. (2001). Listeria pathogenesis and molecular virulence determinants. *Clinical Microbiology Reviews, 14*(3), 584–640.
Wang, S., Weller, D., Falardeau, J., Strawn, L. K., Mardones, F. O., Adell, A. D., et al. (2016). Food safety trends: From globalization of whole genome sequencing to application of new tools to prevent foodborne diseases. *Trends in Food Science & Technology, 57*(Part A), 188–198. https://doi.org/10.1016/j.tifs.2016.09.016.
Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., ... Stevens, R. L. (2017). Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucleic Acids Research, 45*(D1), D535–d542. http://dx.doi.org/10.1093/nar/gkw1017.
Wyres, K. L., Conway, T. C., Garg, S., Queiroz, C., Reumann, M., Holt, K., et al. (2014). WGS analysis and interpretation in clinical and public health microbiology laboratories: What are the requirements and how do existing tools compare? *Pathogens, 3*(2), 437–458.

MICROBIAL GENOMICS

RESEARCH ARTICLE
Lüth *et al*., *Microbial Genomics*
DOI 10.1099/mgen.0.000491

Translatability of WGS typing results can simplify data exchange for surveillance and control of *Listeria monocytogenes*

Stefanie Lüth[1,2,*,†], Carlus Deneke[3]†, Sylvia Kleta[1] and Sascha Al Dahouk[1,4]

Abstract

Where classical epidemiology has proven to be inadequate for surveillance and control of foodborne pathogens, molecular epidemiology, using genomic typing methods, can add value. However, the analysis of whole genome sequencing (WGS) data varies widely and is not yet fully harmonised. We used genomic data on 494 *Listeria monocytogenes* isolates from ready-to-eat food products and food processing environments deposited in the strain collection of the German National Reference Laboratory to compare various procedures for WGS data analysis and to evaluate compatibility of results. Two different core genome multilocus sequence typing (cgMLST) schemes, different reference genomes in single nucleotide polymorphism (SNP) analysis and commercial as well as open-source software were compared. Correlation of allele distances from the different cgMLST approaches was high, ranging from 0.97 to 1, and unified thresholds yielded higher clustering concordance than scheme-specific thresholds. The number of detected SNP differences could be increased up to a factor of 3.9 using a specific reference genome compared with a general one. Additionally, specific reference genomes improved comparability of SNP analysis results obtained using different software tools. The use of a closed or a draft specific reference genome did not make a difference. The harmonisation of WGS data analysis will finally guarantee seamless data exchange, but, in the meantime, knowledge on threshold values that lead to comparable clustering of isolates by different methods may improve communication between laboratories. We therefore established a translation code between commonly applied cgMLST and SNP methods based on optimised clustering concordances. This code can work as a first filter to identify WGS-based typing matches resulting from different methods, which opens up a new perspective for data exchange and thereby accelerates time-critical analyses, such as in outbreak investigations.

DATA SUMMARY

The authors confirm all supporting data, codes and protocols have been provided within the article or through supplementary data files.

Sequencing data have been deposited in the European Nucleotide Archive (ENA) at EMBL–EBI under the accession number PRJEB38495, except for isolate 16-LI00360-0, which is available under the accession number ERS4418852 (SAMEA6659390).

INTRODUCTION

Listeria monocytogenes is the causative agent of the infectious disease listeriosis. While infections may be asymptomatic in otherwise healthy individuals, vulnerable population groups, like immunocompromised or elderly people, pregnant women and newborns, are likely to suffer from severe clinical symptoms, sometimes with a fatal outcome [1]. Although listeriosis is comparatively rare, a hospitalisation rate of 98.6% and a case fatality rate of 13.8% in the European Union (EU) in 2017 clearly show the serious public health hazards [2]. The

Received 10 September 2020; Accepted 21 November 2020; Published 04 December 2020
Author affiliations: [1]National Reference Laboratory for *Listeria monocytogenes*, German Federal Institute for Risk Assessment, Department of Biological Safety, Berlin, Germany; [2]Institute of Biology, Freie Universität Berlin, Department of Biology, Chemistry and Pharmacy, Berlin, Germany; [3]Study Centre for Genome Sequencing and Analysis, German Federal Institute for Risk Assessment, Department of Biological Safety, Berlin, Germany; [4]RWTH Aachen University Hospital, Department of Internal Medicine, Aachen, Germany.
***Correspondence:** Stefanie Lüth, stefanie.lueth@bfr.bund.de
Keywords: core genome MLST; genomic epidemiology; outbreak; single nucleotide polymorphism; standardisation; whole genome sequencing.
Abbreviations: BLASTn, nucleotide BLAST; BLASTp, protein BLAST; CC, clonal complex; CDS, coding sequence; cgMLST, core genome multilocus sequence typing; MLST, multilocus sequence typing; SNP, single nucleotide polymorphism; ST, sequence type; WGS, whole genome sequencing.
The data collected within this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under the accession number PRJEB38495, except for isolate 16-LI00360-0, which is available under the accession number ERS4418852 (SAMEA6659390).
†These authors contributed equally to this work
Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Three supplementary tables and three supplementary files are available with the online version of this article.
000491 © 2020 The Authors

This is an open-access article distributed under the terms of the Creative Commons Attribution License.

vast majority of *L. monocytogenes* infections are foodborne [3]. Consequently, tracing back clinical cases to contaminated food products is one of the key requirements for disease control. However, classical epidemiology alone has proven to be inappropriate for that purpose. The main reasons for this are the very broad range of potentially affected food vehicles and the long incubation period and severity of disease, which complicate patient interviews on food consumption [4–7]. As a result, molecular typing methods have long been applied in *L. monocytogenes* surveillance and outbreak investigations. During recent years, whole genome sequencing (WGS) has revolutionised this field through its unprecedented resolution [8–11].

There are basically two different approaches for WGS-based typing. The first approach is a gene-by-gene comparison where the analysis focuses on allele differences. An example of this approach is core genome multilocus sequence typing (cgMLST), an extension of classical MLST to a larger set of genes that is shared among members of a single species. In the case of *L. monocytogenes*, two main cgMLST schemes are currently in use. One comprises 1701 loci and is built into the software Ridom SeqSphere+ [12], whereas the other one comprises 1748 loci and is built into the software BioNumerics [13]. This incorporation of cgMLST schemes into commercial tools with a graphic user interface has the great advantage of a straightforward operation, including for users lacking bioinformatics skills. However, cost-intensive software licenses might not be affordable for all users. In these cases, it is helpful that both cgMLST schemes are also publicly available and can be used within open-source tools such as the Blast-score-ratio-Based Allele Calling Algorithm (chewBBACA [14]). Although a little bioinformatics training is required, tools like this provide a low-cost alternative. The second approach for WGS-based typing is the single nucleotide polymorphism (SNP) analysis. In this case, single nucleotide variations are used as a distance measure between bacterial sequences. Both, commercial (e.g. BioNumerics) and open-source (e.g. Snippy [15]) solutions are available for analysis. SNP analysis is commonly based on a comparison against a selected reference genome. However, the genome chosen as reference can affect analysis results [16, 17].

For effective surveillance and control of human listeriosis, not only comprehensive molecular typing of *L. monocytogenes* isolates from food, food processing environments and clinical cases, but also communication of results between different sectors (food safety, public health) and countries is needed. However, procedures for WGS-based typing are diverse and not yet fully harmonised. Starting from the sequencing protocol through quality filtering (e.g. read trimming) to algorithms for assembly, mapping or variant calling and finally distance assessment, there is considerable space for variation. The ideal way to go for the future will be the harmonisation of all these methods between different laboratories to enable the seamless exchange of analysis results. To date, several international initiatives have been commenced to deal with this issue [18–20]. However, until a generally accepted solution has been found, an interim solution is urgently needed.

Impact Statement

For effective surveillance and control of human listeriosis, not only comprehensive molecular typing of *Listeria monocytogenes* isolates from food, food processing environments and clinical cases, but also communication of results between different sectors (food safety, public health) and countries is needed. The currently available procedures for WGS-based typing are diverse and not yet fully harmonised. The ideal way to go for the future will be the harmonisation of methods between different laboratories to enable seamless data exchange. However, until a generally accepted solution has been found, an interim solution has to be established. We therefore compared the results of the most commonly used genotyping tools for *L. monocytogenes*, and developed a translation code for the identification of typing matches resulting from different methods. This approach opens up a new perspective for the exchange of WGS analysis results.

In order to assess the transferability of results, we compared the most commonly used WGS-based typing methods for *L. monocytogenes*. Our aim was to provide a translation code as a first filter for the identification of typing matches resulting from the different methods. In addition, we describe a procedure that can also be applied to the comparison of other methods.

METHODS

Study dataset

A total of 494 isolates from ready-to-eat food products and food processing environments sampled in official controls in 2016 were selected from the strain collection of the German National Reference Laboratory for *L. monocytogenes* as a representative dataset for the population structure of *L. monocytogenes* in the food chain in Germany.

Bacterial strain cultivation

Pure cultures of *L. monocytogenes* isolates were routinely stored at −80 °C in brain heart infusion medium with 20v/v % glycerol. Prior to downstream analysis, bacteria were plated onto Sheep Blood Agar and incubated overnight at 37 °C.

Genomic DNA extraction and whole genome sequencing (WGS)

Overnight cultures of *L. monocytogenes* strains were harvested and lysed following the PulseNet standardised laboratory protocol for WGS of Gram-positive bacteria (https://www.cdc.gov/pulsenet/pdf/pnl32-miseq-nextera-xt.pdf). For DNA extraction, the QIAamp DNA Mini Kit (Qiagen) was used following the manufacturer's instructions. Purity of extracted DNA ($OD_{260:280}$ and $OD_{260:230}$ ratio) was measured with the NanoDrop spectrophotometer

(Thermo Fisher Scientific) and extracted DNA was quantified using the Qubit dsDNA BR Assay Kit with a Qubit 2.0 fluorometer (Invitrogen). Sequencing libraries were constructed with the Nextera XT Sample Preparation Kit (Illumina) for sequencing in paired-end mode with 2×300 bp reads on an Illumina MiSeq sequencer using the MiSeq Reagent v3 600-cycle Kit (Illumina).

Sequencing quality control and genome assembly

Raw sequencing reads were quality checked using FastQC version (v) 0.11.5 [21] and trimmed using Trimmomatic v 0.36 [22]. Subsequently, trimmed reads were assembled and analysed using the pipeline Assembly-based QUality Assessment for Microbial Isolate Sequencing (AQUAMIS) v 0.9.0 at default parameters [23]. This pipeline includes the tools unicycler v 0.4.4 for assembly and assembly polishing, mash v 2.1 for reference search, and quast v 4.6.3 for assembly quality control. Genome assemblies obtained from the AQUAMIS pipeline served as a starting point for cgMLST-based typing, whereas trimmed reads were used as the basis for SNP analysis. Detailed information on tools and parameters can be found in Supplementary File S1 (available in the online version of this article).

Classical multilocus sequence typing (MLST)

Classical seven-gene MLST sequence types (STs) and corresponding MLST clonal complexes (CCs) were determined from assembled draft genomes according to the scheme available at https://bigsdb.pasteur.fr/listeria/listeria.html using Ridom SeqSphere+ (Ridom).

Closed genomes of *L. monocytogenes* available at NCBI were analysed with the software mlst [24] and a reference genome was chosen for each MLST CC (https://github.com/crarlus/refseq-MLST/).

Core genome (cg) MLST analysis

Ridom SeqSphere+

Assembled draft genomes were analysed in Ridom SeqSphere+with the 'Process assembled genome data' function at default parameters for *L. monocytogenes*. The integrated 1701 loci scheme was used [12]. A cgMLST allele coverage of at least 98% was set as quality threshold, assuming that this value is representative of the entire genome quality [25]. If the threshold was not reached, sequencing was repeated. Resulting allele profiles were exported in tsv format. This method will be referred to as Ridom_Ruppitsch.

BioNumerics

The WGS tools plugin of BioNumerics v 7.6.3 (Applied Maths) was used for analysis with the integrated 1748 loci cgMLST scheme [13]. Resulting cgMLST allele profiles were exported in tsv format. This method will be referred to as BioNumerics_Moura.

chewBBACA

chewBBACA is a freely available software suite that allows scheme creation, allele calling and scheme evaluation [14]. Allele calling starts with the identification of coding sequences (CDS) using prodigal [26]. If an exact match to the allele database is found, the corresponding allele number is assigned. Otherwise, a BLASTP score ratio (BSR) approach evaluates whether a novel allele is present, or no allele can be inferred. Newly inferred alleles are updated in a local allele database.

Here, we used the pipeline chewieSnake [27] that calls alleles for a set of samples using chewBBACA v 2.0.12, combines their allele profiles and infers an allele distance matrix as well as a minimum spanning tree using GrapeTree v 1.4.1 [28]. Subsequently, samples can be hierarchically clustered and a cgMLST report is compiled. As we used chewieSnake with the cgMLST scheme developed by Ruppitsch and colleagues [12], this method will be referred to as chewBBACA_Ruppitsch.

SNP analysis

Reference genomes

The application of three different kinds of reference genomes was compared: 1) general, species-specific, closed; 2) subgroup-specific, closed; 3) subgroup-specific, draft. The genome of the *L. monocytogenes* reference strain EGDe (NC_003210.1) was used as general reference. Isolates were assigned to subgroups according to MLST CCs. As far as possible, MLST CC-specific closed reference genomes were selected as described by the European Food Safety Authority [16]. To identify reference genomes for those MLST CCs for which the European Food Safety Authority had not specified one, closed genomes of *L. monocytogenes* available at NCBI were subjected to MLST CC determination. In the case of more than one closed genome per MLST CC, the reference genome was selected randomly. For MLST CC-specific draft reference genomes, draft genomes with the best assembly quality (highest N50, lowest number of contigs) per MLST CC were selected from our dataset. Only MLST CCs containing more than five isolates and with a closed reference genome available in NCBI were included in MLST CC-specific analyses.

BioNumerics

For SNP analysis in BioNumerics, the basic version of BioNumerics v 7.6.3 without the WGS tools plugin was used. Strict filtering of SNPs at software default settings was applied. This method will be referred to as SNP_BioNumerics.

Snippy

Snippy was chosen as a representative open-source SNP pipeline since it is recognised as one of the most reliable SNP pipelines [29].

SNPs were detected with the variant calling pipeline snippy-snake [30]. In short, SNPs were called with snippy v 4.0 [15], the core alignment was determined using snippy-core and the SNP distance matrix using snp-dists [31]. Subsequently, the pipeline clustered all samples into cluster types for a range of

thresholds using hierarchical clustering and generated a SNP report. This method will be referred to as SNP_Snippy.

Filtering of isolates for MLST CC-specific analyses

While generally applicable typing methods may give an adequate overview, it can be useful to repeat certain analyses only for a subgroup of isolates to gain deeper insights. We used MLST CC-specific analysis for this purpose. However, isolates belonging to different MLST CCs may vary in diversity, for example depending on the number of individual STs within the CC. This can result in exceptionally large SNP distances, which will distort the results. Actually, the establishment of the largest core genome is required for detailed SNP analysis. Therefore, after initial SNP analysis, very distantly related isolates within each MLST CC were identified (>800 SNPs in SNP_BioNumerics and >18000 SNPs in SNP_Snippy) and excluded from further MLST CC-specific analyses.

Properties and correlations of distance matrices

For all selected cgMLST methods, distance matrices were calculated with GrapeTree v 1.4.1 [28] (using the option '--missing 0' to deal with missing loci) based on allele profiles. Distance matrices for SNP analyses were used as yielded from primary analysis.

All downstream analyses from distance matrices were performed in R using the packages *plyr*, *reshape2* and *ggplot2*. Distance matrices were linearized and sorted to compare pairwise distances. For MLST CC-specific analyses, the resulting sub-settings of distance matrices per MLST CC were merged to one. Boxplots were generated to visualise the magnitude of detected distances and Spearman correlation was used to quantify the similarity of pairwise distances between different methods. Correlations were visualised using the package *corrplot*.

Concordance of clustering

Distance matrices were used for single linkage clustering at different threshold values in R. The clustering results dependent on methods and thresholds served as inputs for the Comparing Partitions online tool available at http://www.comparingpartitions.info [32]. The adjusted Wallace coefficient was selected as pairwise agreement measure because it directly indicates the concordance between clusters. The coefficient can be regarded as the probability that a cluster calculated by method 1 matched that calculated by method 2, and vice versa. Always two adjusted Wallace coefficients deriving from two comparison directions were determined.

Establishment of a translation code between methods

We assessed the degree of concordance between three different cgMLST methods (comprising three software solutions and two cgMLST schemes) and six different SNP methods (two software solutions with three types of reference genomes each). Our aim was to define threshold values that can be communicated for a comparable interpretation of clustering results. As they are epidemiologically well defined, we chose two published allele distance thresholds for the cgMLST schemes (seven and ten allele differences, referred to in previous work [12, 13]) as references to establish our translation code. More precisely, the clustering information derived from one of the three cgMLST methods with one of the two threshold values was set as a reference and was compared with the clustering at various threshold values in a comparison method (other cgMLST method or SNP method) using the Comparing Partitions online tool as described above. The threshold value of the comparison method, at which the sum of the two adjusted Wallace coefficients reached a maximum, was defined as the 'adjusted threshold'.

Practical test of the translation code

The translation code was tested with the cgMLST dataset retrieved from BioNumerics_Moura at an allele threshold of seven as reference method. The five largest clusters, for which also MLST CC-specific analyses had been performed, were selected. Thus, one cgMLST cluster each from MLST CC9, CC121, CC3, CC8 and CC2 was included in our analysis. One isolate per cgMLST cluster was randomly selected and used for cluster search at adjusted threshold values in the other methods. This approach reflects the generally valid workflows during international disease outbreaks, when the sequence of an individual reference isolate is shared between laboratories as the basis for local cluster identification.

RESULTS

Quality control

All 494 isolates were sequenced with coverage between 32 and 231 (median 99). Raw reads could be assembled into 15 to 72 contigs (median 29) with an N50 between 9.6×10^4 and 1.5×10^6 (median 3.6×10^5). Median cgMLST allele coverage using Ridom SeqSphere+was 99.8%.

Comparison based on distance matrices

Generally applicable methods

In cgMLST analysis, pairwise allele distance between isolates ranged from 0 to 1687 (median 1347) using Ridom_Ruppitsch, from 0 to 1687 (median 1351) using chewBBACA_Ruppitsch and from 0 to 1740 (median 1409) using BioNumerics_Moura. The differences in pairwise distances resulting from Ridom_Ruppitsch and BioNumerics_Moura varied between −89 and 24 (median −55), from chewBBACA_Ruppitsch and BioNumerics_Moura between −87 and 27 (median −54), and from Ridom_Ruppitsch and chewBBACA_Ruppitsch between −12 and 4 (median −1). Method correlations were 0.97 and 0.98 using different cgMLST schemes and 1 with the same scheme (Fig. 1). A visual comparison between distances derived from different methods can be found in Supplementary file 2.

Pairwise SNP distance between isolates with EGDe as the reference genome ranged from 0 to 12694 (median 3504) using SNP_BioNumerics and from 0 to 107646 (median 26884) using SNP_Snippy. Method correlation was 0.89.

Correlations of cgMLST_Ridom_Ruppitsch and cgMLST_chewBACCA_Ruppitsch were 0.89 and 0.88 to

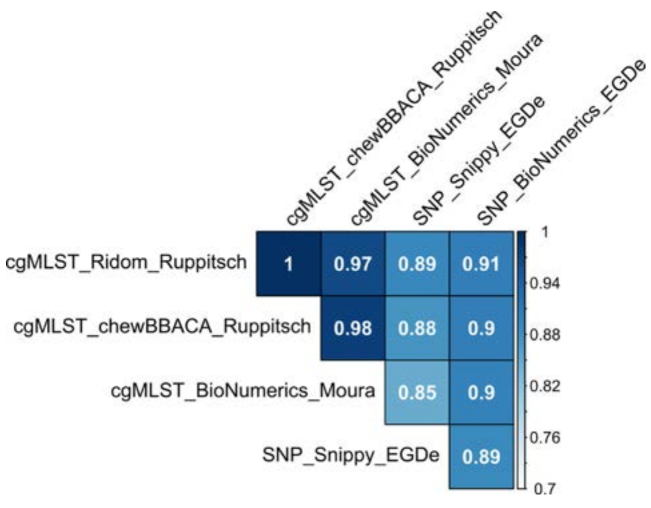


Fig. 1. Correlations of generally applicable typing methods, based on linearized distance matrices. Colour scale indicates the strength of correlation.

SNP_Snippy_EGDe and 0.91 and 0.9 to SNP_BioNumerics_EGDe, respectively (Fig. 1). Correlation of cgMLST_BioNumerics_Moura was 0.85 to SNP_Snippy_EGDe and 0.9 to SNP_BioNumerics_EGDe.

Subgroup (MLST CC)-specific methods

The 494 isolates belonged to 39 different MLST CCs (Tables S1 and S2), out of which 19 MLST CCs contained at least five isolates, but a closed reference genome was only available for 16 of them at NCBI. Accordingly, 409 isolates were selected for initial MLST CC-specific analyses. After filtering out those isolates with too large SNP distances within an MLST CC, 394 isolates from 15 different MLST CCs were left (Table 1). Filtered isolates came from CC8 (n=3), CC4 (n=1) and CC14 (n=9). As for CC14, only two isolates were left after filtering, the entire MLST CC was excluded from further analyses.

In SNP_BioNumerics, use of an MLST CC-specific closed reference genome led to pairwise SNP distances between 0 and 292 (median 68), whereas use of a specific draft reference genome yielded 0 to 290 (median 70) pairwise SNP distances (Fig. 2). Applying SNP_Snippy, SNP distances with a specific closed reference genome ranged between 0 and 622 (median 68) and between 0 and 714 (median 69) with a specific draft reference genome. In the MLST CC-specific analyses with EGDe as a reference, SNP distances were 0 to 64 (median 17) using SNP_BioNumerics and 0 to 240 (median 59) using SNP_Snippy.

The median ratio of pairwise distances resulting from SNP analysis with the same software but with a closed or draft specific reference genome was 1. The median ratio between a closed or a draft specific reference and EGDe was 3.8 and 3.9, respectively, using SNP_BioNumerics, and 1.2 using SNP_Snippy. When EGDe was used as reference genome, the median ratio of pairwise distances between SNP_Snippy and SNP_BioNumerics was 3.1.

Overall, there was near perfect (0.99) to perfect (1) correlation between different software and closed or draft specific reference genome usage in SNP analysis. Lowest method correlations were found with SNP_BioNumerics when using

Table 1. MLST CCs and references used for MLST CC-specific analyses (sorted by frequency in our dataset)

MLST CC	Closed reference (GenBank Accession)	Draft reference	Coverage	Contigs
CC121	HG813249	16-LI01132-0	91	21
CC9	FR733649	16-LI00873-0	77	17
CC8	CP006862	16-LI00415-0	84	19
CC2	CP006046	16-LI01038-0	119	25
CC3	CP006594	16-LI00227-0	148	27
CC1	AE017262	16-LI00258-0	61	19
CC37	CP011397	16-LI00295-0	113	20
CC6	CP006047	16-LI00782-0	85	16
CC5	CP006592	16-LI00750-0	133	21
CC101	CP025221	16-LI00284-0	117	20
CC18	CP020830	16-LI00319-0	119	15
CC155	CP002004	16-LI00862-0	90	25
CC224	CP016629	16-LI00391-0	91	24
CC7	CP002002	17-LI00007-0	112	21
CC4	FM242711	16-LI00480-0	93	27

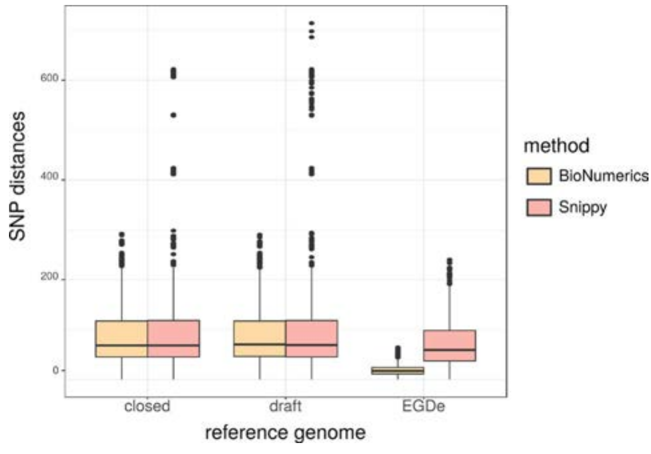


Fig. 2. Boxplot of SNP distances from BioNumerics and Snippy using different reference genomes for SNP analysis (applied to a subset of 394 isolates of 15 different MLST CCs), based on linearized distance matrices.

EGDe as reference (0.77 to 0.85). All other correlations were larger than 0.96 (Fig. 3).

Distances to the reference and size of the core genome

When having a more detailed look into the results from SNP_Snippy (Table S3), isolates had a smaller SNP distance to the draft than to the closed MLST CC-specific reference genomes. On the one hand, more SNP positions were missing when the isolate reads were mapped to the closed references. On the other hand, however, the size of the closed reference genomes tended to be larger than that of the draft ones (by 26 kbp on average). Altogether, the core genome size (defined as the number of positions in the reference that are neither missing nor masked in any of the isolate's mapping to the reference) was in the end very similar between draft (median size 2809303 bp) and closed reference (median size 2802508 bp) genomes. In SNP_Snippy, the isolates had a distance of 25000 SNPs to the EGDe reference on average. Furthermore, we observed a substantial increase in the number of missing as well as masked positions compared with the specific reference

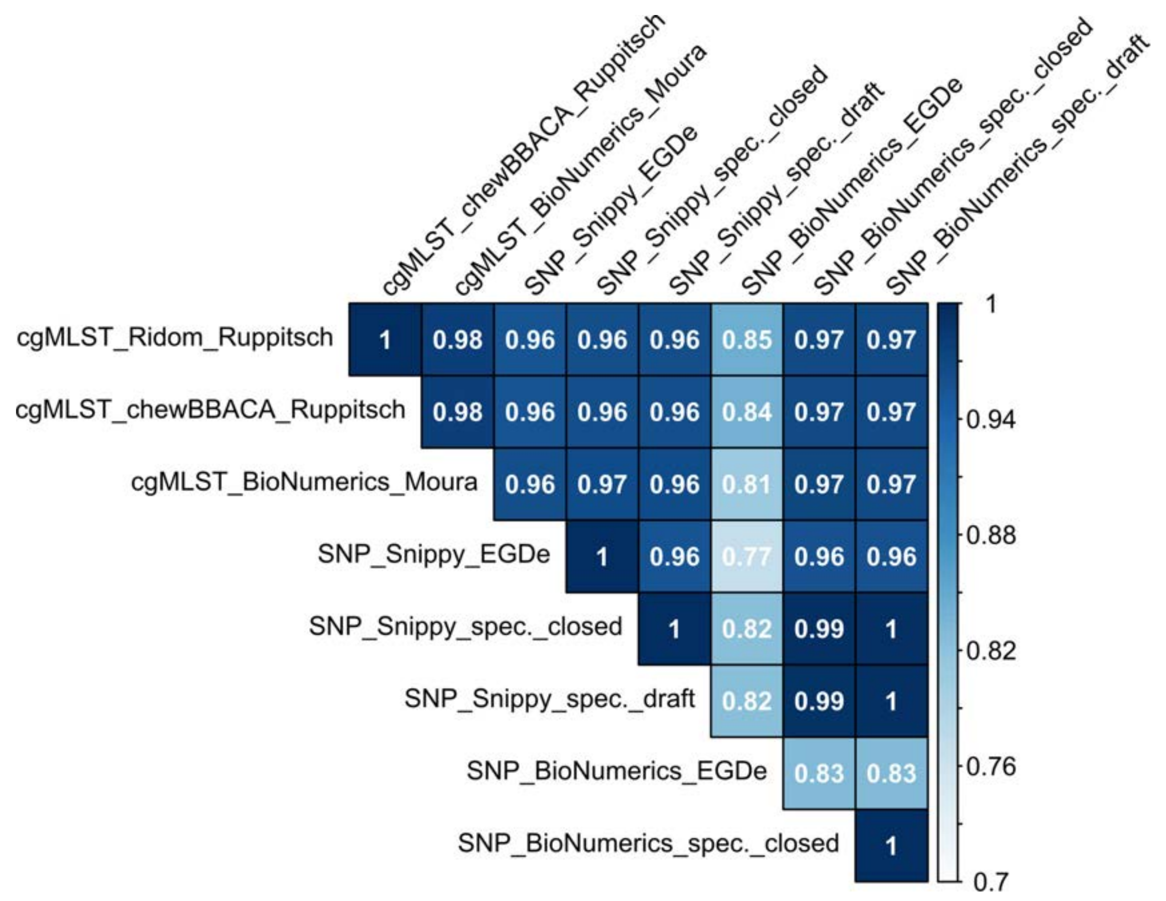


Fig. 3. Correlations of MLST CC-specific typing methods (applied to a subset of 394 isolates of 15 different MLST CCs), based on linearized distance matrices. Colour scale indicates the strength of correlation.

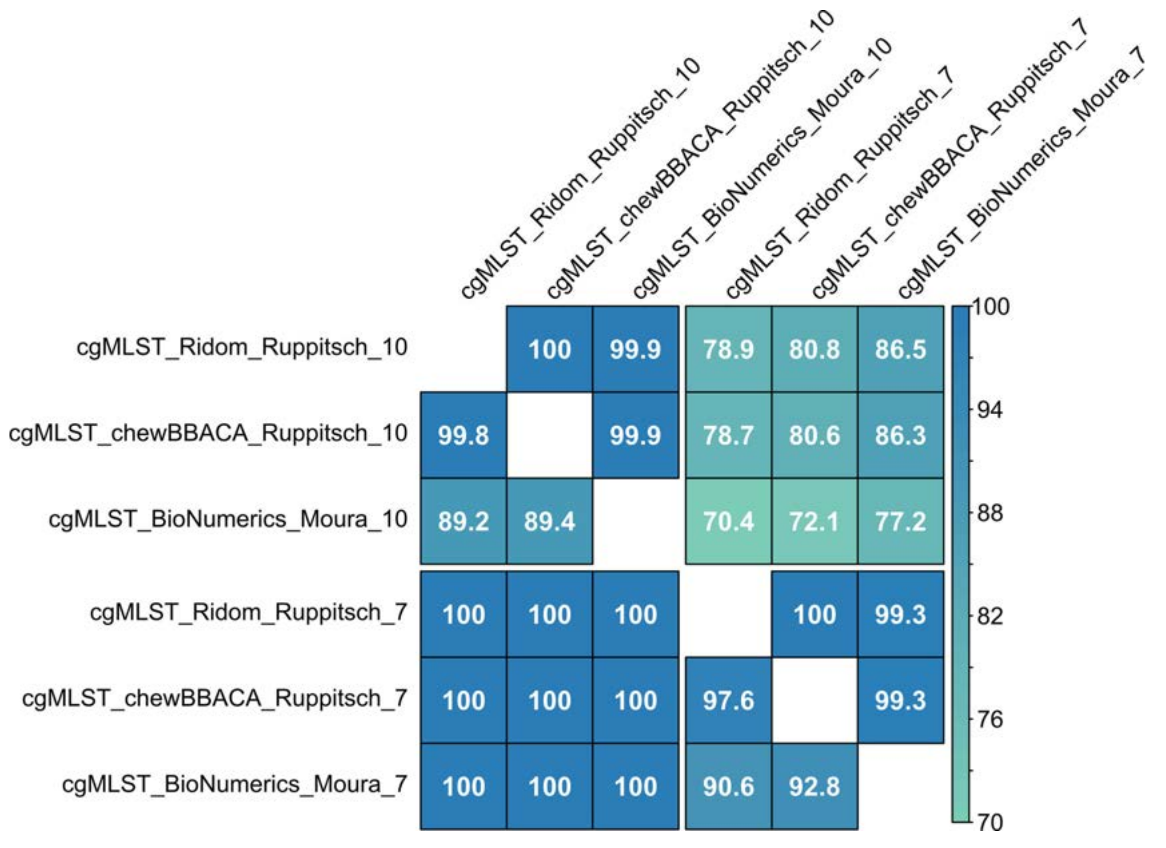


Fig. 4. Matrix of adjusted Wallace coefficients (direction-dependent values) for cgMLST methods at common thresholds (seven and ten alleles). Colour scale indicates percentage of concordance.

genomes. Therefore, the core genome size when using EGDe as reference was only 2281008 bp.

Comparison of clustering

cgMLST methods

To compare the clustering of isolates, threshold values published for the two cgMLST schemes were applied to the different cgMLST approaches. For the Ruppitsch scheme this is ten alleles [12], and for the Moura scheme seven alleles [13] between neighbouring isolates. Agreement was perfect when comparing clusters at a seven-allele threshold with clusters at a ten-allele threshold (adjusted Wallace coefficient 100%). The other way around (from ten to seven), however, concordance was only between 70.4 and 86.5% (Fig. 4). When using the same threshold values for different methods, overall concordance was higher than with different thresholds. At a threshold of seven alleles, concordance was 97.6 and 100% when using the same cgMLST scheme in different software and between 90.6 and 92.8 and 99.3% (depending on the direction) for different schemes. At a threshold of ten alleles, concordance was 99.8 and 100% when using the same cgMLST scheme in different software and between 89.2 and 89.4 and 99.9% (depending on the direction) for different schemes.

Optimisation of clustering and translation code between cgMLST and SNP methods

The cgMLST clustering at described threshold values [12, 13] was set as the reference for the adjustment of clustering thresholds for other methods. Our idea was to define threshold values, which allow for the communication of clustering information between laboratories. Table 2 displays the resulting translation code, which can be applied as follows: in a case in which Laboratory A uses cgMLST analysis with BioNumerics_Moura at the published allele threshold of seven, an allele threshold of eight in cgMLST analysis with Ridom_Ruppitsch used in Laboratory B would result in the best cluster agreement. The corresponding adjusted Wallace coefficients, 97.3 and 98.2%, can be found in Fig. 5a. If Laboratory C uses SNP analysis with Snippy_EGDe, a threshold of 13 SNPs should be applied to yield comparable clustering to Laboratory A. If Laboratory C, however, uses a different reference genome in SNP analysis, for example a draft MLST CC-specific one (Snippy_draft), a threshold of 12 SNPs will be best suited to fit the clustering from Laboratory A.

Although use of identical thresholds in different cgMLST methods already led to higher concordance of clustering than use of different cgMLST scheme-specific thresholds, slight adjustment of thresholds (±1 allele) could increase

Table 2. Adjusted thresholds for optimised clustering concordance between cgMLST methods and between cgMLST and SNP methods. Clustering by cgMLST methods at published thresholds [12, 13] (in bold type on the left) was set as reference for the adjustment of clustering thresholds for other methods. The columns show the different comparison methods and the threshold values (alleles or SNPs) at which the greatest possible agreement among the clustering with the respective reference method was achieved based on adjusted Wallace coefficients presented in Fig. 5. As cluster comparison is direction-dependent, the table must be read from the left to the right

			cgMLST			SNP					
						General reference		MLST CC-specific reference			
		Allele threshold	Ridom_Ruppitsch	chewBBACA_Ruppitsch	BioNumerics_Moura	Snippy_EGDe	BioNumerics_EGDe	Snippy_closed	Snippy_draft	BioNumerics_closed	BioNumerics_draft
cgMLST	Ridom_ Ruppitsch	**7**		6	6	9–10	4	11	12	12	11–12
	chewBBACA_ Ruppitsch		7		7	9–10	4	11	12	12	11–12
	BioNumerics_ Moura		8	7		13	4	12	12	11	11–12
	Ridom_ Ruppitsch	**10**		10	10	15	5	19	19	18	18
	chewBBACA_ Ruppitsch		10		10	15	5	20	18–19	18	18
	BioNumerics_ Moura		11	11		18	6	20	22	19	20

concordance even more (Table 2, Fig. 5). For example, clustering at an allele threshold of seven in BioNumerics_Moura, compared with Ridom_Ruppitsch setting a threshold of eight instead of seven alleles, led to a method concordance of at least 97.3% (Fig. 5a) instead of only 90.6% (Fig. 4).

Overall, at an allele threshold of seven, achievable method concordance with cgMLST and SNP methods was at least 90.9% (Fig. 5a) and at a threshold of ten alleles in cgMLST at least 77.9% (Fig. 5b).

When using a general reference genome (EGDe) in SNP analysis, threshold values for optimised clustering concordance with cgMLST were lower than with an MLST CC-specific reference. Additionally, thresholds differed between SNP_Snippy_EGDe and SNP_BioNumerics_EGDe, but threshold values were similar for MLST CC-specific approaches irrespective of whether closed or draft references or the two different software tools were applied (Table 2).

Practical test of the translation code

For the cgMLST cluster from CC121 retrieved from BioNumerics_Moura at an allele threshold of seven (16 isolates), clustering differed by one to six isolates (median 3.5) when using other methods (Fig. 6).

In contrast, for the cgMLST clusters from CC3 (15 isolates) and CC2 (seven isolates), agreement was perfect, apart from a single isolate that was missing in the clustering results from SNP_Snippy_EGDe. For the clusters from CC9 (26 isolates) and CC8 (eight isolates), exactly the same isolates were found to form a cluster at the adjusted threshold values in all methods.

DISCUSSION

cgMLST

Use of the Moura cgMLST scheme mostly resulted in higher allele distances than the Ruppitsch scheme. Given that the number of loci included in the two schemes differs by 47, this was to be expected. Overall correlation of the different cgMLST approaches was high, probably due to the 1261 loci overlap between the two schemes [13]. However, use of the Ruppitsch scheme either in Ridom SeqSphere+or in chewBBACA resulted in slightly different allele distances. This can be attributed to differences in the way the two algorithms work. While the query sequence is compared to the loci via a nucleotide BLAST (BLASTn) in Ridom SeqSphere+, chewBBACA is centred on the prediction of CDS and a subsequent protein BLAST (BLASTp). The idea behind making a BLASTp instead of a BLASTn is that silent mutations are ignored because they are biologically irrelevant [14]. However, a major problem with BLASTp arises from frameshifts – either biologically present or due to assembly errors – which can change an entire protein (all amino acids). In a BLASTn approach, a frameshift is perceived as a single nucleotide change.

Apart from the pure distances between isolates, in outbreak investigations especially, clustering of isolates is important to

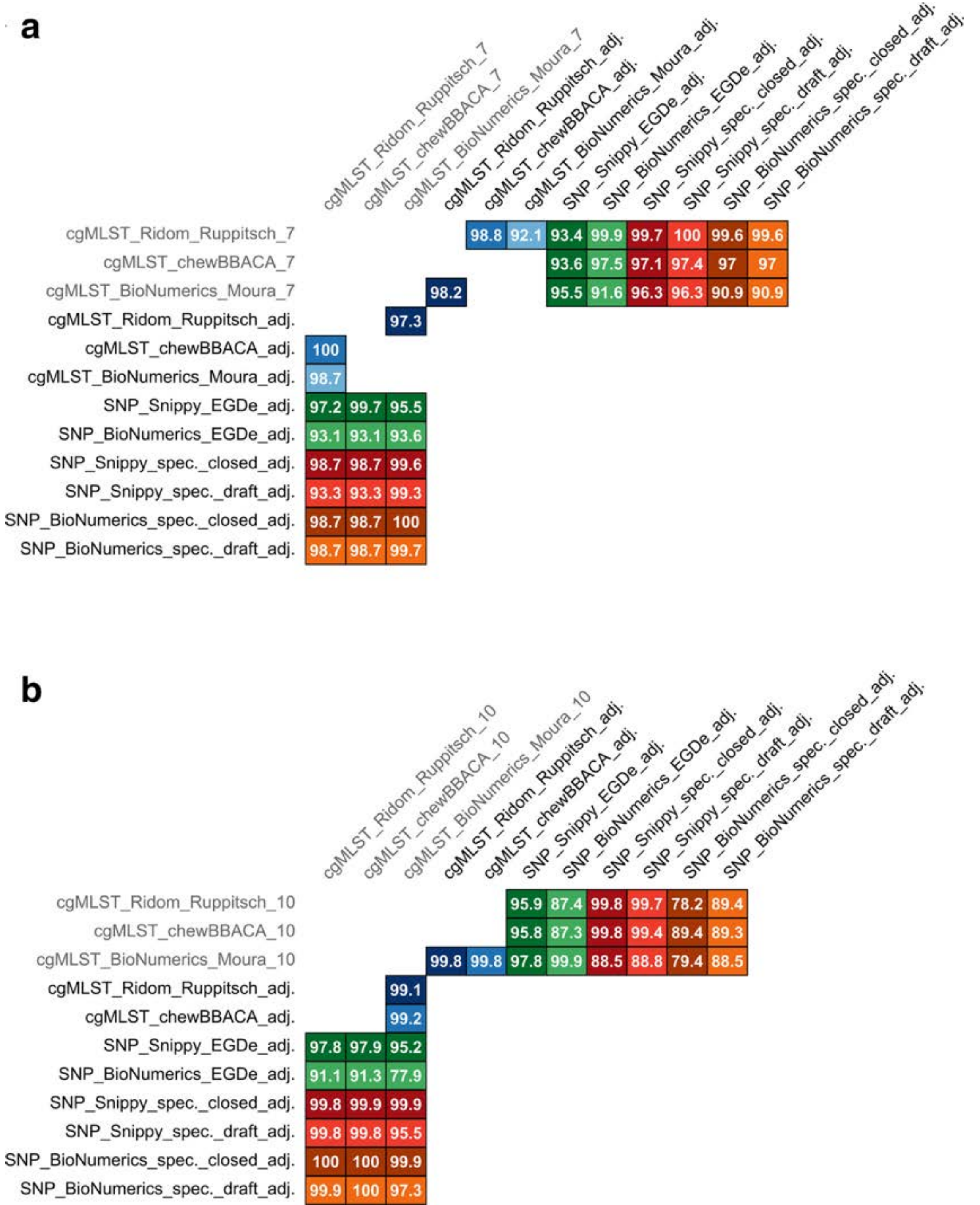


Fig. 5. Adjusted Wallace coefficients (direction-dependent values) at optimised clustering thresholds. (a) Threshold seven alleles, (b) Threshold ten alleles. Grey text colour indicates that the method was used as the reference for threshold adjustment. Percentage values of concordance are presented. Each method has a specific colour and rows and columns of the same colour represent the two directions of cluster comparison. adj.: adjusted threshold from Table 2.

provide enough evidence for potential epidemiological links. Due to its ease of use and the possibility of a unified nomenclature, gene-by-gene approaches are recommended for that purpose by the PulseNet International global consortium [20, 25]. We, therefore, applied cgMLST together with epidemiologically well-defined clustering threshold values to establish a translation code between different WGS typing approaches (Table 2). It is important to note that the idea

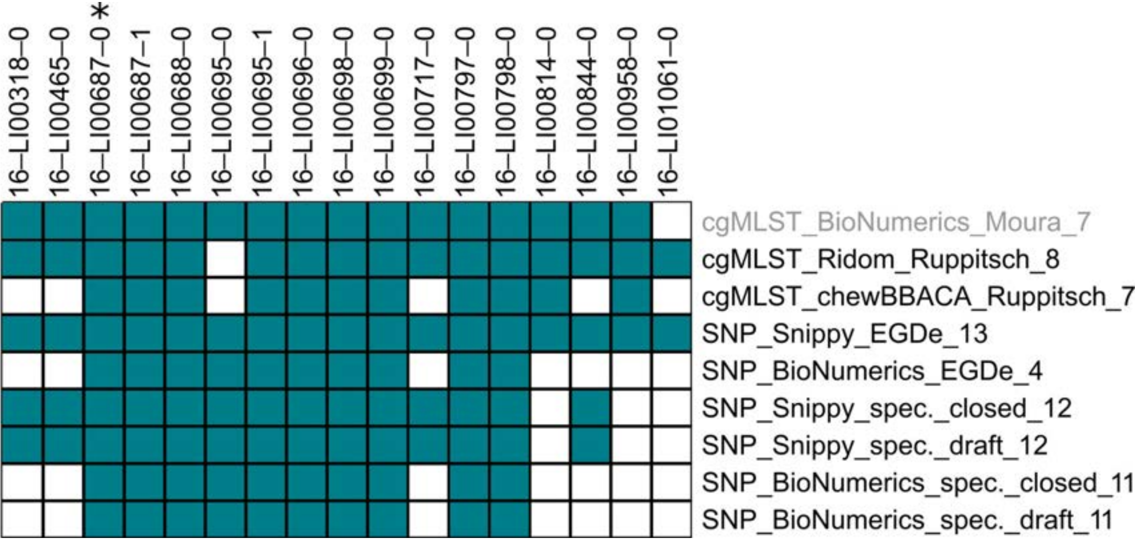


Fig. 6. Practical test of the translation code taking a cgMLST cluster of 16 isolates belonging to MLST CC121 as an example. The cgMLST dataset retrieved from BioNumerics_Moura at an allele threshold of seven (grey text colour) was used as reference method for clustering. Labelling on the right, 'method_threshold'. Upper labels: isolate identifiers. An asterisk indicates the isolate that was used for cluster search in the different methods. Members of a cluster are coloured. Corresponding distance matrices can be found in Supplementary file S3.

of fixed clustering thresholds is controversial, as has been discussed previously [33, 34]. WGS trace-back analysis always has to be used in combination with epidemiological evidence and published thresholds should be seen more as guidelines than as absolute rules [35]. Isolates that fall into a cluster at a certain threshold do not necessarily have to be epidemiologically linked [33]. Nevertheless, threshold values can be a valuable tool for a first delimitation of possibly linked isolates.

Despite the different numbers of loci in the different cgMLST schemes, application of unified thresholds yielded higher clustering concordance than application of scheme-specific thresholds. Slight adjustment of the thresholds could further increase concordance and led to identical clustering in four out of five tests of the translation code. However, method concordance did not reach 100%, even when using the same cgMLST scheme. This shows that not only the agreement on a specific scheme but also on specific software is important to achieve unambiguous comparability of clustering results.

SNP

While generally applicable typing methods may provide a valuable overview, further analyses on a subgroup of isolates will help to gain deeper insights. A potential outbreak cluster, for instance, can be initially identified by using cgMLST or SNP analysis with a general reference genome. A higher resolution of the closely related isolates within the cluster can be achieved by subsequent SNP analysis with a specific reference genome. To this end, the use of MLST CC-specific references has been proposed [16].

SNP analysis, limited to closely related isolates (in our case isolates of a specific MLST CC) using a closely related reference genome, reduced differences between the tools Snippy and BioNumerics when compared with analysis with a general reference. This indicates that such a restriction to closely related genomes improves method robustness. Additionally, in agreement with results from previous studies [17], using a closed or a draft specific reference genome did not have a decisive effect (neither on distance matrix nor on clustering). Both approaches have advantages and disadvantages. While a closed genome resolves repetitive regions, those will most probably not be present in a draft assembly. This phenomenon could decrease the number of detected SNPs actually present in unresolved regions and close to contig borders in a draft reference genome. Conversely, a draft genome from a certain study population is likely to have a higher degree of similarity to the rest of the isolates than a closed genome from a public repository, which may increase the core genome size and thus potentially the number of detected SNPs. As we have shown above (similar size of core genomes in SNP_Snippy with draft and closed reference genomes), the two effects (closeness and completeness) appeared to offset each other in our dataset. Therefore, if closed reference genomes specific for MLST CCs are not available, draft genomes from the dataset to be analysed can be used equivalently as references without losing analytical accuracy.

As an alternative to reference-based SNP calling, also reference-free, k-mer based approaches exist [36, 37]. They may have the advantage that no bias is introduced due to the selection of a certain reference. However, results are thereby highly dependent on the dataset and more difficult to compare than results derived from standardised, pre-defined references. Therefore, reference-based SNP calling using

pre-defined references in the form of MLST CC-specific reference genomes will lead to more standardised results when different datasets are compared.

While there was no difference between different software when using a specific reference genome, differences were large between Snippy and BioNumerics when using a general reference genome. The clustering threshold in the translation code for SNP_BioNumerics, in turn, was generally lower than for SNP_Snippy. As filter settings were similar in the two tools, the reasons for this effect remain unclear. Differences in the size of the core genome used in SNP analysis might have played a role. At this point, a major disadvantage of commercial over open-source tools becomes obvious. Although a closed-source software solution may be easier to use, open-source tools offer higher transparency since they allow for full comprehension of all steps in the analysis and provide intermediate and final results in standardised bioinformatics file formats.

CONCLUSIONS

In case of international disease outbreaks, for instance, one country needs to know whether related strains are found in other countries so that appropriate measures can be taken to prevent human infections. However, different laboratories frequently have different preferences for WGS data analysis. Such missing standards might hamper collaboration between sectors and countries [38]. Although web servers can be used for shared data analysis, the great advantage of local data analysis over submitting results to a web server is that the period between sequencing experiment and analysis results can be influenced. Of course, time also depends on the computational infrastructure at a certain institute. However, especially in time-critical applications like outbreak investigations, this could be a limiting factor. Even when primary analysis is performed locally, use of harmonised methods would open the opportunity to exchange intermediate results, like allele profiles in the case of cgMLST or variant files in the case of SNP analysis. These could then be used for global clustering with little computational effort.

Until harmonisation of methods is achieved, a translation code based on method concordances can work as a first filter to identify typing matches resulting from the different WGS analysis methods. This gives a new perspective for data exchange. The main advantage of our approach is the free choice of analysis tools, provided that there is good concordance with comparison methods. In this way, methods already established in a laboratory can be applied and uptake challenges of a method prescribed by another party are avoided.

Our translation code represents an average over the population structure of *L. monocytogenes* in the food chain in Germany. Despite the predominantly encouraging results from our practical test, we have seen that the WGS analysis methods may show better or worse agreement for individual clusters and different combinations of methods. This is also reflected in the range of adjusted Wallace coefficients achievable (Fig. 5). These coefficients must always be kept in mind when using the translation code, since they provide information about the probability of exact cluster matches between two methods of analysis. If these values are too low, the use of an alternative method on either side should be considered in order to improve adjusted Wallace coefficients before exchanging cluster information. However, good translatability (high clustering concordances at adjusted threshold values) between the majority of tested methods offers the valuable opportunity to minimise the amount of sequence data that needs to be exchanged and individually re-analysed. In this way, processes can be accelerated, which is an enormous advantage, especially in time-critical analyses of supraregional outbreaks.

Funding information

This work was supported by a grant of the Federal Ministry of Health (GE 2016 03 26) in the framework of the German Research Platform for Zoonoses and by the German Federal Institute for Risk Assessment (1322–668).

Acknowledgements

We would like to thank the European Reference Laboratory for *L. monocytogenes* and the French Agency for Food, Environmental and Occupational Health and Safety and especially Benjamin Félix for kindly granting us access to their BioNumerics calculation engine for cgMLST analysis. We also thank Ralf Dieckmann for his very valuable comments on the manuscript and Anna-Louisa Hauffe for her excellent technical assistance.

Author contributions

S. L., Conceptualisation, Methodology, Software, Formal Analysis, Investigation, Data Curation, Writing – Original Draft Preparation, Visualisation, Project administration; C. D., Conceptualisation, Methodology, Software, Formal Analysis, Data Curation, Writing – Original Draft Preparation; Visualisation; S. K., Conceptualisation, Resources, Writing – Review and Editing, Supervision, Funding; S. A. D., Conceptualisation, Resources, Writing – Review and Editing, Supervision, Funding.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. Allerberger F, Wagner M. Listeriosis: a resurgent foodborne infection. *Clin Microbiol Infect* 2010;16:16–23.
2. ECDC. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *Efsa J* 2018;16:e05500.
3. Swaminathan B, Gerner-Smidt P. The epidemiology of human listeriosis. *Microbes Infect* 2007;9:1236–1243.
4. Angelo KM, Conrad AR, Saupe A, Dragoo H, West N *et al*. Multistate outbreak of *Listeria monocytogenes* infections linked to whole apples used in commercially produced, prepackaged caramel apples: United States, 2014-2015. *Epidemiol Infect* 2017;145:848–856.
5. Chen Y, Luo Y, Curry P, Timme R, Melka D *et al*. 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* 2017;12:e0171389.
6. EFSA. Multi-country outbreak of *Listeria monocytogenes* serogroup IV B, multi-locus sequence type 6, infections linked to frozen corn and possibly to other frozen vegetables–first update. *EFSA Supporting Publications* 2018;15:1448E.
7. Goulet V, King LA, Vaillant V, de Valk H. What is the incubation period for listeriosis? *BMC Infect Dis* 2013;13:11.

8. Kleta S, Hammerl JA, Dieckmann R, Malorny B, Borowiak M *et al*. Molecular tracing to find source of protracted invasive listeriosis outbreak, southern Germany, 2012–2016. *Emerg Infect Dis* 2017;23:1680–1683.
9. Gelbíčová T, Zobaníková M, Tomáštíková Z, Van Walle I, Ruppitsch W *et al*. An outbreak of listeriosis linked to turkey meat products in the Czech Republic, 2012-2016. *Epidemiol Infect* 2018;146:1407–1412.
10. Schjørring S, Gillesberg Lassen S, Jensen T, Moura A, Kjeldgaard JS *et al*. Cross-border outbreak of listeriosis caused by cold-smoked salmon, revealed by integrated surveillance and whole genome sequencing (WGS), Denmark and France, 2015 to 2017. *Eurosurveillance* 2017;22:17–762.
11. Pietzka A, Allerberger F, Murer A, Lennkh A, Stöger A *et al*. Whole genome sequencing based surveillance of *L. monocytogenes* for early detection and investigations of Listeriosis outbreaks. *Front Public Health* 2019;7:139.
12. Ruppitsch W, Pietzka A, Prior K, Bletz S, Fernandez HL *et al*. Defining and evaluating a core genome multilocus sequence typing scheme for whole-genome sequence-based typing of *Listeria monocytogenes*. *J Clin Microbiol* 2015;53:2869–2876.
13. Moura A, Criscuolo A, Pouseele H, Maury MM, Leclercq A *et al*. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat Microbiol* 2016;2:16185.
14. Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J *et al*. chewBBACA: a complete suite for gene-by-gene schema creation and strain identification. *Microb Genom* 2018;4
15. Seemann T. Snippy - Rapid haploid variant calling and core genome alignment. https://github.com/tseemann/snippy 2015.
16. Møller Nielsen E, Björkman JT, Kiil K, Grant K, Dallman T *et al*. 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* 2017;14.
17. Kwong JC, Mercoulia K, Tomita T, Easton M, Li HY, *et al*. Prospective whole-genome sequencing enhances national surveillance of *Listeria monocytogenes*. *J Clin Microbiol* 2016;54:333–342.
18. Allard MW, Strain E, Melka D, Bunning K, Musser SM *et al*. Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J Clin Microbiol* 2016;54:1975–1983.
19. Wielinga PR, Hendriksen RS, Aarestrup FM, Lund O, Smits SL *et al*. *Global Microbial Identifier*. Springer: Applied Genomics of Food-borne Pathogens; 2017. pp. 13–31.
20. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I *et al*. Pulsenet international: vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. *Euro Surveill* 2017;22:30544.
21. Babraham Bioinformatics. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/ 2019.
22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 2014;30:2114–2120.
23. Deneke C, Tausch S. AQUAMIS (Assembly-based quality assessment for microbial isolate sequencing). https://gitlab.com/bfr_bioinformatics/AQUAMIS 2019.
24. Seemann T. mlst - Scan contig files against traditional PubMLST typing schemes. https://github.com/tseemann/mlst 2019.
25. Van Walle I, Björkman JT, Cormican M, Dallman T, Mossong J *et al*. Retrospective validation of whole genome sequencing-enhanced surveillance of listeriosis in Europe, 2010 to 2015. *Euro Surveill* 2018;23.
26. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW *et al*. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010;11:119.
27. Deneke C. chewieSnake - snakemake pipeline based on chewbbaca. https://gitlab.com/bfr_bioinformatics/chewieSnake 2019.
28. Zhou Z, Alikhan N-F, Sergeant MJ, Luhmann N, Vaz C *et al*. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Res* 2018;28:1395–1404.
29. Bush SJ, Foster D, Eyre DW, Clark EL, De Maio N *et al*. Genomic diversity affects the accuracy of bacterial single-nucleotide polymorphism-calling pipelines. *GigaScience* 2020;9
30. Deneke C. Variant calling pipeline with snippy. https://gitlab.com/bfr_bioinformatics/snippy-snake 2019.
31. Seemann T. 2018. snp-dists - Pairwise SNP distance matrix from a FASTA sequence alignment. https://github.com/tseemann/snp-dists
32. Carriço JA, Silva-Costa C, Melo-Cristino J, Pinto FR, de Lencastre H *et al*. Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. *J Clin Microbiol* 2006;44:2524–2532.
33. Pightling AW, Pettengill JB, Luo Y, Baugher JD, Rand H *et al*. Interpreting whole-genome sequence analyses of foodborne bacteria for regulatory applications and outbreak investigations. *Front Microbiol* 2018;9:1482.
34. Jackson BR, Tarr C, Strain E, Jackson KA, Conrad A *et al*. Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. *Clin Infect Dis* 2016;63:380–386.
35. Schürch AC, van Schaik W. Challenges and opportunities for whole-genome sequencing-based surveillance of antibiotic resistance. *Ann N Y Acad Sci* 2017;1388:108–120.
36. Jagadeesan B, Baert L, Wiedmann M, Orsi RH. Comparative analysis of tools and approaches for source tracking *Listeria monocytogenes* in a food facility using whole-genome sequence data. *Front Microbiol* 2019;10:947.
37. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics* 2015;31:2877–2878.
38. Rossen JWA, Friedrich AW, Moran-Gilad J, ESCMID Study Group for Genomic and Molecular Diagnostics (ESGMD). Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. *Clin Microbiol Infect* 2018;24:355–360.

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.

Supplement 1: Tools and parameters

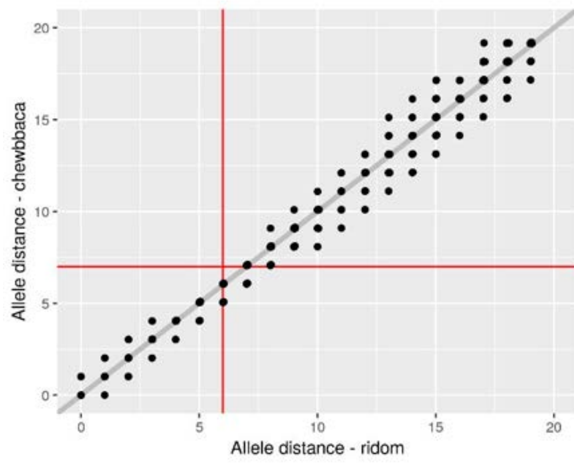
Tool	Version	Parameters
Trimmomatic	0.36	ILLUMINACLIP:NexteraPE-PE.fa:2:30:10 SLIDINGWINDOW:4:20 MINLEN:50
unicycler	0.4.4	--mode conservative --min_fasta_length 300
chewBBACA	2.0.12	--bsr: 0.6 # *BLAST score ratio (BSR) threshold* --ptf Listeria_monocytogenes.trn # *prodigal training file*
GrapeTree	1.4.1	--method distance --missing 0 # *(Number of different alleles)/(Number of loci that present in both genomes)*(Total number of loci)*
Snippy	4.0	mapqual: 60 # *Minimum read mapping quality to consider* basequal: 13 # *Minimum base quality to consider* mincov: 10 # *Minimum site depth to for calling alleles* minfrac: 0 # *Minumum proportion for variant evidence (0=AUTO)* minqual: 100 # *Minumum QUALITY in VCF column 6* maxsoft: 10 # *Maximum soft clipping to allow (default '10')*

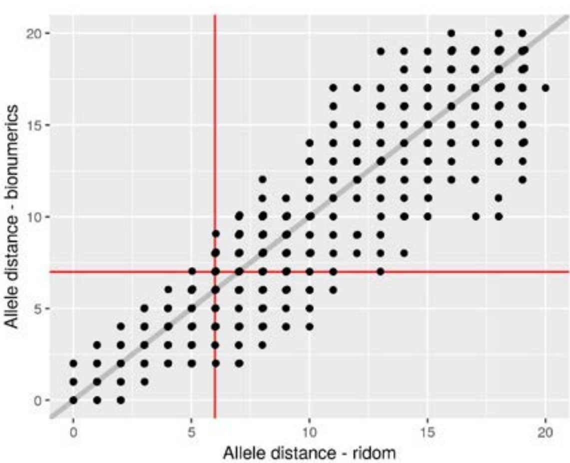
Supplement 2: Distance comparison

Visual comparison of the distances derived from different methods.
Each point represents a pairwise difference, with the point’s x-value being the distance in method 1 (marked on the x-axis) and the y-value being the distance in method 2 (marked on the y-axis). If the points are found on the diagonal regression line (grey), this indicates a high degree of agreement between the methods (also reflected in a high correlation value).
To highlight how many isolates are clustered differently by different tools, exemplary cluster thresholds for the individual methods are included in the figures, represented by a horizontal or vertical line (red). All points located in the lower left quadrant are assigned to the same cluster by both methods at the thresholds applied. If a point is either above this quadrant or to the right of it, this is a case where the clustering does not match. All points in the upper right quadrant are correctly assigned to no cluster.

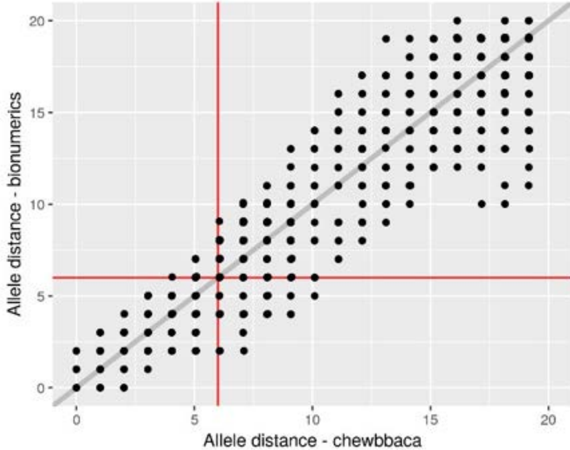
1. *cgMLST versus cgMLST*

cgMLST_Ridom_Ruppitsch versus
cgMLST_chewBBACA_Ruppitsch





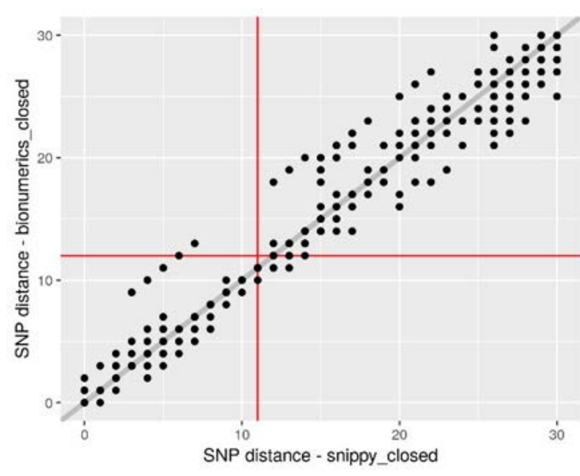
cgMLST_chewBBACA_Ruppitsch versus
cgMLST_BioNumerics_Moura



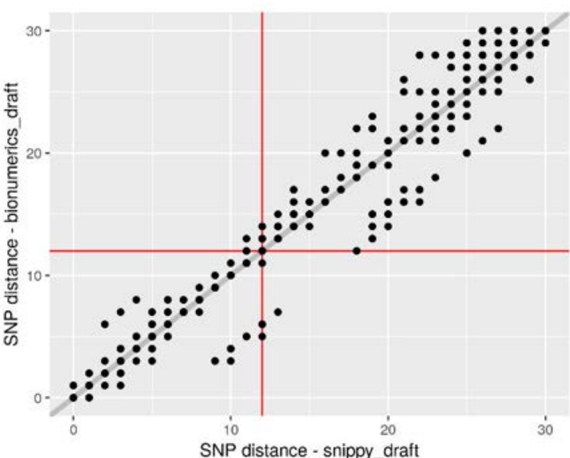
2. SNP versus SNP

2.1. Software-wise

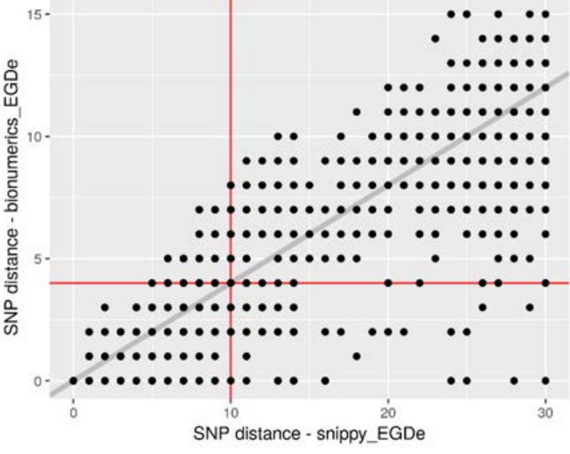
SNP_Snippy_spec._closed versus
SNP_BioNumerics_spec._closed



SNP_Snippy_spec._draft versus
SNP_BioNumerics_spec._draft

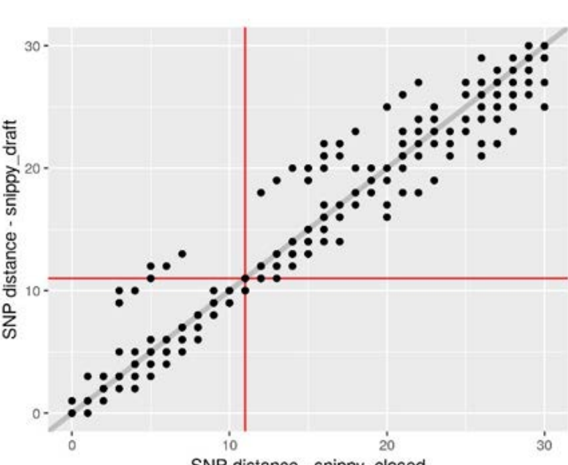


SNP_Snippy_EGDe versus
SNP_BioNumerics_EGDe

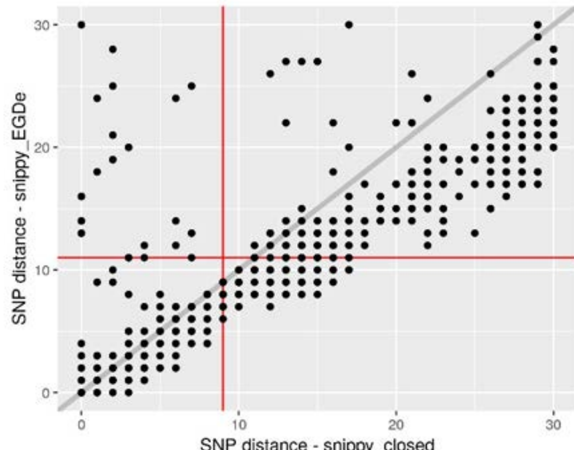


2.2. Reference genome-wise

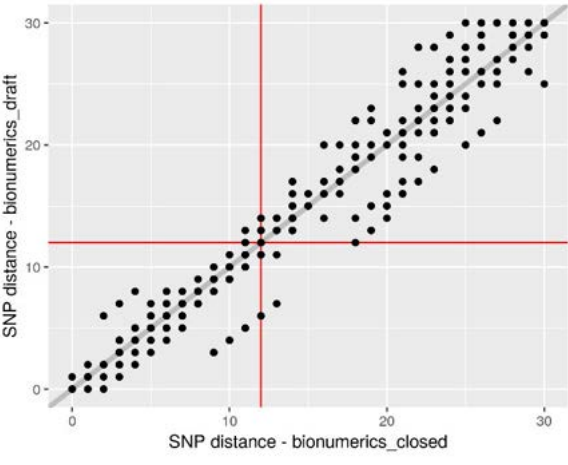
SNP_Snippy_spec._closed versus
SNP_Snippy_spec._draft



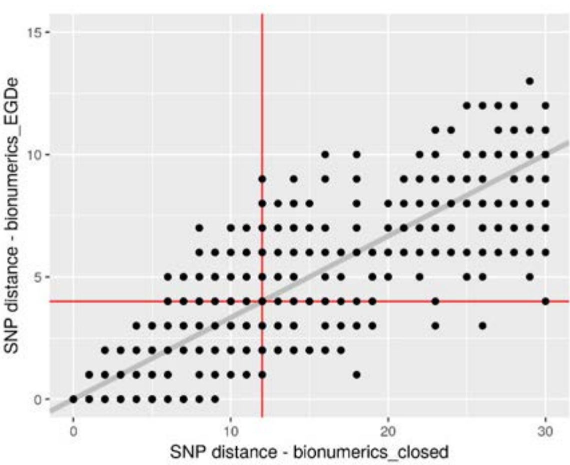
SNP_Snippy_spec._closed versus
SNP_Snippy_EGDe



SNP_BioNumerics_spec._closed versus
SNP_BioNumerics_spec._draft



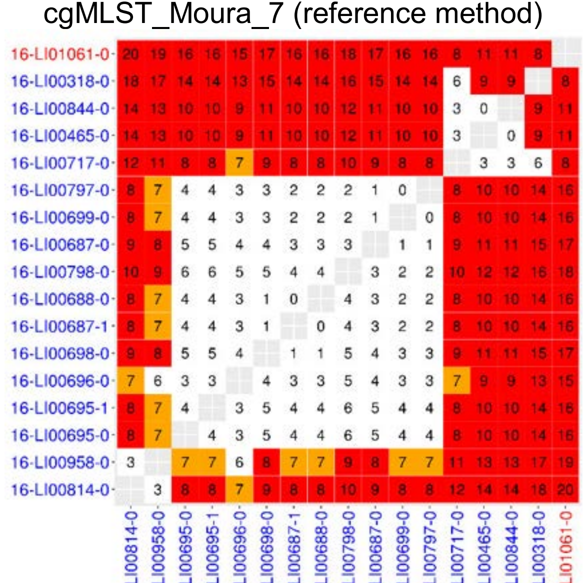
SNP_BioNumerics_spec._closed versus
SNP_BioNumerics_EGDe

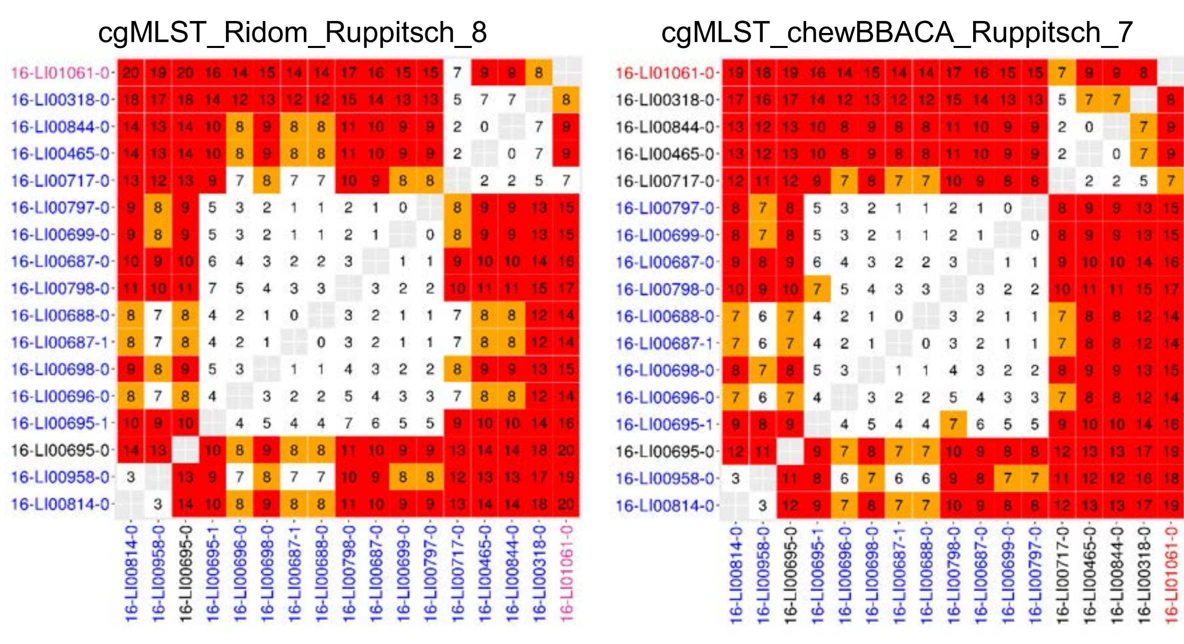


Supplement 3: Distance matrices for the translation code test on MLST CC121 isolates

Allele distances and SNP distances were compared for all applied methods.
Distances are coloured according to the method-specific (translational) threshold. White: distance value within threshold; red: above threshold; orange: borderline case where actual cgMLST distances are above thresholds (e.g. 7.2) but distances were rounded for better visibility.

The reference method was cgMLST_BioNumerics_Moura with an allele threshold of seven. Correct and wrong clustering of the different test methods is indicated through the colour code of isolate identifiers. Blue: correctly identified as a member of the cluster (compared to the reference method); black: incorrectly not assigned to cluster; red: correctly not assigned to cluster; pink: incorrectly assigned to cluster.

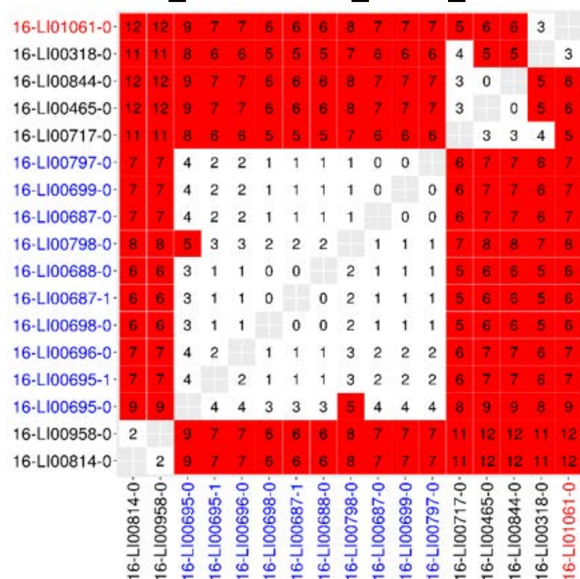




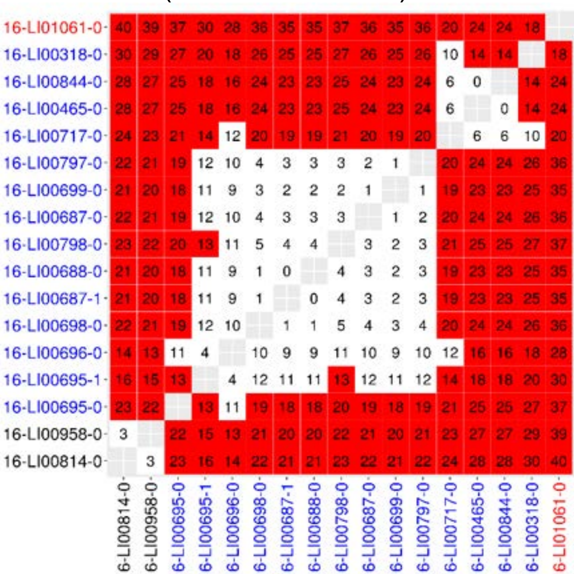
SNP_Snippy_EGDe_13

	16-LI00814-0	16-LI00958-0	16-LI00695-0	16-LI00695-1	16-LI00696-0	16-LI00698-0	16-LI00687-1	16-LI00688-0	16-LI00798-0	16-LI00687-0	16-LI00699-0	16-LI00797-0	16-LI00717-0	16-LI00465-0	16-LI00844-0	16-LI00318-0	16-LI01061-0
16-LI01061-0	30	28	32	24	22	23	23	23	25	32	23	24	14	16	16	12	
16-LI00318-0	24	22	26	18	16	17	17	17	19	26	17	18	8	10	10		12
16-LI00844-0	22	20	24	16	14	15	15	15	17	24	15	16	4	0		10	16
16-LI00465-0	22	20	24	16	14	15	15	15	17	24	15	16	4		0	10	16
16-LI00717-0	20	18	22	14	12	13	13	13	15	22	13	14		4	4	8	14
16-LI00797-0	12	10	14	6	4	3	3	3	3	10	1		14	16	16	18	24
16-LI00699-0	11	9	13	5	3	2	2	2	2	9		1	13	15	15	17	23
16-LI00687-0	20	18	22	14	12	11	11	11	11		9	10	22	24	24	26	32
16-LI00798-0	13	11	15	7	5	4	4	4		11	2	3	15	17	17	19	25
16-LI00688-0	11	9	13	5	3	0	0		4	11	2	3	13	15	15	17	23
16-LI00687-1	11	9	13	5	3	0		0	4	11	2	3	13	15	15	17	23
16-LI00698-0	11	9	13	5	3		0	0	4	11	2	3	13	15	15	17	23
16-LI00696-0	10	8	12	4		3	3	3	5	12	3	4	12	14	14	16	22
16-LI00695-1	12	10	14		4	5	5	5	7	14	5	6	14	16	16	18	24
16-LI00695-0	20	18		14	12	13	13	13	15	22	13	14	22	24	24	26	32
16-LI00958-0	2		18	10	8	9	9	9	11	18	9	10	18	20	20	22	28
16-LI00814-0		2	20	12	10	11	11	11	13	20	11	12	20	22	22	24	30

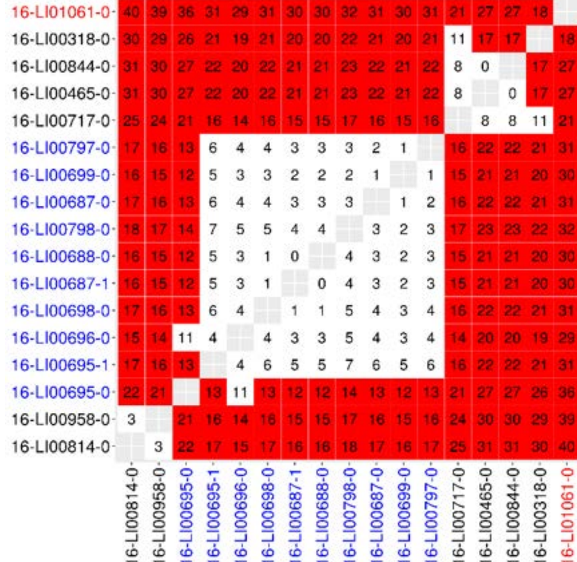
SNP_BioNumerics_EGDe_4



SNP_Snippy_spec._closed_12 and SNP_Snippy_spec._draft_12 (matrices identical)

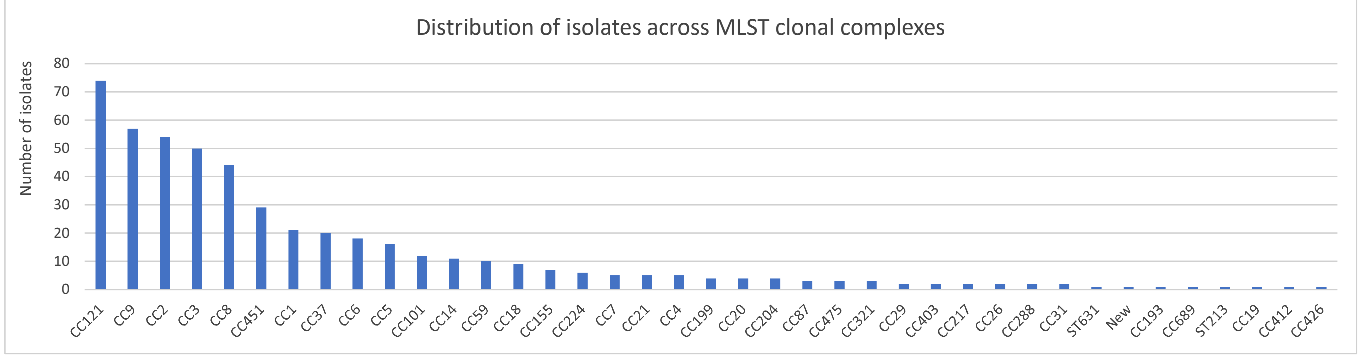


SNP_BioNumerics_spec._closed_11 and SNP_BioNumerics_spec._draft_11 (matrices identical)



Supplementary Table 1

MLST clonal complex	Number of isolates
CC121	74
CC9	57
CC2	54
CC3	50
CC8	44
CC451	29
CC1	21
CC37	20
CC6	18
CC5	16
CC101	12
CC14	11
CC59	10
CC18	9
CC155	7
CC224	6
CC7	5
CC21	5
CC4	5
CC199	4
CC20	4
CC204	4
CC87	3
CC475	3
CC321	3
CC29	2
CC403	2
CC217	2
CC26	2
CC288	2
CC31	2
ST631	1
New	1
CC193	1
CC689	1
ST213	1
CC19	1
CC412	1
CC426	1



Supplementary Table 2

Sample ID	MLST ST	MLST CC
16-LI00231-0	8	CC8
16-LI00232-0	9	CC9
16-LI00402-0	8	CC8
16-LI00415-0	8	CC8
16-LI00437-0	8	CC8
16-LI00508-0	29	CC29
16-LI00542-0	8	CC8
16-LI00583-0	8	CC8
16-LI00762-0	580	CC9
16-LI00781-0	8	CC8
16-LI00794-0	189	CC29
16-LI00851-0	8	CC8
16-LI00918-1	8	CC8
16-LI00962-0	9	CC9
16-LI00968-0	8	CC8
16-LI01044-0	8	CC8
16-LI01044-2	8	CC8
16-LI01091-0	8	CC8
16-LI01172-0	2117	CC8
16-LI01175-0	16	CC8
16-LI00843-0	8	CC8
16-LI00880-0	8	CC8
16-LI01065-0	451	CC451
16-LI01010-0	18	CC18
16-LI00820-0	37	CC37
16-LI00390-0	new	?
16-LI00254-0	2	CC2
16-LI00537-0	7	CC7
16-LI00248-0	18	CC18
16-LI00258-0	1	CC1
16-LI00280-0	101	CC101
16-LI00281-0	101	CC101
16-LI00538-0	1	CC1
16-LI00539-0	1	CC1
16-LI00540-0	1	CC1
16-LI00809-0	?	CC1
16-LI00991-0	18	CC18
16-LI01182-0	1	CC1
17-LI00032-0	26	CC26
16-LI00878-0	7	CC7
16-LI00287-0	1	CC1
16-LI00293-0	3	CC3
16-LI00315-0	4	CC4
16-LI00317-0	3	CC3
16-LI00356-0	87	CC87
16-LI00370-0	3	CC3
16-LI00377-0	1	CC1
16-LI00383-0	1	CC1
16-LI00411-0	1	CC1
16-LI00434-0	1	CC1
16-LI00434-1	1	CC1
16-LI00439-0	330	CC288
16-LI00466-0	2	CC2
16-LI00543-0	399	CC14
16-LI00600-0	3	CC3
16-LI00613-0	3	CC3
16-LI00632-0	5	CC5
16-LI00703-0	3	CC3
16-LI00706-0	3	CC3
16-LI00723-0	3	CC3
16-LI00792-0	?	CC288
16-LI00840-0	2	CC2
16-LI00859-0	3	CC3
16-LI00881-0	3	CC3
16-LI00908-0	3	CC3
16-LI00959-0	87	CC87
16-LI00979-0	2	CC2
16-LI00990-0	3	CC3
16-LI00996-0	1	CC1
16-LI01037-0	193	CC193
16-LI01043-3	2	CC2
16-LI01059-0	3	CC3
16-LI01073-0	8	CC8
16-LI01079-0	1	CC1
16-LI01081-0	87	CC87

16-LI01088-0	20	CC20
16-LI01098-0	3	CC3
16-LI01137-0	1	CC1
16-LI01141-0	5	CC5
16-LI01154-0	3	CC3
16-LI01159-0	3	CC3
16-LI01167-0	3	CC3
16-LI01173-0	1	CC1
16-LI01198-0	3	CC3
17-LI00005-0	3	CC3
17-LI00010-0	3	CC3
17-LI00024-0	2	CC2
17-LI00031-0	9	CC9
17-LI00068-0	3	CC3
16-LI00876-0	9	CC9
16-LI00897-0	?	CC2
16-LI00898-0	2	CC2
16-LI00884-0	2	CC2
16-LI00227-0	3	CC3
16-LI00220-0	3	CC3
16-LI00245-0	3	CC3
16-LI00253-0	2	CC2
16-LI00259-0	1	CC1
16-LI00279-0	101	CC101
16-LI00270-0	5	CC5
16-LI00282-0	101	CC101
16-LI00283-0	101	CC101
16-LI00284-0	101	CC101
16-LI00285-0	101	CC101
16-LI00300-0	3	CC3
16-LI00314-0	2	CC2
16-LI00322-0	2	CC2
16-LI00330-0	5	CC5
16-LI00332-0	5	CC5
16-LI00333-0	5	CC5
16-LI00334-0	5	CC5
16-LI00335-0	5	CC5
16-LI00336-0	5	CC5
16-LI00337-0	5	CC5
16-LI00338-0	5	CC5
16-LI00341-0	1	CC1
16-LI00347-0	3	CC3
16-LI00352-0	1	CC1
16-LI00354-0	3	CC3
16-LI00355-0	59	CC59
16-LI00298-0	3	CC3
16-LI00360-0	5	CC5
16-LI00366-0	1	CC1
16-LI00369-0	3	CC3
16-LI00386-0	3	CC3
16-LI00399-0	3	CC3
16-LI00412-1	4	CC4
16-LI00418-0	2	CC2
16-LI00428-0	3	CC3
16-LI00442-0	2	CC2
16-LI00455-0	224	CC224
16-LI00456-0	224	CC224
16-LI00457-0	3	CC3
16-LI00545-0	2	CC2
16-LI00565-0	2	CC2
16-LI00601-0	2	CC2
16-LI00632-1	5	CC5
16-LI00684-0	3	CC3
16-LI00691-0	59	CC59
16-LI00705-0	2	CC2
16-LI00707-0	2	CC2
16-LI00708-0	2	CC2
16-LI00726-0	2	CC2
16-LI00750-0	5	CC5
16-LI00775-0	2	CC2
16-LI00784-0	3	CC3
16-LI00788-0	2	CC2
16-LI00861-0	9	CC9
16-LI00864-0	2	CC2
16-LI00926-0	224	CC224
16-LI00938-0	2	CC2
16-LI00955-0	224	CC224
16-LI00957-0	31	CC31

16-LI00961-0	2	CC2
16-LI00966-0	9	CC9
16-LI00969-0	2	CC2
16-LI00972-0	145	CC2
16-LI00982-0	3	CC3
16-LI00985-0	3	CC3
16-LI01014-0	2	CC2
16-LI01025-0	2	CC2
16-LI01032-0	2	CC2
16-LI01038-0	2	CC2
16-LI01043-0	2	CC2
16-LI01043-1	2	CC2
16-LI01043-2	2	CC2
16-LI01075-0	9	CC9
16-LI01080-0	3	CC3
16-LI01096-0	2	CC2
16-LI01103-0	426	CC426
16-LI01109-0	3	CC3
16-LI01125-0	3	CC3
16-LI01138-0	9	CC9
16-LI01139-0	3	CC3
16-LI01142-0	2	CC2
16-LI01165-0	9	CC9
16-LI01168-0	2	CC2
16-LI01169-0	3	CC3
16-LI01179-0	2	CC2
16-LI01183-0	631	ST631
16-LI01185-0	2	CC2
16-LI01189-0	3	CC3
16-LI01195-0	3	CC3
16-LI01199-0	3	CC3
16-LI01211-0	2	CC2
16-LI01213-0	219	CC4
17-LI00001-0	3	CC3
17-LI00003-0	3	CC3
17-LI00004-0	3	CC3
17-LI00026-0	1	CC1
17-LI00049-0	2	CC2
16-LI00956-0	31	CC31
16-LI00863-0	9	CC9
16-LI00867-0	2	CC2
16-LI00868-0	2	CC2
16-LI00872-0	2	CC2
16-LI00882-0	2	CC2
16-LI00883-0	2	CC2
16-LI00909-0	3	CC3
17-LI00002-0	3	CC3
16-LI00222-0	59	CC59
16-LI00244-0	59	CC59
16-LI00262-0	2	CC2
16-LI00223-0	59	CC59
16-LI00288-0	2	CC2
16-LI00303-0	6	CC6
16-LI00312-0	6	CC6
16-LI00358-0	101	CC101
16-LI00363-0	?	CC2
16-LI00367-0	6	CC6
16-LI00391-0	224	CC224
16-LI00423-0	2	CC2
16-LI00426-0	6	CC6
16-LI00427-0	6	CC6
16-LI00438-0	6	CC6
16-LI00480-0	4	CC4
16-LI00494-0	6	CC6
16-LI00495-0	6	CC6
16-LI00498-0	59	CC59
16-LI00578-0	504	CC475
16-LI00584-0	59	CC59
16-LI00685-0	6	CC6
16-LI00689-0	101	CC101
16-LI00701-0	5	CC5
16-LI00702-0	5	CC5
16-LI00704-0	122	CC9
16-LI00749-0	155	CC155
16-LI00751-0	6	CC6
16-LI00778-0	6	CC6
16-LI00780-0	6	CC6
16-LI00782-0	6	CC6

16-LI00786-0	155	CC155
16-LI00847-0	451	CC451
16-LI00849-0	6	CC6
16-LI00850-0	101	CC101
16-LI00853-0	101	CC101
16-LI00862-0	155	CC155
16-LI00925-0	6	CC6
16-LI00963-0	504	CC475
16-LI00971-0	59	CC59
16-LI00983-0	504	CC475
16-LI00993-0	217	CC217
16-LI01013-0	217	CC217
16-LI01035-0	2	CC2
16-LI01036-0	59	CC59
16-LI01078-0	7	CC7
16-LI01104-0	6	CC6
16-LI01126-0	6	CC6
16-LI01140-0	224	CC224
16-LI01156-0	2	CC2
16-LI01157-0	4	CC4
16-LI01162-0	155	CC155
16-LI01204-0	6	CC6
16-LI01209-0	689	CC689
17-LI00028-0	213	ST213
17-LI00029-0	26	CC26
17-LI00030-0	101	CC101
16-LI00899-0	59	CC59
16-LI00900-0	2	CC2
16-LI00216-0	91	CC14
16-LI00353-0	155	CC155
16-LI00778-1	321	CC321
16-LI00858-0	37	CC37
16-LI01099-0	155	CC155
16-LI01105-0	2079	CC199
16-LI01106-0	121	CC121
17-LI00050-0	121	CC121
16-LI00911-0	155	CC155
16-LI00292-0	321	CC321
16-LI00304-0	321	CC321
16-LI00309-0	18	CC18
16-LI00319-0	18	CC18
16-LI00359-0	121	CC121
16-LI00361-0	204	CC204
16-LI00368-0	121	CC121
16-LI00385-0	173	CC19
16-LI00400-0	199	CC199
16-LI00401-0	199	CC199
16-LI00500-0	18	CC18
16-LI00502-0	199	CC199
16-LI00544-0	21	CC21
16-LI00563-0	?	CC121
16-LI00569-0	121	CC121
16-LI00621-0	18	CC18
16-LI00654-0	9	CC9
16-LI00710-0	9	CC9
16-LI00725-0	20	CC20
16-LI00778-3	451	CC451
16-LI00795-0	451	CC451
16-LI00797-0	121	CC121
16-LI00812-0	403	CC403
16-LI00838-0	451	CC451
16-LI00848-0	91	CC14
16-LI00922-0	204	CC204
16-LI00943-0	9	CC9
16-LI00992-0	18	CC18
16-LI01005-0	14	CC14
16-LI01042-0	18	CC18
16-LI01072-0	121	CC121
16-LI01094-0	8	CC8
16-LI01100-0	121	CC121
16-LI01110-0	91	CC14
16-LI01135-0	121	CC121
16-LI01155-0	412	CC412
16-LI01161-0	14	CC14
16-LI01180-0	121	CC121
17-LI00006-0	7	CC7
17-LI00007-0	7	CC7
17-LI00051-0	9	CC9

17-LI00056-0	451	CC451
16-LI00926-1	451	CC451
16-LI00865-1	37	CC37
16-LI00912-0	9	CC9
16-LI00226-0	9	CC9
16-LI00289-0	9	CC9
16-LI00228-0	37	CC37
16-LI00246-0	121	CC121
16-LI00263-0	37	CC37
16-LI00264-0	9	CC9
16-LI00200-0	451	CC451
16-LI00294-0	37	CC37
16-LI00295-0	37	CC37
16-LI00296-0	9	CC9
16-LI00306-0	20	CC20
16-LI00307-0	451	CC451
16-LI00308-0	9	CC9
16-LI00318-0	121	CC121
16-LI00325-0	121	CC121
16-LI00326-0	121	CC121
16-LI00328-0	8	CC8
16-LI00339-0	9	CC9
16-LI00340-0	8	CC8
16-LI00345-0	9	CC9
16-LI00346-0	9	CC9
16-LI00350-0	451	CC451
16-LI00362-0	204	CC204
16-LI00365-0	204	CC204
16-LI00375-0	121	CC121
16-LI00384-0	121	CC121
16-LI00395-0	?	CC37
16-LI00406-0	121	CC121
16-LI00407-0	9	CC9
16-LI00408-0	37	CC37
16-LI00410-1	121	CC121
16-LI00412-0	9	CC9
16-LI00464-0	451	CC451
16-LI00465-0	121	CC121
16-LI00499-0	121	CC121
16-LI00501-0	9	CC9
16-LI00503-0	8	CC8
16-LI00504-0	8	CC8
16-LI00506-0	8	CC8
16-LI00507-0	8	CC8
16-LI00541-0	121	CC121
16-LI00564-0	451	CC451
16-LI00571-0	451	CC451
16-LI00572-0	9	CC9
16-LI00582-0	121	CC121
16-LI00625-0	121	CC121
16-LI00639-0	14	CC14
16-LI00651-0	121	CC121
16-LI00652-0	121	CC121
16-LI00667-0	120	CC8
16-LI00686-0	121	CC121
16-LI00687-0	121	CC121
16-LI00687-1	121	CC121
16-LI00688-0	121	CC121
16-LI00690-0	8	CC8
16-LI00695-0	121	CC121
16-LI00695-1	121	CC121
16-LI00696-0	121	CC121
16-LI00698-0	121	CC121
16-LI00699-0	121	CC121
16-LI00716-0	121	CC121
16-LI00717-0	121	CC121
16-LI00730-0	451	CC451
16-LI00732-0	9	CC9
16-LI00733-0	9	CC9
16-LI00734-0	121	CC121
16-LI00738-0	451	CC451
16-LI00739-0	451	CC451
16-LI00740-0	451	CC451
16-LI00741-0	451	CC451
16-LI00742-0	451	CC451
16-LI00748-0	8	CC8
16-LI00752-0	121	CC121
16-LI00759-0	451	CC451

16-LI00778-2	37	CC37
16-LI00780-1	236	CC121
16-LI00783-0	37	CC37
16-LI00785-0	9	CC9
16-LI00798-0	121	CC121
16-LI00801-0	121	CC121
16-LI00802-0	16	CC8
16-LI00803-0	121	CC121
16-LI00804-0	14	CC14
16-LI00808-0	14	CC14
16-LI00810-0	9	CC9
16-LI00813-0	121	CC121
16-LI00814-0	121	CC121
16-LI00821-0	9	CC9
16-LI00822-0	451	CC451
16-LI00836-0	399	CC14
16-LI00837-0	451	CC451
16-LI00839-0	451	CC451
16-LI00841-0	9	CC9
16-LI00842-0	9	CC9
16-LI00857-0	8	CC8
16-LI00860-0	8	CC8
16-LI00904-0	236	CC121
16-LI00905-0	121	CC121
16-LI00906-0	37	CC37
16-LI00915-0	121	CC121
16-LI00916-0	9	CC9
16-LI00917-0	121	CC121
16-LI00918-0	9	CC9
16-LI00919-0	9	CC9
16-LI00920-0	9	CC9
16-LI00921-0	121	CC121
16-LI00923-0	121	CC121
16-LI00939-0	121	CC121
16-LI00942-0	9	CC9
16-LI00944-0	9	CC9
16-LI00950-0	37	CC37
16-LI00951-0	121	CC121
16-LI00952-0	121	CC121
16-LI00953-0	121	CC121
16-LI00954-0	21	CC21
16-LI00958-0	121	CC121
16-LI00960-0	9	CC9
16-LI00962-1	121	CC121
16-LI00964-0	121	CC121
16-LI00965-0	20	CC20
16-LI00969-1	9	CC9
16-LI00975-0	8	CC8
16-LI00976-0	451	CC451
16-LI00977-0	451	CC451
16-LI00978-0	8	CC8
16-LI00984-0	9	CC9
16-LI00986-0	9	CC9
16-LI00994-0	403	CC403
16-LI01004-0	121	CC121
16-LI01034-0	451	CC451
16-LI01044-1	8	CC8
16-LI01045-0	21	CC21
16-LI01054-0	8	CC8
16-LI01055-0	8	CC8
16-LI01056-0	8	CC8
16-LI01061-0	121	CC121
16-LI01062-0	121	CC121
16-LI01064-0	9	CC9
16-LI01066-0	37	CC37
16-LI01066-1	9	CC9
16-LI01067-0	121	CC121
16-LI01068-0	8	CC8
16-LI01069-0	121	CC121
16-LI01074-0	9	CC9
16-LI01075-1	14	CC14
16-LI01076-0	9	CC9
16-LI01077-0	9	CC9
16-LI01090-0	8	CC8
16-LI01095-0	451	CC451
16-LI01107-0	121	CC121
16-LI01108-0	121	CC121
16-LI01127-0	451	CC451

16-LI01131-0	121	CC121
16-LI01132-0	121	CC121
16-LI01134-0	9	CC9
16-LI01145-0	16	CC8
16-LI01153-0	8	CC8
16-LI01158-0	9	CC9
16-LI01163-0	121	CC121
16-LI01164-0	121	CC121
16-LI01178-0	121	CC121
16-LI01193-0	121	CC121
16-LI01194-0	?	CC37
16-LI01201-0	9	CC9
16-LI01206-0	37	CC37
16-LI01212-0	21	CC21
17-LI00008-0	9	CC9
17-LI00009-0	121	CC121
17-LI00012-0	8	CC8
17-LI00025-0	8	CC8
17-LI00027-0	8	CC8
16-LI00844-0	121	CC121
16-LI00865-0	37	CC37
16-LI00865-2	37	CC37
16-LI00865-3	37	CC37
16-LI00865-4	37	CC37
16-LI00869-0	451	CC451
16-LI00873-0	9	CC9
16-LI00888-0	21	CC21
16-LI00910-0	9	CC9
16-LI00913-0	121	CC121

Supplementary Table 3

MLST clonal complex	Reference_type	Reference accesion	Reference_length	Sample_number_in_group	Number_maskedPositions	Number_missingPositions	Number_missingmaskedPositions	CoreGenomeSize	RelativeCoreGenomeSize
all	EGDe	NC003210	2,944,528	495	536,844	449,218	663,520	2,281,008	0.77
CC1	closed	AE017262	2,905,187	22	65,267	36,827	99,343	2,805,844	0.97
CC101	closed	CP025221	2,987,434	13	95,162	112,440	184,926	2,802,508	0.94
CC121	closed	HG13249	3,010,620	75	136,427	162,971	249,064	2,761,556	0.92
CC155	closed	CP002004	2,874,431	8	23,392	32,334	53,941	2,820,490	0.98
CC18	closed	CP020830	2,905,907	10	47,871	69,242	110,415	2,795,492	0.96
CC2	closed	CP006046	3,032,269	55	133,886	160,311	254,033	2,778,236	0.92
CC224	closed	CP016629	2,935,033	7	34,722	62,397	94,941	2,840,092	0.97
CC3	closed	CP006594	3,034,043	51	126,010	162,536	240,248	2,793,795	0.92
CC37	closed	CP011397	2,918,170	21	52,551	51,570	97,265	2,820,905	0.97
CC4	closed	FM242711	2,912,690	5	32,108	29,556	60,360	2,852,330	0.98
CC5	closed	CP006592	2,943,218	17	59,489	67,821	121,211	2,822,007	0.96
CC6	closed	CP006047	2,947,460	19	40,018	72,254	98,902	2,848,558	0.97
CC7	closed	CP002002	2,903,106	6	46,455	70,094	112,138	2,790,968	0.96
CC8	closed	CP006862	2,994,351	42	103,811	161,739	227,003	2,767,348	0.92
CC9	closed	FR733649	2,972,172	58	96,248	127,109	189,525	2,782,647	0.94
CC1	draft	16-LI00258-0	2,870,559	22	66,919	2,694	67,645	2,802,914	0.98
CC101	draft	16-LI00284-0	2,912,769	13	82,122	38,248	111,209	2,801,560	0.96
CC121	draft	16-LI01132-0	2,986,125	75	106,012	68,601	160,351	2,825,774	0.95
CC155	draft	16-LI00862-0	3,042,694	8	34,411	199,376	221,871	2,820,823	0.93
CC18	draft	16-LI00319-0	2,922,840	10	51,935	70,677	113,537	2,809,303	0.96
CC2	draft	16-LI01038-0	2,956,695	55	117,839	84,401	178,962	2,777,733	0.94
CC224	draft	16-LI00391-0	2,875,971	7	33,304	1,344	33,882	2,842,089	0.99
CC3	draft	16-LI00227-0	2,931,884	51	94,262	62,323	141,858	2,790,026	0.95
CC37	draft	16-LI00295-0	2,887,752	21	51,130	19,649	66,222	2,821,530	0.98
CC4	draft	16-LI00480-0	2,914,700	5	33,718	41,217	73,849	2,840,851	0.97
CC5	draft	16-LI00750-0	2,916,879	17	60,558	44,813	100,589	2,816,290	0.97
CC6	draft	16-LI00782-0	2,918,716	19	38,476	41,902	69,617	2,849,099	0.98
CC7	draft	17-LI00007-0	2,911,815	6	47,096	79,094	121,689	2,790,126	0.96
CC8	draft	16-LI00415-0	3,005,032	42	106,469	169,227	233,389	2,771,643	0.92
CC9	draft	16-LI00873-0	2,962,301	58	93,688	116,649	179,581	2,782,720	0.94

Article

Contamination Pathways can Be Traced along the Poultry Processing Chain by Whole Genome Sequencing of *Listeria innocua*

Mayada Gwida [1,†], Stefanie Lüth [2,3,†,*], Maged El-Ashker [4], Amira Zakaria [5], Fatma El-Gohary [1], Mona Elsayed [1], Sylvia Kleta [2] and Sascha Al Dahouk [2,6]

1 Department of Hygiene and Zoonoses, Faculty of Veterinary Medicine, Mansoura University, Mansoura 35516, Egypt; mayada.gwida@gmail.com (M.G.); dr.fatmagohary@gmail.com (F.E.-G.); dr.monamohy@yahoo.com (M.E.)
2 Department of Biological Safety, German Federal Institute for Risk Assessment, Max-Dohrn-Straße 8-10, 10589 Berlin, Germany; sylvia.kleta@bfr.bund.de (S.K.); Sascha.Al-Dahouk@gmx.de (S.A.D.)
3 Institute for Biology, Freie Universität Berlin, Königin-Luise-Straße 1-3, 14195 Berlin, Germany
4 Department of Internal Medicine, Infectious and Fish Diseases, Faculty of Veterinary Medicine, Mansoura University, Mansoura 35516, Egypt; maged.elashker1978@gmail.com
5 Department of Food Hygiene and Control, Faculty of Veterinary Medicine, Mansoura University, Mansoura 35516, Egypt; amera.zakaria@yahoo.com
6 RWTH Aachen University Hospital, Pauwelsstraße 30, 52074 Aachen, Germany
* Correspondence: stefanie.lueth@bfr.bund.de
† These authors contributed equally to the work.

Received: 29 January 2020; Accepted: 13 March 2020; Published: 14 March 2020

Abstract: Foodborne infection with *Listeria* causes potentially life-threatening disease listeriosis. *Listeria monocytogenes* is widely recognized as the only species of public health concern, and the closely related species *Listeria innocua* is commonly used by the food industry as an indicator to identify environmental conditions that allow for presence, growth, and persistence of *Listeria* spp. in general. In our study, we analyze the occurrence of *Listeria* spp. in a farm-to-fork approach in a poultry production chain in Egypt and identify bacterial entry gates and transmission systems. Prevalence of *Listeria innocua* at the three production stages (farm, slaughterhouse, food products) ranged from 11% to 28%. The pathogenic species *Listeria monocytogenes* was not detected, and *Listeria innocua* strains under study did not show genetic virulence determinants. However, the close genetic relatedness of *Listeria innocua* isolates (maximum 63 SNP differences) indicated cross-contamination between all stages from farm to final food product. Based on these results, chicken can be seen as a natural source of *Listeria*. Last but not least, sanitary measures during production should be reassessed to prevent bacterial contamination from entering the food chain and to consequently prevent human listeriosis infections. For this purpose, surveillance must not be restricted to pathogenic species.

Keywords: *Listeria innocua*; poultry production; food safety; whole-genome sequencing; single nucleotide polymorphism; listeriosis; *Listeria monocytogenes*

1. Introduction

Within the bacterial genus *Listeria*, *Listeria (L.) monocytogenes*, the causative agent of listeriosis, is widely recognized as the only species of public health concern. In humans, the clinical picture of listeriosis varies from mild to life-threatening illness with a mortality rate of 20% to 30% on average [1]. Severe cases appear mainly in vulnerable populations like pregnant women, newborns, immunocompromised, or older people [2]. The majority of listeriosis cases are foodborne [1]. Although

closely related to *L. monocytogenes* (perfect synteny of genome organization), the species *L. innocua* is considered as non-pathogenic because it lacks the typical virulence genes [3]. However, rare atypical hemolytic *L. innocua* strains have been described [4,5] and proven to be virulent in in vivo assays using mouse or zebrafish models [6]. Furthermore, *L. innocua* was isolated from two human patients suffering from fatal sepsis [7] and acute meningitis [8].

Listeria monocytogenes, as well as *L. innocua*, are widespread in natural environments such as soil, surface water, sewage, or feces of mammals and birds [9,10]. Livestock animals like poultry can be asymptomatic carriers of *Listeria* spp., and thereby lead to the unnoticed entry of the bacteria into the food chain via contaminated raw animal products [11,12]. Due to its undemanding nature and high adaptability, *Listeria* is able to persist in food production plants, which may lead to continuous contamination of food through contact with previously contaminated surfaces [13]. Raw animal products are either primarily contaminated or cross-contaminated during food-processing [12]. In this way, *Listeria* can finally end up in the food chain, posing a health threat for consumers. Detailed knowledge of entry gates and transmission routes is, therefore, indispensable to prevent food contamination and human listeriosis cases.

Because of the high similarity of *Listeria* species in terms of distribution and adaptability, the transmission path identified for one species is assumed to be transferable to other species. Hence, *L. innocua* is commonly used by the food industry as an indicator to identify environmental conditions that allow for the presence, growth, and persistence of the relevant human pathogen *L. monocytogenes* [9,14]. Using *L. innocua* as a model for *Listeria* contamination in general can help to improve surveillance and hygiene measures, which will consequently prevent human infections. However, systematic data on the prevalence of *Listeria* spp. in poultry slaughterhouses are still limited. In our study, we therefore analyze the occurrence of *Listeria* spp. in a farm-to-fork approach from the primary production stage to the final food product to reveal bacterial transmission routes. To trace *Listeria* spp. along the food chain, we applied whole-genome sequencing techniques.

2. Materials and Methods

2.1. Sample Collection

In total, 210 samples were collected from a single commercial poultry farm with five chicken flocks in separate henhouses and an affiliated slaughterhouse in 2017. The farm and slaughterhouse were located in Dakahlia Governorate, Egypt. The farm owners were asked to sign consent forms after being informed about the aims and goals of our research project and the potential health risks associated with the contamination of food products with *Listeria* spp. The study followed the ethical guidelines of Mansoura University and was approved by the responsible ethics committee (Code No. R/15).

Samples were collected from three different sources along the food production chain: on the farm, in the slaughterhouse, and from the final chicken products. Two weeks before slaughtering, 25 samples were collected on the chicken farm, all on the same day in November 2017. For soiled litter (containing fresh fecal droppings), poultry feed, drinking water, and the walls of the henhouses, samples were pooled per sampling site (n = 5) and per flock (n = 5). Five samples (20 g each) were collected from the top few centimeters of soiled litter in different locations of every henhouse (close to drinking troughs and feeding stations, from walls and near the center) and then mixed to form a composite sample of 100 g for the specific flock. The poultry feed samples were taken from the five different feeding stations of each henhouse. The water samples (20 mL each) were collected from five different drinking troughs of each henhouse and pooled to yield a composite sample of 100 mL. Farm walls were swabbed on five different sites inside each henhouse using a sterile template of 25 cm^2 and samples from workers' hands were collected from five individual workers. In the slaughterhouse, 13 cloacal swabs were collected from each of the five flocks just before slaughter (n = 65). The birds were randomly selected. In addition, 20 surface swabs (10 from slaughterhouse walls, 5 from tables and 5 from knives) were

taken during processing. Another 100 samples were taken after slaughter, including swabs from 80 whole carcasses, 10 chicken fillets, and 10 livers.

Hand swabs (palm, between fingers, fingertips, and fingernails) were essentially carried out according to the protocol of Genigeorgis and colleagues [15]. We used buffered peptone water (BPW; Oxoid, Basingstoke, UK) to moisten the cotton swabs and as enrichment broth. Soiled litter, poultry feed, and drinking water were sampled following standard procedures using sterilized spatulas or syringes [16,17]. The samples were homogenized with a stomacher and stored in sterile bags at 4 to 8 °C until transport to the laboratory. Swabs from surfaces of walls and tables in the slaughterhouse were collected according to the guidelines of the American Public Health Association [18]. Briefly, four 100 cm^2 regions of the sampling site were swabbed with sterile sponges moistened with 40 mL of BPW in several horizontal and vertical movements. The sponges were then transferred to sterile bags containing 160 mL of BPW to yield a final volume of 200 mL. The farm walls were swabbed using the same technique. The two sides of the butcher's knives were swabbed with BPW-moistened cotton swabs instead of sterile sponges. Cloacal swabs (from the mucosal wall) were collected from living chicken with sterile cotton-tipped swabs pre-moistened in BPW. Swabs from chicken carcasses were collected after evisceration using the method described by McEvoy and colleagues [19]. Briefly, each swab was moistened just before use with 25 mL of BPW and put into a sterile plastic bag after sampling. Chicken fillet and liver samples (~25 g each) were sliced with a sterile scalpel and put into a sterile stomacher bag. All samples were processed under aseptic conditions and then directly sent to the laboratory for further analyses.

2.2. *Listeria Isolation and Identification*

Listeria spp. were isolated and identified as described in the Bacteriological Analytical Manual of the U.S. Food and Drug Administration [20]. For solid samples (poultry feed, chicken fillet, and liver) and water samples, 25 g or ml were added to 225 mL *Listeria* enrichment broth without antibiotic supplement, pH 8.6 (Oxoid) and homogenized in a stomacher for two minutes. The swab samples were transferred to 10 mL *Listeria* enrichment broth. Homogenates of solid samples, water samples, and swabs were incubated at 30 °C for 4 h. Then, *Listeria* selective enrichment supplements (Oxoid), including nalidixic acid, cycloheximide, and acriflavine, were added and the broth cultures were incubated at the same temperature for another 24 to 48 h. An inoculation loop of the enriched sample was incubated on Oxford agar (Oxoid) at 35 °C for 24 to 48 h. At least five colonies showing a black halo characteristic for *Listeria* spp. were picked, transferred onto tryptic soy agar plates with 0.6% yeast extract and incubated at 30 °C for 24 to 48 h. These presumptive *Listeria* isolates were stored at –80 °C in brain heart infusion with 20% glycerol.

2.3. *Matrix-Assisted Laser Desorption/Ionisation Time-of-Flight (MALDI-TOF) Mass Spectrometry*

After thawing, the bacteria were plated onto sheep blood agar and incubated overnight at 37 °C. Mass spectrometry samples were prepared using the direct smear method [21]. Species identification was performed using the MALDI Biotyper® Subtyping Module (Bruker Daltonik, Bremen, Germany) according to the manufacturer's instructions.

2.4. *Genomic DNA Extraction and Next Generation Sequencing*

All strains identified as *Listeria* spp. by MALDI–TOF MS were again grown on sheep blood agar overnight at 37 °C. Bacterial cells were harvested and lysed following the Pulse Net standardized laboratory protocol for whole-genome sequencing of Gram-positive bacteria (https://www.cdc.gov/pulsenet/pdf/pnl32-miseq-nextera-xt.pdf). We extracted DNA with the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions.

Sequencing libraries were constructed using the Nextera XT Sample Preparation Kit (Illumina, Inc., San Diego, CA, USA). Sequencing was performed in paired-end mode with 2 × 300 bp reads on the Illumina MiSeq sequencer using the MiSeq Reagent v3 600-cycle Kit (Illumina).

2.4.1. Multilocus Sequence Typing (MLST)

Multilocus sequences types (ST) and clonal complexes (CC) were determined according to the scheme available at https://bigsdb.pasteur.fr/listeria/listeria.html.

2.4.2. Single Nucleotide Polymorphism (SNP) Analysis

Sequences were trimmed with Trimmomatic version 0.36 [22] using default parameters. Trimmed reads were mapped to the closed reference genome of *L. innocua* (NC_003212.1) in BioNumerics version 7.6 (Applied Maths, Gent, Belgium), followed by SNP calling. Strict SNP filtering with software default parameters was applied.

2.4.3. In Silico Screening for Virulence Factors

Trimmed reads were de novo assembled with SPAdes version 3.11.1 [23]. Assembled genomes were used for virulence gene screening with ABRicate version 0.8 [24] using the Virulence Factor Database (VFDB) (2597 sequences, [25], last updated 9 July 2019). The *L. monocytogenes* reference strain EGDe (NC_003210.1) was included in the screening as a representative for a pathogenic strain. A cut-off of at least 80% gene identity was applied for gene presence.

2.4.4. Data Storage

The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB36384.

3. Results

3.1. Prevalence of Listeria spp. along the Poultry Production Chain

Listeria innocua was the only *Listeria* species identified by MALDI–TOF MS and was isolated from 17% (36/210) of the samples along the poultry production chain (Table 1). The prevalence of *L. innocua* on the farm level was 28% (7/25). *Listeria innocua* was found in 100% (5/5) of the swabs from farm walls and in 20% (1/5) of the samples from soiled litter or poultry feed. Both workers' hands and drinking water tested negative for *L. innocua*. A total of 9 out of 85 samples (11%) collected in the slaughterhouse revealed *L. innocua*. Tables and abattoir walls were contaminated, with 40% (2/5) and 20% (2/10) positive swabs, respectively, whereas knives were tested negative. Five out of 85 (7%) cloacal swabs taken from three of the five chicken flocks tested positive. Finally, *L. innocua* was identified in 20% (20/100) of the food samples, with 11% (9/80) of the carcasses, 50% (5/10) of raw chicken fillets, and 60% (6/10) of the liver samples being positive.

Table 1. Occurrence of *Listeria innocua* along the poultry production chain

Source of Sample	Number of Samples Tested	Number of Positive Samples	% of Positive Samples
Chicken farm	**25** [1]	**7**	**28**
soiled litter	5	1	20
drinking water	5	0	0
poultry feed	5	1	20
farm wall	5	5	100
workers' hands	5 [2]	0	0
Slaughterhouse	**85** [2]	**9**	**11**
chicken cloaca	65	5	8
slaughterhouse wall	10	2	20
knife	5	0	0
table	5	2	40

Table 1. *Cont.*

Source of Sample	Number of Samples Tested	Number of Positive Samples	% of Positive Samples
Food product	**100** [2]	**20**	**20**
carcass	80	9	11
chicken fillet	10	5	50
chicken liver	10	6	60
total	**210**	**36**	**17**

[1] pooled samples; [2] individual samples.

3.2. Genomic Analysis

Whole-genome sequencing of the 36 *Listeria* isolates confirmed the MALDI–TOF MS results and clearly assigned them to the species *L. innocua* (94% to 100% of reads mapping to the *L. innocua* Clip11262 complete genome, NC_003212.1). All isolates belonged to the same *L. innocua*-specific MLST ST 530 (corresponding to CC ST530, Lineage: *L. innocua*). Sequencing coverage ranged between 43- and 132-fold (median 78). The 36 isolates showed 0 to 63 SNPs difference (median 41) and formed four distinct clusters of isolates differing by no more than 0, 3, 5, or 10 SNPs (Figure 1). Clusters were not restricted to a specific sampling site or sampling stage except for one cluster (no. 4), which included two isolates from carcasses (Figures 1 and 2).

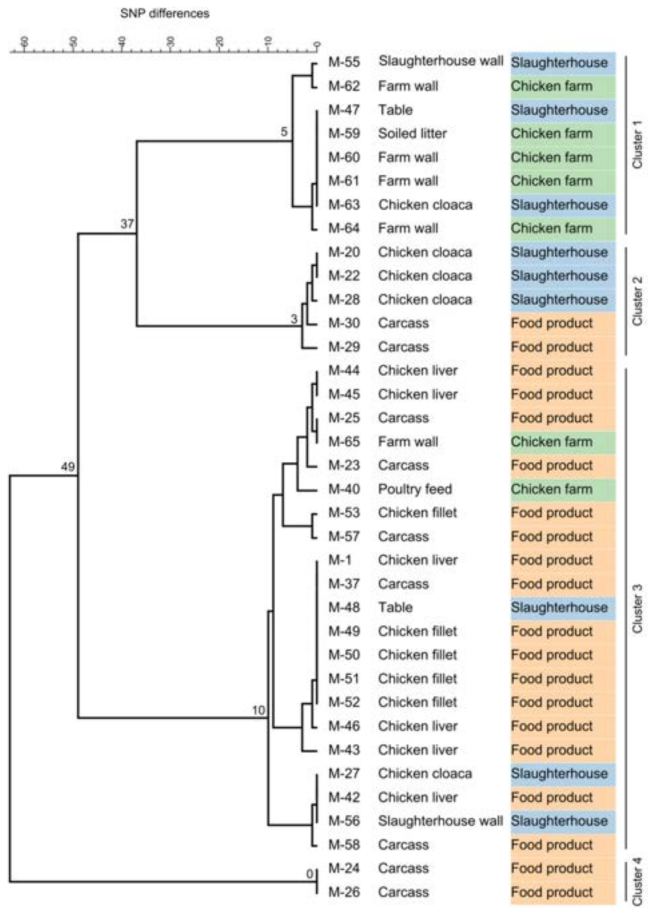


Figure 1. Complete linkage tree summarizing SNP analysis results from 36 *Listeria innocua* isolates. Node labels indicate the maximum SNP difference in the branch. Isolates fell into four distinct clusters differing by no more than 5, 3, 10, or 0 SNPs. Clusters were not restricted to a specific sampling site (first column) or sampling stage (second column) except for Cluster 4.

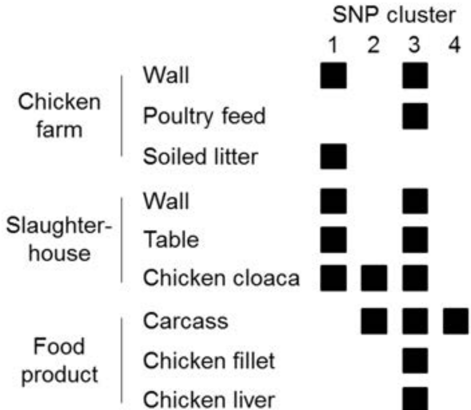


Figure 2. Visualization of contaminated sites and presumable transmission routes based on SNP clusters. Cross-contamination is likely to have happened between all production stages.

The *L. innocua* isolates under study harbored 12 to 13 *L. monocytogenes* virulence genes (Figure 3). However, the *Listeria* pathogenicity island LIPI-1 and internalins, especially *inlA*, which are genetic determinants for virulence in either atypical hemolytic *L. innocua* strains or in the pathogenic species *L. monocytogenes* [6], were only found in the sequence of the *L. monocytogenes* reference strain EGDe and were missing in all *L. innocua* isolates.

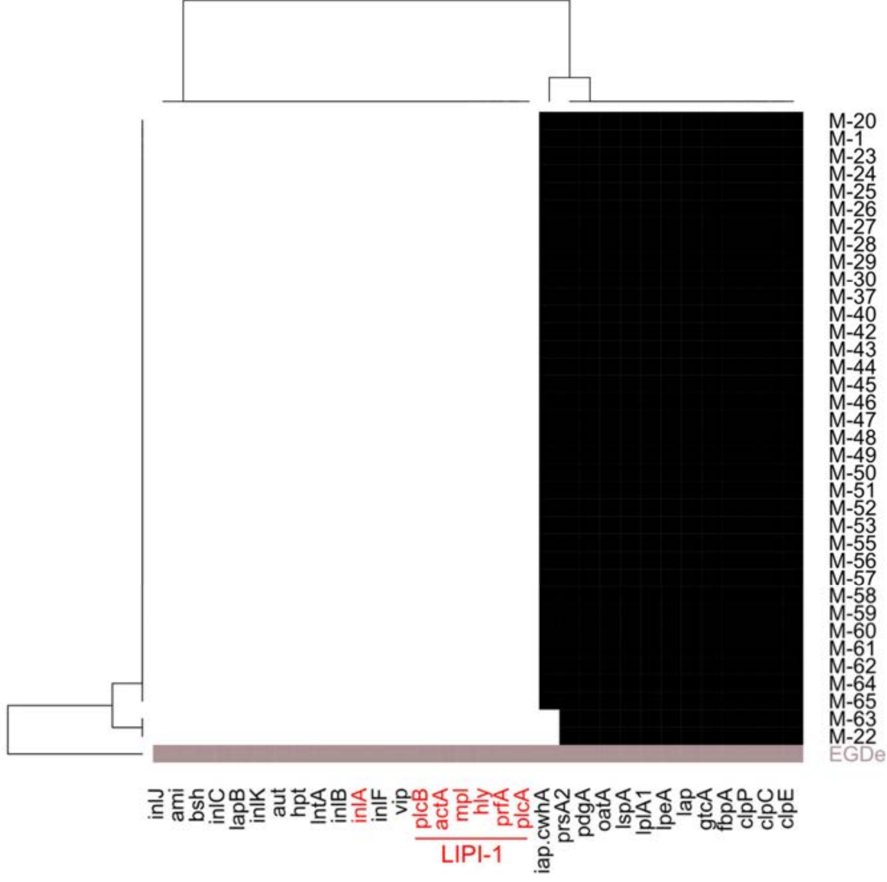


Figure 3. Heatmap of in silico detected virulence genes in the *Listeria (L.) innocua* study population (black: gene present; white: gene absent) compared to the *L. monocytogenes* reference strain EGDe (grey: gene present). None of the *L. innocua* study isolates contained virulence factors like *inlA* or the *Listeria* pathogenicity island 1 (LIPI-1; in red) that are found in atypical hemolytic *L. innocua* [6].

4. Discussion

At the three stages of the poultry production chain investigated (farm, slaughterhouse, food products), the prevalence of *L. innocua* ranged from 11% to 28%.

As we only looked into one specific farm-to-fork continuum (one farm, one slaughterhouse) in our study, the prevalences found at the various production stages cannot necessarily be generalized. However, our prevalence rates on farm level were in agreement with a previous study on a poultry farm in Egypt [26] where *L. innocua* was found in 20% (4/20) of samples from poultry feed but not in drinking water (0/20). However, *L. innocua* was more prevalent in soiled litter than in our study (35% (28/80) vs. 20% (1/5)). The observed difference could be explained either by the very different sample sizes or by variable efficiency of biosecurity practices on the farms. In our study, the overall prevalence of *L. innocua* on the farm was 28%. The prevalence reported for *Listeria* spp. (especially *L. monocytogenes* and *L. innocua*) on farms in various countries varied widely, ranging from 1.4% to 53% [9,27–31]. Accordingly, the detection rates for *Listeria* spp. in soiled litter (10%–53%), poultry feed (70%), drinking water (10%), soil (30%), and grass (6%–43%) were quite variable. In accordance with our results, *L. innocua* was the predominant species, representing up to 78% of the *Listeria* isolates.

The slaughterhouse prevalence of *L. innocua* in our study was 11%. The highest rate was found in swabs collected from tables (40%, 2/5), indicating that hygiene measures were not always properly applied to remove surface contamination. Lower contamination rates were found on the walls (20%, 2/10) and in cloacal swabs (8%, 5/65). A positive cloacal swab indicates intestinal colonization of the chicken. Hence, the positive rate of cloacal swabs is actually an indicator of the prevalence of carrier animals in the livestock population that form the basis for zoonotic entry of *Listeria* into the food chain. In our study, according to the microbiological results from cloacal swabs, three of the five chicken flocks tested were carriers of *L. innocua*. A lower detection rate (2%, 7/400) for *L. innocua* in cloacal swabs from laying hens was reported from Bavaria, Germany [28]. In the suburbs of Tokyo, *Listeria* spp. were found in 5% of 150 fresh fecal droppings collected on four chicken farms [32]. In a Danish study, *Listeria* spp. could not be isolated at all from 50 cloacal swabs taken at the abattoir of 71 broiler flocks [27]. Obviously, there are major differences in the *Listeria* prevalence rates in chicken livestock among countries.

In our dataset, the overall occurrence of *L. innocua* in raw poultry meat and chicken products was 20% (20/100). Higher prevalence rates for *L. innocua* in raw poultry meat were reported from Spain (66%), Turkey (58%), Italy (40%), Jordan (50%), and Egypt (31%) [26,33–36] while lower rates were reported from Morocco (14%) and Iran (19%) [37,38]. Contamination of raw animal products may occur after slaughter or during food processing. Major influencing factors are the primary prevalence of the pathogen in the livestock population, on the one hand, and hygiene measures such as surface disinfection during processing, on the other hand.

To get detailed insights into possible transmission routes along the poultry production chain, we analyzed single nucleotide polymorphisms of *L. innocua* isolates at various stages of production (Figure 1). Based on the fact that natural mutation rates in the genus *Listeria* are low, a very low number of SNP differences between two or more strains are commonly used as an indicator of their epidemiological relatedness [39,40]. The guiding assumption is that a small genetic difference between strains indicates a common origin. In our study, based on the high genetic relatedness of *L. innocua* isolates from different sampling sites from farm to final food product, transfer between all stages of the production chain appears very likely. Although no directional information can be extracted from the genetic data alone, accompanying metadata can be used to speculate about causal relationships between contaminated sites and transmission routes (Figure 2). For instance, zero SNP differences were identified between *L. innocua* isolates from chicken liver, fillet, carcass, and a slaughterhouse table (Figure 1; within Cluster 3). Presumably, the same bacterial strain has been transferred between all these sampling sites, indicating cross-contamination during the processing of the slaughtered chicken. Of positive note is the fact that all samples from knives and hands were negative for *Listeria* spp., indicating that hygienic measures are already successfully applied to a certain extent, although the

relatively low number of samples taken from both sites may qualify this statement. However, knowing that there is likely entry of *L. innocua* into the food processing plant from the farm, other surfaces should be under proper internal monitoring as well. For example, the same strain (zero SNPs difference) could be found on a table in the slaughterhouse and in cloacal swabs, as well as on farm walls and in the soiled litter (Figure 1; within Cluster 1). Consequently, bacterial contamination on the farm has reached the processing level, thereby posing a risk for further cross-contamination during processing steps performed on the table. Since contamination with *L. innocua* was already present on the farm, the food operator should reassess the sanitary measures applied and the way how chickens are introduced into the processing stage to prevent contamination from entering the production chain. If no reasonable measures are taken, *Listeria* can establish persistence in food processing plants, which may form the basis for repeated re-contamination [13,41,42].

Listeria innocua strains isolated in our study did not show any genetic virulence determinants needed for human or animal infection, as described for *L. monocytogenes* strains and rare hemolytic *L. innocua* strains [6]. In addition, all isolates were non-hemolytic on sheep blood agar. Therefore, it is highly unlikely that they would have been able to cause human infection. Furthermore, the pathogenic species *L. monocytogenes* was not detected. However, as previously mentioned, the Oxford medium used in our study does not allow for a distinction between colonies of different *Listeria* species [43]. A chromogenic agar, as described in ISO 11290:1:2017, would have been able to improve *L. monocytogenes* detection but it was not available in Egypt at the time of the study. Therefore, presumptive *Listeria* spp. colonies were randomly selected and *L. monocytogenes* isolates could have been missed. Additionally, *L. innocua* can produce a bacteriocin-like substance against *L. monocytogenes* and usually grows faster in enrichment broth, leading to an underestimation of the prevalence of *L. monocytogenes* [27,43].

While several studies from Egypt have reported frequent contamination of foodstuffs such as meat and dairy products with different *Listeria* spp. [26,44–47] such food has not yet been associated with documented outbreaks of listeriosis. A major reason for this is probably the lack of a surveillance system for human listeriosis in Egypt and hence underreporting of cases. As a result, the real public health burden caused by *Listeria* contamination throughout the food chain is very difficult to assess. Further close monitoring of slaughtering and company hygiene practices and their continuous adjustment and improvement will help to gain insights into the risks emerging from different food sources and will make an essential contribution to prevent listeriosis cases.

5. Conclusions

Our study did not reveal any *L. monocytogenes* contamination, but *L. innocua* existed throughout the entire chicken meat processing chain from stable to table. Given that *L. innocua* was not only isolated from environmental samples on the farm and in the slaughterhouse but also from cloacal swabs, chicken can be seen as a natural source of *L. innocua*. The presence of any non-pathogenic *Listeria* spp. like *L. innocua* in processing lines and foodstuffs is a good indicator for poor hygienic conditions and serves as an alarming sign for the need to implement appropriate hygiene practices. Through knowing and eliminating risk factors, contamination of poultry food products with the pathogenic species *L. monocytogenes* can be effectively prevented.

Author Contributions: Conceptualization, M.G., S.L., and S.A.D.; Data curation, S.L.; Formal analysis, M.G. and S.L.; Funding acquisition, S.K. and S.A.D.; Investigation, M.G., M.E.-A., A.Z., F.E.-G., and M.E.; Project administration, M.G.; Resources, S.K. and S.A.D.; Supervision, M.G. and S.A.D.; Visualization, S.L.; Writing—original draft, M.G. and S.L.; Writing—review and editing, M.E.-A., S.K., and S.A.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Federal Institute for Risk Assessment (grant no. 1322-668).

Acknowledgments: We would like to thank Anna-Louisa Hauffe and Marett Splett for their excellent technical assistance throughout the study.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Swaminathan, B.; Gerner-Smidt, P. The epidemiology of human listeriosis. *Microbes Infect.* **2007**, *9*, 1236–1243. [CrossRef]
2. Schuchat, A.; Swaminathan, B.; Broome, C.V. Epidemiology of human listeriosis. *Clin. Microbiol. Rev.* **1991**, *4*, 169–183. [CrossRef]
3. Buchrieser, C.; Rusniok, C.; Consortium, L.; Kunst, F.; Cossart, P.; Glaser, P. Comparison of the genome sequences of Listeria monocytogenes and Listeria innocua: Clues for evolution and pathogenicity. *FEMS Immunol. Med. Microbiol.* **2003**, *35*, 207–213. [CrossRef]
4. Johnson, J.; Jinneman, K.; Stelma, G.; Smith, B.; Lye, D.; Messer, J.; Ulaszek, J.; Evsen, L.; Gendel, S.; Bennett, R. Natural atypical Listeria innocua strains with Listeria monocytogenes pathogenicity island 1 genes. *Appl. Environ. Microbiol.* **2004**, *70*, 4256–4266. [CrossRef] [PubMed]
5. Moreno, L.Z.; Paixao, R.; Gobbi, D.D.; Raimundo, D.C.; Ferreira, T.P.; Hofer, E.; Matte, M.H.; Moreno, A.M. Characterization of atypical Listeria innocua isolated from swine slaughterhouses and meat markets. *Res. Microbiol.* **2012**, *163*, 268–271. [CrossRef] [PubMed]
6. Moura, A.; Disson, O.; Lavina, M.; Thouvenot, P.; Huang, L.; Leclercq, A.; Fredriksson-Ahomaa, M.; Eshwar, A.K.; Stephan, R.; Lecuit, M. Atypical hemolytic Listeria innocua are virulent, albeit less than Listeria monocytogenes. *Infect. Immun.* **2019**. [CrossRef] [PubMed]
7. Perrin, M.; Bemer, M.; Delamare, C. Fatal case of Listeria innocua bacteremia. *J. Clin. Microbiol.* **2003**, *41*, 5308–5309. [CrossRef] [PubMed]
8. Favaro, M.; Sarmati, L.; Sancesario, G.; Fontana, C. First case of Listeria innocua meningitis in a patient on steroids and eternecept. *JMM Case Rep.* **2014**, *1*, e003103. [CrossRef]
9. Milillo, S.R.; Friedly, E.C.; Saldivar, J.C.; Muthaiyan, A.; O'Bryan, C.; Crandall, P.G.; Johnson, M.G.; Ricke, S.C. A Review of the Ecology, Genomics, and Stress Response of Listeria innocua and Listeria monocytogenes. *Crit. Rev. Food Sci. Nutr.* **2012**, *52*, 712–725. [CrossRef]
10. Sauders, B.D.; Overdevest, J.; Fortes, E.; Windham, K.; Schukken, Y.; Lembo, A.; Wiedmann, M. Diversity of Listeria species in urban and natural environments. *Appl. Environ. Microbiol.* **2012**, *78*, 4420–4433. [CrossRef]
11. Mbata, T.I. Poultry meat pathogens and its control. *Internet J. Food Saf.* **2005**, *7*, 20–28.
12. Lakicevic, B.; Nastasijevic, I.; Raseta, M. Sources of Listeria monocytogenes contamination in retail establishments. *Procedia Food Sci.* **2015**, *5*, 160–163. [CrossRef]
13. Carpentier, B.; Cerf, O. Persistence of Listeria monocytogenes in food industry equipment and premises. *Int. J. Food Microbiol.* **2011**, *145*, 1–8. [CrossRef] [PubMed]
14. Orsi, R.H.; Wiedmann, M. Characteristics and distribution of Listeria spp., including Listeria species newly described since 2009. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 5273–5287. [CrossRef]
15. Genigeorgis, C.A.; Dutulescu, D.; Garayzabal, J.F. Prevalence of Listeria spp. in poultry meat at the supermarket and slaughterhouse level. *J. Food Prot.* **1989**, *52*, 618–624. [CrossRef]
16. Terzich, M.; Pope, M.J.; Cherry, T.E.; Hollinger, J. Survey of pathogens in poultry litter in the United States. *J. Appl. Poult. Res.* **2000**, *9*, 287–291. [CrossRef]
17. American Public Health Association. *Standard methods for the examination of water and wastewater*, 21st ed.; Am. Public Health Assoc. (APHA): Washington, DC, USA, 2005.
18. American Public Health Association. *Compendium of Methods for the Microbial Examination of Foods*; Vanderzant, C., Splittstoesser, D.F., Eds.; APHA: Washington, DC, USA, 1992.
19. McEvoy, J.; Nde, C.; Sherwood, J.; Logue, C. An evaluation of sampling methods for the detection of Escherichia coli and Salmonella on turkey carcasses. *J. Food Prot.* **2005**, *68*, 34–39. [CrossRef]
20. Hitchins, A.D.; Jinneman, K.; Chen, Y. BAM: Detection and Enumeration of Listeria monocytogenes. *Bacteriol. Anal. Man.* **2016**.
21. Freiwald, A.; Sauer, S. Phylogenetic classification and identification of bacteria by mass spectrometry. *Nat. Protoc.* **2009**, *4*, 732. [CrossRef]
22. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]

23. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef] [PubMed]
24. Seemann, T. Abricate—Mass Screening of Contigs for Antimicrobial and Virulence Genes. Available online: https://github.com/tseemann/abricate (accessed on 9 August 2019).
25. Liu, B.; Zheng, D.; Jin, Q.; Chen, L.; Yang, J. VFDB 2019: A comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **2019**, *47*, D687–D692. [CrossRef] [PubMed]
26. Dahshan, H.; Merwad, A.; Mohamed, T.S. Listeria species in broiler poultry farms: Potential public health hazards. *J. Microbiol. Biotechnol.* **2016**, *26*, 1551–1556. [CrossRef] [PubMed]
27. Petersen, L.; Madsen, M. Listeria spp. in broiler flocks: Recovery rates and species distribution investigated by conventional culture and the EiaFoss method. *Int. J. Food Microbiol.* **2000**, *58*, 113–116. [CrossRef]
28. Schwaiger, K.; Schmied, E.M.; Bauer, J. Comparative analysis on antibiotic resistance characteristics of Listeria spp. and Enterococcus spp. isolated from laying hens and eggs in conventional and organic keeping systems in Bavaria, Germany. *Zoonoses Public Health* **2010**, *57*, 171–180. [CrossRef]
29. Jones, D.; Anderson, K.; Guard, J. Prevalence of coliforms, Salmonella, Listeria, and Campylobacter associated with eggs and the environment of conventional cage and free-range egg production. *Poult. Sci.* **2012**, *91*, 1195–1202. [CrossRef]
30. Milillo, S.; Stout, J.; Hanning, I.; Clement, A.; Fortes, E.; Den Bakker, H.; Wiedmann, M.; Ricke, S. Listeria monocytogenes and hemolytic Listeria innocua in poultry. *Poult. Sci.* **2012**, *91*, 2158–2163. [CrossRef]
31. Dhama, K.; Verma, A.K.; Rajagunalan, S.; Kumar, A.; Tiwari, R.; Chakraborty, S.; Kumar, R. Listeria monocytogenes infection in poultry and its public health importance with special reference to food borne zoonoses. *Pak. J. Biol. Sci.* **2013**, *16*, 301–308. [CrossRef]
32. Iida, T.; Kanzaki, M.; Maruyama, T.; Inoue, S.; Kaneuchi, C. Prevalence of Listeria monocytogenes in intestinal contents of healthy animals in Japan. *J. Vet. Med. Sci.* **1991**, *53*, 873–875. [CrossRef]
33. Capita, R.; Alonso-Calleja, C.; Moreno, B.; Garcıa-Fernández, M.A.C. Occurrence of Listeria species in retail poultry meat and comparison of a cultural/immunoassay for their detection. *Int. J. Food Microbiol.* **2001**, *65*, 75–82. [CrossRef]
34. Yücel, N.; Cıtak, S.; Önder, M. Prevalence and antibiotic resistance of Listeria species in meat products in Ankara, Turkey. *Food Microbiol.* **2005**, *22*, 241–245. [CrossRef]
35. Pesavento, G.; Ducci, B.; Nieri, D.; Comodo, N.; Nostro, A.L. Prevalence and antibiotic susceptibility of Listeria spp. isolated from raw meat and retail foods. *Food Control.* **2010**, *21*, 708–713. [CrossRef]
36. Osaili, T.M.; Alaboudi, A.R.; Nesiar, E.A. Prevalence of Listeria spp. and antibiotic susceptibility of Listeria monocytogenes isolated from raw chicken and ready-to-eat chicken products in Jordan. *Food Control.* **2011**, *22*, 586–590. [CrossRef]
37. Amajoud, N.; Leclercq, A.; Soriano, J.M.; Bracq-Dieye, H.; El Maadoudi, M.; Senhaji, N.S.; Kounnoun, A.; Moura, A.; Lecuit, M.; Abrini, J. Prevalence of Listeria spp. and characterization of Listeria monocytogenes isolated from food products in Tetouan, Morocco. *Food Control.* **2018**, *84*, 436–441. [CrossRef]
38. Hamidiyan, N.; Salehi-Abargouei, A.; Rezaei, Z.; Dehghani-Tafti, R.; Akrami-Mohajeri, F. The prevalence of Listeria spp. food contamination in Iran: A systematic review and meta-analysis. *Food Res. Int.* **2018**, *107*, 437–450. [CrossRef] [PubMed]
39. Orsi, R.H.; Borowsky, M.L.; Lauer, P.; Young, S.K.; Nusbaum, C.; Galagan, J.E.; Birren, B.W.; Ivy, R.A.; Sun, Q.; Graves, L.M. Short-term genome evolution of Listeria monocytogenes in a non-controlled environment. *BMC Genom.* **2008**, *9*, 539. [CrossRef]
40. Pightling, A.W.; Pettengill, J.B.; Luo, Y.; Baugher, J.D.; Rand, H.; Strain, E. Interpreting Whole-Genome Sequence Analyses of Foodborne Bacteria for Regulatory Applications and Outbreak Investigations. *Front. Microbiol.* **2018**, *9*, 1482. [CrossRef]
41. Møretrø, T.; Langsrud, S. Listeria monocytogenes: Biofilm formation and persistence in food-processing environments. *Biofilms* **2004**, *1*, 107–121. [CrossRef]
42. Ferreira, V.; Wiedmann, M.; Teixeira, P.; Stasiewicz, M.J. Listeria monocytogenes persistence in food-associated environments: Epidemiology, strain characteristics, and implications for public health. *J. Food Prot.* **2014**, *77*, 150–170. [CrossRef]
43. Azevedo, I.; Regalo, M.; Mena, C.; Almeida, G.; Carneiro, L.; Teixeira, P.; Hogg, T.; Gibbs, P.A. Incidence of Listeria spp. in domestic refrigerators in Portugal. *Food Control.* **2005**, *16*, 121–124. [CrossRef]

44. Ahmed, A.M.; El-Atti, N.M.A. Existence of Listeria species in broiler carcasses with an attempt to control Listeria monocytogenes using trisodium phosphate. *Afr. J. Food Sci.* **2010**, *4*, 046–051.
45. El-Malek, A.M.A.; Ali, S.F.H.; Hassanein, R.; Mohamed, M.A.; Elsayh, K.I. Occurrence of Listeria species in meat, chicken products and human stools in Assiut city, Egypt with PCR use for rapid identification of Listeria monocytogenes. *Vet. World* **2010**, *3*, 353. [CrossRef]
46. Ismaiel, A.A.-R.; Ali, A.E.-S.; Enan, G. Incidence of Listeria in Egyptian meat and dairy samples. *Food Sci. Biotechnol.* **2014**, *23*, 179–185. [CrossRef]
47. Osman, K.M.; Samir, A.; Abo-Shama, U.H.; Mohamed, E.H.; Orabi, A.; Zolnikov, T. Determination of virulence and antibiotic resistance pattern of biofilm producing Listeria species isolated from retail raw milk. *BMC Microbiol.* **2016**, *16*, 263. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Emerging Microbes & Infections
2020, VOL. 9
https://doi.org/10.1080/22221751.2020.1784044

Backtracking and forward checking of human listeriosis clusters identified a multiclonal outbreak linked to *Listeria monocytogenes* in meat products of a single producer

Stefanie Lüth [a,b], Sven Halbedel [c], Bettina Rosner[d], Hendrik Wilking[d], Alexandra Holzer[d], Alice Roedel[a], Ralf Dieckmann[a], Szilvia Vincze[a], Rita Prager[c], Antje Flieger[c], Sascha Al Dahouk [a,e] and Sylvia Kleta[a]

[a]Department of Biological Safety, German Federal Institute for Risk Assessment, Berlin, Germany; [b]Department of Biology, Chemistry and Pharmacy, Freie Universität Berlin, Berlin, Germany; [c]Department of Infectious Diseases, Robert Koch-Institute, Wernigerode, Germany; [d]Department of Infectious Disease Epidemiology, Robert Koch-Institute, Berlin, Germany; [e]Department of Internal Medicine III, RWTH Aachen University Hospital, Aachen, Germany

ABSTRACT
Due to its high case fatality rate, foodborne listeriosis is considered a major public health concern worldwide. We describe one of the largest listeriosis outbreaks in Germany with 83 cases of invasive listeriosis between 2013 and 2018. As part of the outbreak investigation, we identified a highly diverse *Listeria monocytogenes* population at a single producer of ready-to-eat meat products. Strikingly, the extensive sampling after identification of a first match between a cluster of clinical isolates and a food isolate allowed for a linkage between this producer and a second, previously unmatched cluster of clinical isolates. Bacterial persistence in the processing plant and indications of cross-contamination events explained long-term contamination of food that led to the protracted outbreak. Based on screening for virulence factors, a pathogenic phenotype could not be ruled out for other strains circulating in the plant, suggesting that the outbreak could have been even larger. As most isolates were sensitive to common biocides used in the plant, hard to clean niches in the production line may have played a major role in the consolidation of the contamination. Our study demonstrates how important it is to search for the origin of infection when cases of illness have occurred (backtracking), but also clearly highlights that it is equally important to check whether a contamination at food or production level has caused disease (forward checking). Only through this two-sided control strategy, foodborne disease outbreaks such as listeriosis can be minimized, which could be a real improvement for public health.

ARTICLE HISTORY Received 21 April 2020; Revised 8 June 2020; Accepted 12 June 2020

KEYWORDS Listeriosis; outbreak; whole genome sequencing; source tracking; disinfection

Introduction

The bacterium *Listeria (L.) monocytogenes* is ubiquitous in nature and the causative agent of human listeriosis, a comparatively rare but potentially life-threatening foodborne disease [1]. The pathogen enters the food chain either through raw products or through contamination of food during processing [2]. Especially ready-to-eat (RTE) products pose a risk for infection [3]. Listeriosis may lead to a self-limiting gastrointestinal disease, to cerebral and bloodstream infections in predominantly immunocompromised patients or to fetal complications in pregnant women [4]. In Germany, the number of notified cases has been constantly increasing from 337 cases in 2011 to 770 cases in 2017, corresponding to an incidence increase from 0.4 to 0.9 cases per 100,000 population [5,6]. In 2018, the number of cases has fallen to 701 again (incidence: 0.8 cases per 100,000 population), with a case fatality rate of 5% [7]. Within the European Union, case fatality was even higher with 15.6% in the same year [8]. Thus, listeriosis represents a considerable burden to society which requires effective surveillance and prevention strategies by close collaboration between public health and food authorities. In Germany, the binational consultant laboratory for *L. monocytogenes* at the German Robert Koch-Institute and the Austrian Agency for Health and Food Safety collects *L. monocytogenes* strains isolated from clinical infections. During the last years, approximately 450 clinical isolates were collected annually, corresponding to approximately two thirds of all listeriosis cases notified in Germany. The National Reference Laboratory (NRL) for *L. monocytogenes*, hosted at the German Federal Institute for Risk Assessment, on the other hand, receives isolates sampled from food and food processing plants.

Molecular surveillance of *L. monocytogenes* using whole genome sequencing (WGS), combined with

CONTACT Stefanie Lüth stefanie.lueth@bfr.bund.de
Supplemental data for this article can be accessed at https://doi.org/10.1080/22221751.2020.1784044

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

epidemiological evidence, has greatly facilitated listeriosis outbreak clarification [9–13]. In addition, WGS analysis enables detailed insights into industrial hygiene and forms the basis for in-depth root cause analysis [14].

In our study, we analyzed the diverse *L. monocytogenes* population of a German food processing plant (isolates from food and environment) that was linked to a large long-lasting listeriosis outbreak consisting of two distinct clusters. In addition to backtracking and forward checking, we estimated the virulence potential of the strains circulating in the production facility. Last but not least, we addressed the question of how the contamination has been persisting for years despite periodic hygiene measures.

Materials and methods

Bacterial cultivation

L. monocytogenes strains were routinely cultured in brain heart infusion (BHI) broth, on BHI agar plates or on sheep blood agar plates at 37°C overnight.

Pulsed-field gel electrophoresis

Pulsed-field gel electrophoresis (PFGE) was performed according to the PulseNet protocol (https://www.cdc.gov/pulsenet/pdf/listeria-pfge-protocol-92 508c.pdf). Restriction patterns were analyzed with BioNumerics, version 7.1 (Applied Maths, Sint-Martens-Latem, Belgium).

Whole genome sequencing

Genomic DNA was extracted using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich, St. Louis, MO, United States; clinical isolates) or the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany; non-clinical isolates) following the PulseNet protocol for gram-positive bacteria (https://www.cdc.gov/pulsenet/pdf/pnl32-miseq-nextera-xt.pdf). Extracted DNA was quantified on a fluorescence microplate reader using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA, United States; clinical isolates) or using the Qubit dsDNA BR Assay Kit with a Qubit 2.0 fluorometer (Invitrogen, Carlsbad, CA, United States; non-clinical isolates). Sequencing libraries from genomic DNA were prepared with the Nextera XT DNA Library Prep Kit (Illumina, San Diego, CA, United States). Sequencing was performed on the MiSeq sequencer in paired-end mode with 2 × 300 bp reads or single-end mode with 1 × 300 bp or 1 × 150 bp reads or on an Illumina HiSeq 1500 sequencer generating 250 bp paired-end reads in a dual flow cell run.

Sequencing data analysis

Data preparation

Sequencing reads were trimmed with Trimmomatic version 0.36 at default parameters [15]. Trimmed reads were either directly used for single nucleotide polymorphism (SNP)-mapping or *de novo* assembled with SPAdes version 3.11.1 [16].

Multi locus sequence typing and molecular serogrouping

Multi locus sequence types (MLST STs) as well as corresponding MLST clonal complexes (CCs) and PCR-serogroups were determined from *de novo* assemblies according to the seven house-keeping gene MLST scheme and the PCR-serogrouping scheme, respectively, available at http://bigsdb.pasteur.fr/listeria.

Core genome MLST

Core genome MLST (cgMLST) was performed based on assembled genomes in the software Ridom SeqSphere+ (Münster, Germany) with the integrated 1701 genes cgMLST scheme [17]. CgMLST allele coverage of at least 98% was set as quality threshold. CgMLST allelic profiles were imported into BioNumerics version 7.6 to perform single linkage clustering. Isolates with a maximum of ten allele differences from each other were assigned to the same cluster [17]. Trees were visualized and annotated in iTOL version 4 [18].

Single nucleotide polymorphism analysis

Trimmed reads were mapped against the sequence of the *L. monocytogenes* strain EGDe (NC_003210.1) using Snippy version 4.0 at default settings [19].

In silico screening for antimicrobial resistance and virulence genes

Antimicrobial resistance (AMR) and virulence genes were identified from assembled genomes with ABRicate version 0.8 [20] using the databases ncbi (AMR, 4528 sequences) and vfdb (virulence, 2597 sequences) [21,22], last updated 9 July 2019. To reduce assembly bias, gene coverages were summed up when a gene was split across multiple contigs (visualization with Geneious Prime 2020.0.3). A cutoff of at least 75% gene coverage in total was applied for gene presence.

Data availability

The sequence data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accessions listed in Supplementary Table 1.

Biocide susceptibility testing

Except for benzalkonium chloride (BAC), only active substances of the cleaning agents and disinfectants used in the food processing plant under investigation

were tested. BAC was included due to genetically encoded BAC tolerance mechanisms found in the study isolates. Minimum inhibitory concentrations (MICs) of BAC (0.3 to 20 mg/L), sodium hypochlorite (62.5 to 8000 mg/L), peracetic acid (22.3 to 2850 mg/L), hydrogen peroxide (7.8 to 999 mg/L) and phosphoric acid (444.4 to 28440 mg/L) were determined in triplicates as previously described [23]. Biocide susceptibility of nine cgMLST cluster 1 outbreak isolates and two cluster 3 non-outbreak isolates were analyzed exemplarily. Two more isolates were selected for phenotypic testing of BAC tolerance because of their AMR genotype (cluster 9 and 15). The MIC breakpoint ≥4 mg/L was used to classify isolates as BAC tolerant [24,25]. Minimal in-use concentrations of the other biocides were calculated based on manufacturer specifications about biocide concentrations in stock solutions and application concentrations specified in the cleaning and disinfection plan of the processing plant.

Statistical analysis

Statistical analysis was performed in IBM SPSS Statistics version 21 (IBM, Armonk, NY, United States). Case–control-study was analyzed with Stata 15.0 (StataCorp LLC, TX, United States). In general, analyses with p-values lower than 0.05 were considered as statistically significant. In chi-squared tests, p-values were adjusted using Bonferroni correction. Strength and direction of a relationship between variables were measured by Spearman correlation.

Epidemiological analysis

Case definition

Outbreak cases were defined as listeriosis patients reported to public health authorities with disease onset in 2013 or later, and isolation of *L. monocytogenes* from normally sterile body fluids revealing either characteristic PFGE profiles (typing method applied before 2015) or belonging to the cgMLST clusters 1 and 2 (typing method applied after 2015, Figure 2). To ensure compatibility of molecular typing results, selected strains with PFGE profiles typical for the outbreak were sequenced retrospectively. In case of affiliation to cgMLST cluster 1 or 2, other isolates showing the same PFGE profile were considered as associated as well.

Case–control study

In order to support molecular typing data with epidemiological evidence, a case–control study was conducted in 2017. Recent outbreak cases were asked about their diet in the two weeks prior to disease onset, while earlier cases, from 2016, were asked about general consumption habits during the time of infection (total $n = 8$). Healthy controls from Germany with a similar age and sex distribution as outbreak cases were interviewed by a social research institute using random digit dialling ($n = 32$). Our hypothesis was that the outbreak was caused by the consumption of plastic packaged RTE meatballs. This assumption was compared to consumption of two foodstuffs classically considered at risk for *L. monocytogenes* contamination: plastic-packaged sliced cheese and plastic-packaged smoked fish from the supermarket [1].

Results

Epidemiological outbreak description

The German surveillance system for *L. monocytogenes* identified an outbreak with 83 invasive listeriosis cases between 2013 and 2018. It consisted of two distinct cgMLST clusters (cluster 1 and cluster 2) which could be traced to the same food processing plant. Because of the common source, both clusters were treated as one outbreak (Figure 1). Cluster 1 had been communicated to European member states on 29/09/2016 via the Epidemic Intelligence Information System (EPIS UI-376), but none of the other participating countries reported cases.

Cluster 1 comprised 72 cases between 2013 and 2018, spread over 12 out of 16 German federal states. Median age of cases in this cluster was 69 years (range 0–96 years) and more male ($n = 44$) than female ($n = 28$) persons were affected. The cluster included six pregnancy-associated cases (mother and child were treated as two separate cases). Five (non-pregnancy-associated) cases died. For three of them, listeriosis was reported as the major cause. Cluster 2 was smaller, included 11 cases between 2015 and 2017 and was reported only from six German federal states. Median age of cases in this cluster was 70 years (range 16–86 years) and again, more male ($n = 8$) than female ($n = 3$) persons were affected. In this cluster, there were no pregnancy-associated cases and no deaths. Cluster 1 reached a peak in 2016 with 26 cases, whereas most cases of cluster 2 appeared in 2017 ($n = 7$) (Figure 1).

From 2013 through 2015, most outbreak cases were male (26 out of 29 cases, 90%). After that the ratio between female and male was almost balanced (26 male versus 28 female cases).

Clinical L. monocytogenes *isolates linked to the outbreak*

Since clinical *L. monocytogenes* isolates could not always be assigned to a notified case [7], the number of isolates was higher than the number of cases in the outbreak. Some of the isolates were exclusively typed with PFGE, but only sequenced clinical isolates ($n = 77$) were included in our study, with 65 isolates in cgMLST cluster 1, and 12 isolates in cluster 2.

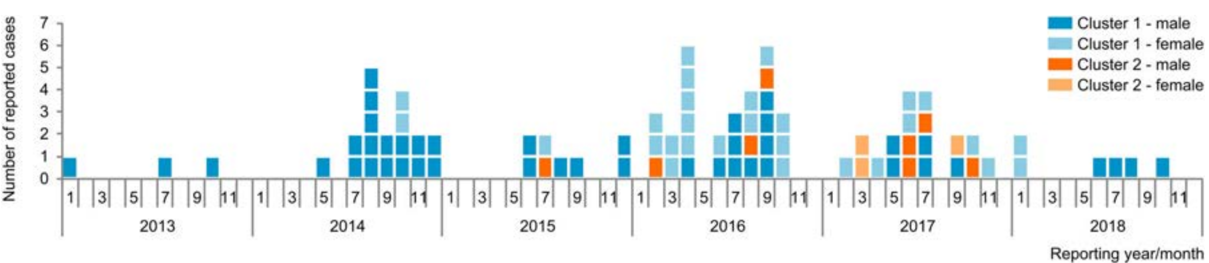


Figure 1. Epidemiological curve of the listeriosis outbreak under investigation. The outbreak comprised 83 cases between 2013 and 2018 and consisted of two distinct core genome MLST clusters (cluster 1, blue, and cluster 2, orange). During the outbreak, a shift from mainly male cases (dark shading) to a balanced ratio between male and female cases (light shading) occurred.

Source identification and sampling at a food producer

During retrospective investigations, an isolate sampled in 2016 from RTE meatballs was found to match the clinical isolates of cgMLST cluster 1 (one allele difference). After checking the NRL database, a second isolate from meatballs of the same producer, sampled in 2014, could be assigned to the same cluster. Both isolates were collected in official controls while the respective product had already been put on the market. Based on these findings, extensive sampling at the food producer was initiated, and isolates from the production facility subsequently matched a second cluster of clinical isolates, namely cluster 2 (Figure 2).

Altogether, 235 non-clinical isolates originating from a single producer were included in our study. Except for the two initial food isolates from 2014 and 2016, isolates were collected in 2017 and 2018. A total of 210 originated from the food processing environment and 25 from food products sampled either at the retail level ($n = 3$) or at the producer level ($n = 22$). Swabs were taken from conveyor belts, pulleys, freezers, accompanying parts like condensate lines or cable ducts, and from gullies. Food isolates originated from RTE meat products such as meatballs or burger patties made from pork, poultry or unknown type of meat and chicken nuggets. The food samples were contaminated below 100 CFU/g [3], except for the one sample from 2014 which contained 3×10^4 CFU/g. Of the 235 non-clinical isolates, 216 (92%) were sampled within self-controls by the producer (21 from food, 195 from food processing environment) and the rest in official controls.

Case–control study

In order to underpin WGS-based typing results with epidemiological findings, 8 cases and 32 healthy controls were interviewed in 2016 and 2017 concerning their consumption of RTE meatballs, sliced cheese or

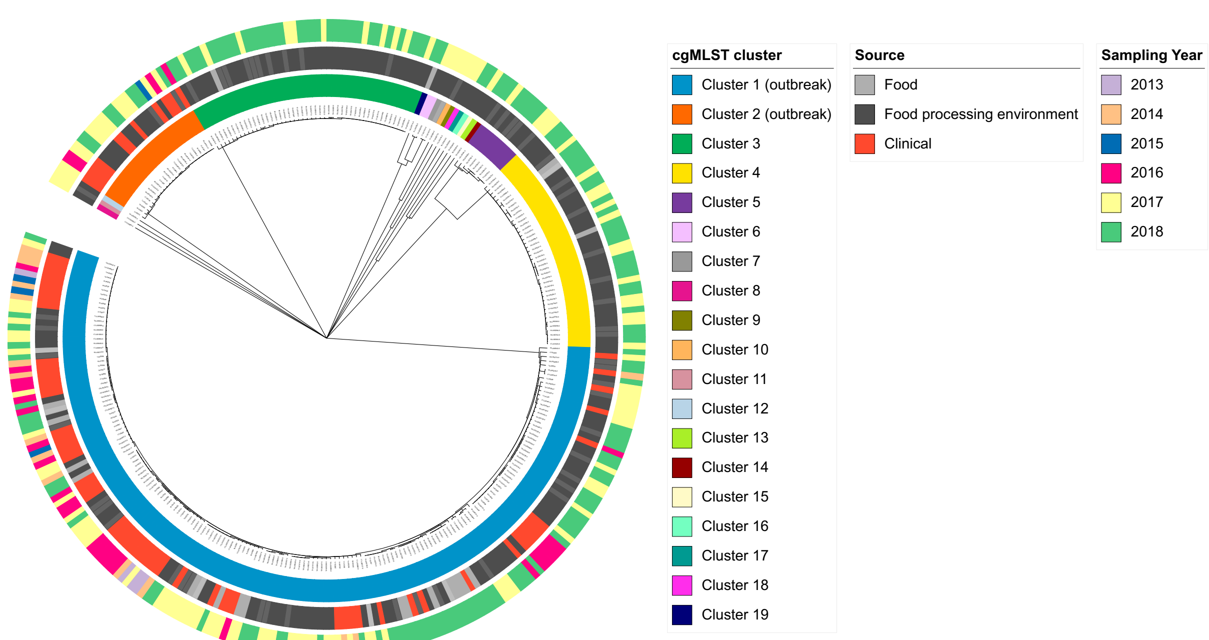


Figure 2. Core genome (cg) MLST-based single linkage clustering of 312 *Listeria monocytogenes* isolates from clinical cases and a single food producer. Colored rings indicate clustering results and metadata of isolates. From inside to outside: cgMLST cluster number, source and sampling year.

smoked fish (Table 1). One case could not recall the consumption of plastic packaged RTE meatballs, but exposure to the unpacked product while eating out was not entirely excluded. A second case answered the question on consumption of smoked fish with "I don't know". Both cases were classified as not exposed to the respective food category in our analysis. Altogether, six out of eight patients (75%) remembered eating plastic packaged RTE meatballs from retail whereas only two of the 32 controls (6%) did. This corresponds to an odds ratio of 102.4 at a p-value of 0.001, indicating a strong association between listeriosis outbreak cases and consumption of plastic-packaged RTE meatballs. In contrast to that, odds ratios between cases and controls for consumption of sliced cheese and smoked fish were 4.4 and 1.3, respectively, and not statistically significant.

Molecular typing and cluster analysis of strains

The 312 isolates included in our study were assigned to 11 different MLST CCs spanning four serogroups (IIb, IIa, IVb and IIc, with decreasing prevalence). The majority (>95%) of isolates fell into CC5 ($n = 176$), CC121 ($n = 50$), CC31 ($n = 46$) and CC7 ($n = 24$). Clinical isolates were either CC5 ($n = 65$, serogroup IIb) or CC7 ($n = 12$, serogroup IIa).

The isolates fell into 7 cgMLST clusters and 12 singletons (containing only one non-clinical isolate) (Figure 2). The outbreak-associated cgMLST clusters 1 and 2 contained 176 and 24 isolates, respectively. Cluster 1 was composed of 65 clinical isolates, 20 isolates from food and 91 isolates from food processing environment. All isolates in this cluster were closely related with an overall allelic distance between 0 and 18 (median 7). The close genetic relatedness of isolates in this cluster could be confirmed by SNP analysis (overall SNP distance 0–60, median 10). Cluster 2 contained only 24 isolates, 12 of clinical origin and 12 from food processing environment. Allelic differences ranged between 0 and 8 (median 2). Corresponding SNP distances were between 0 and 8 with a median of 3. With an allelic distance of 1633, the genetic difference between cluster 1 and cluster 2 was large. The two outbreak clusters did not show sub-clustering according to the source of isolates or the time of sampling (Figure 2). For the other cgMLST clusters of the non-clinical isolates, no match to clinical isolates from Germany could be found.

Four cgMLST clusters (cluster 1, 3, 4 and 19) contained isolates from RTE food products and three of those (cluster 1, 3 and 4) also isolates from food processing environment. In the remaining 15 cgMLST clusters, non-clinical isolates originated only from the food processing environment. Four cgMLST clusters (cluster 2–5) contained isolates which have been sampled over a period of 9–12 months, in 2017 and 2018. Cluster 1 has been detected over four years.

Table 1. Univariate analysis of factors associated with listeriosis, cgMLST cluster 1, Germany 2017.

Food product (plastic packaged, RTE)	Cases exposed no./total no. (%)	Controls exposed no./total no. (%)	Logistic regression, adjusted for age and sex	
			Odds ratio (95% CI[a])	p-value
Meatballs	6/8 (75)[b]	2/32 (6)	102.4 (7.0–1509.6)	0.001
Sliced cheese	6/8 (75)	17/32 (53)	4.4 (0.5–42.0)	0.200
Smoked fish	1/8 (13)[b]	3/32 (9)	1.3 (0.1–14.7)	0.851

[a]CI = confidence interval; [b]A case in this category has been classified as not exposed due to unclear information.

Virulence genes

A total of 40 different virulence genes could be identified in the 312 isolates (Figure 3). A single isolate contained 29–39 virulence genes (median 32). Virulence gene counts were significantly different between outbreak and non-outbreak clusters (Mann–Whitney-U test, $p < 0.001$). A set of 26 virulence genes was found in all isolates.

Genes positively associated with the two outbreak clusters were *aut, inlF, lapB* and *vip* (correlation 0.2, 0.58, 0.56, 0.34, respectively; 1-tailed $p < 0.01$). LIPI-3 (*llsAGHXBYDP*, [26]) was only detected in non-outbreak clusters (cluster 6, 7 and 17–19) with a correlation of 0.2 (1-tailed $p < 0.01$).

Clusters 9, 10, and 12–16 had the same virulence factor composition as outbreak cluster 1. The virulence factors of cluster 8 and cluster 11 isolates were identical to outbreak cluster 2 isolates lacking the *vip* gene.

Antimicrobial susceptibility

In all 312 isolates, the AMR genes *lin*, coding for the lincomycin resistance ABC-F type ribosomal protection protein, and *fosX*, coding for the fosfomycin resistance thiol transferase FosX, were identified (Figure 4).

In 48 isolates of three different cgMLST clusters (3, 9 and 15), none of them associated with clinical cases, the BAC tolerance genes *bcrB* and *bcrC* were found [27] (Figure 4). At least one representative isolate per *bcrB-bcrC* containing cgMLST cluster was phenotypically tested and confirmed as BAC tolerant (MIC values ranging from 5 to 10 mg/L). Representative isolates of the outbreak cluster 1, lacking *bcrB* and *bcrC*, showed a BAC MIC value of 2.5 mg/L and were classified as susceptible. MIC values of the other biocides tested did not differ between isolates. MIC of sodium hypochlorite was 500 mg/L, of peracetic acid 356.3 mg/L, of hydrogen peroxide 249.8 mg/L and of phosphoric acid 3555 mg/L. All MIC values were below in-use concentrations.

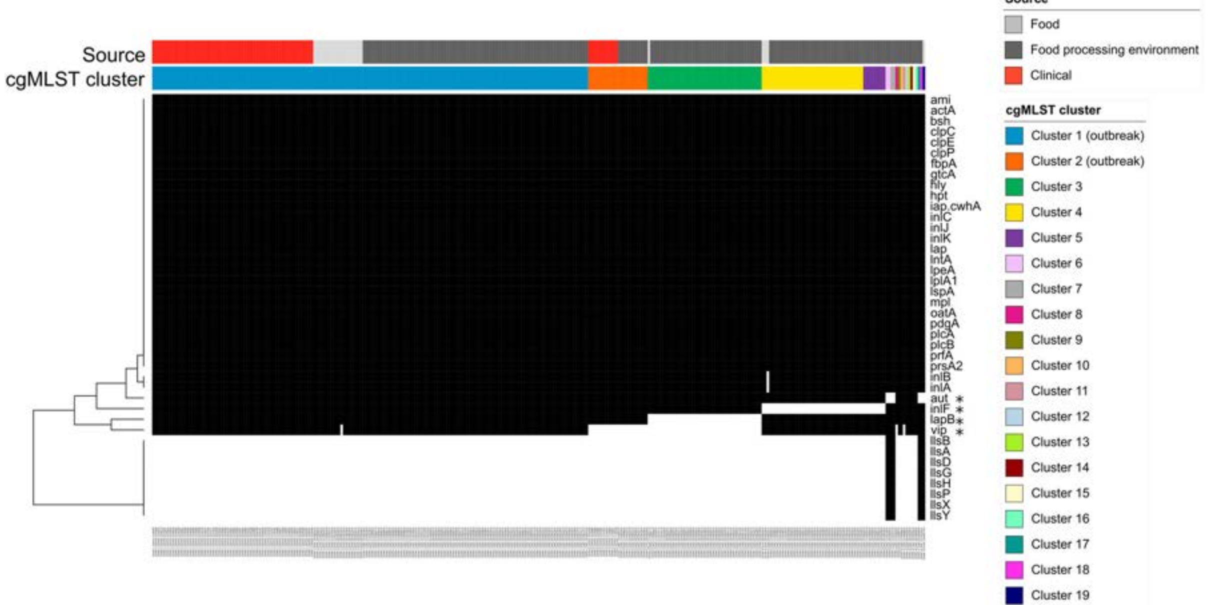


Figure 3. Heatmap of *in silico* detected virulence genes. Black: gene present; white: gene absent. Isolates are sorted by source and by cgMLST cluster number. A set of 26 virulence genes was present in all 312 study isolates. Genes positively associated with the outbreak clusters are marked by an asterisk. CgMLST clusters 9, 10, and 12–16 had the same virulence factor composition as outbreak cluster 1. Isolates in the clusters 8 and 11 were lacking the *vip* gene and were hence identical to outbreak cluster 2.

Discussion

With 83 confirmed cases, the outbreak described in our study represents one of the largest listeriosis outbreaks identified in Germany since cgMLST has been implemented for molecular surveillance [28]. From 2017 to 2018, for the first time after the introduction of mandatory reporting of human listeriosis, the number of notified cases in Germany declined [5,7]. Using the potential of WGS to resolve and to stop large outbreaks [13,29] might have already made a decisive contribution towards reducing the burden of listeriosis in Germany. More males than females were affected within the outbreak described here. This is in line with the general trend of a significantly higher incidence rate of listeriosis among men in Germany [5,7,30–32]. Interestingly, there was a gender shift of cases during the outbreak. General changes in consumption habits as well as the introduction of novel food products could have played a role. Both factors are likely to influence the ingested dose and hence the relevant dose–response relationship for a consumer group of interest – in this case women [33]. The age distribution was typical for a listeriosis outbreak, representing the population groups at risk: elderly people (older than 69 years of age), pregnant women (women of fertile age), and newborns.

In the first outbreak cluster, the entire hypothetical transmission chain could be traced back from clinical cases to food product to food processing environment due to the high genetic relatedness of isolates. Indeed, the cgMLST-based suspicion that RTE meatballs were the causative food vehicle for listeriosis infections could be epidemiologically confirmed by a case–control study. In the second outbreak cluster, the food product was missing in the transmission chain. However, as soon as the contamination in the food processing environment was detected and eliminated, further contamination of food products and hence transmission to consumers was effectively prevented. Consequently, forward checking, along with backtracking, and the

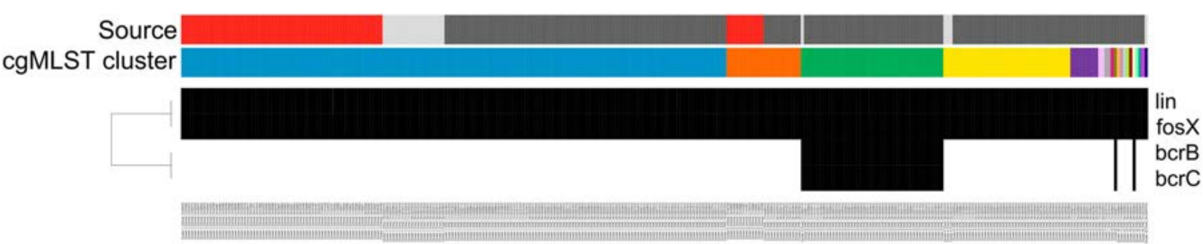


Figure 4. Heatmap of *in silico* detected antimicrobial resistance genes. Black: gene present; white: gene absent. Isolates are sorted by source and by cgMLST cluster number (for legend see Figure 3). All 312 study isolates contained the *lin* gene (lincomycin resistance) and the *fosX* gene (fosfomycin resistance). Only 48 isolates of three different cgMLST clusters (cluster 3, 9 and 15) additionally contained the benzalkonium chloride tolerance genes *bcrB* and *bcrC*.

resulting recall of products from the market as well as stopping of the production, finally terminated the outbreak.

Because of the great variety of *L. monocytogenes* strains at the producer, the question arose as to why only two of them have caused human infections although probably more of them had reached consumers. One possible explanation could be that strains differ in pathogenicity. In line with a recent French study showing associations of certain MLST CCs with either infection or food [34], CC121 was the second most common MLST CC in our non-clinical isolates. The two outbreak clusters, however, belonged to CC5 and CC7, which were classified as intermediate and even rarely responsible for clinical cases, respectively. Analysis of clinical *L. monocytogenes* isolates from Germany [28] had confirmed that CC5 and CC7 have not frequently been associated with listeriosis cases so far. To gain a deeper insight into strain pathogenicity, we screened the genomes of all isolates for known virulence factors [21]. The genes *aut*, *vip*, *inlF* and *lapB*, which are all critical for host cell entry [35–37], were found to be positively associated with outbreak isolates and thus appeared to be involved in a strain's ability to infect humans. However, those virulence genes were also found in non-outbreak clusters including infection-associated MLST CCs, such as CC1 and CC6 [34]. Additionally, LIPI-3 (*llsAGHXBYDP*, [26]) was detected in non-outbreak clusters, a genetic island which is linked to increased invasiveness of *L. monocytogenes* [38]. Last but not least, pathogenicity is a multifactorial process, and cannot merely be derived from the presence or absence of virulence genes [39]. In summary, the variety of potentially pathogenic strains circulating at the producer, as already described in other food processing plants [40], clearly shows that selectively removing the source of contamination for one specific outbreak cluster is neither sufficient nor sustainable. Instead, the entire *L. monocytogenes* population in a food processing plant must be controlled and eliminated to not only stop present outbreaks but to also prevent future ones.

The high diversity of the *L. monocytogenes* population found in the food processing plant is not necessarily due to an extraordinary extent of contamination, but very likely results from extensive sampling ("Who seeks shall find"). Nevertheless, it highlights the problem of recurring *L. monocytogenes* contamination at the producer and insufficiently established hygiene measures. Strains of at least five of the *L. monocytogenes* clusters were persistent in the food processing plant and have been detected for nine months to four years. This long period of time may explain the comparatively large allele and SNP distance [12,41–43] between the epidemiologically linked isolates in our study. Furthermore, in three of those clusters, strains were found in both, food and food processing environment, verifying cross-contamination events. Detection of twelve cgMLST singletons showed a snapshot which does not necessarily exclude persistence of these strains. This observation may also provide evidence for periodic entry events of *L. monocytogenes* into the production line, for example via raw meat from various suppliers.

To gain a better understanding of the reason for long-lasting persistence of *L. monocytogenes* strains in the food processing plant, we screened for AMR genes. Inadequate disinfection practices may expose bacteria to sub-lethal biocide concentrations and thus select for tolerant strains which may then persist in niches [44]. As *L. monocytogenes* is known to be naturally resistant to lincomycin and fosfomycin [45–47], full length detection of these two AMR genes in all isolates illustrated the effectiveness of our in silico screening method. Forty-eight out of 235 *L. monocytogenes* isolates (20%) from food or food processing environment, found in cgMLST clusters 3, 9 and 15, carried the BAC tolerance genes *bcrB* and *bcrC*. These figures are in agreement with a recent study that found a prevalence rate of BAC tolerance of 16% in 93 isolates from German food processing environments collected from 2008 through 2016 [23]. However, since November 2016, BAC has been listed as an unapproved disinfectant and preservative in the EU (implementation decision 2016/1950), and most of our study population and importantly, all isolates in outbreak clusters, were susceptible. All tested isolates exhibited MIC values lower than the in-use concentrations of biocides in the cleaning agents and disinfectants applied in the high care area of the food processing plant. We therefore assume that all substances were suitable for cleaning and disinfection if hygiene measures met the guidelines. Hence, retention in hard-to-reach and consequently hard-to-clean niches may have played a major role in the establishment of persistence. Indeed, *L. monocytogenes* contamination was found in such niches along production lines and included surfaces in high risk areas where previously heat-treated meat products were chilled before packaging. The knowledge gained about hotspots of contamination should help to improve cleaning regimes including periodic disassembly of production lines and/or to redesign the manufacturing equipment so that hard-to-clean areas are minimized.

As observed in other studies [48,49], neither adjustment of the hygiene management concepts, nor infrastructural changes were successful to get the *L. monocytogenes* contamination at the producer under control. As a result, the entire processing plant was shut down in autumn 2018. With the last clinical case in October 2018, the outbreak was considered as terminated.

In the beginning of 2019, however, an isolate from a second producer located in another German federal

state matched outbreak cluster 1 of our study according to cgMLST. Throughout the year, 16 isolates from this second producer, 3 from RTE meat products and 13 from the food processing environment, were found to be highly genetically related to the isolates of the outbreak. Both producers did not have a direct supply relationship and, in so far as this is known, neither equipment nor staff has been transferred between them. However, they shared some of their suppliers, supporting the hypothesis that the outbreak strain has been introduced via contaminated raw animal products in both plants. On the one hand, this shows that the search for an outbreak source does not necessarily end at the level of final food processing but needs to be extended to the level of slaughterhouses and cutting plants to really address the root of the problem. Ultimately, this means that not only sharing of sequencing data is needed, but that integration of information on commodity chains into a common database would be important as well. On the other hand, it also shows the need for epidemiological investigations in addition to molecular surveillance. Despite the high genetic relatedness of the new isolates to the former outbreak, no further clinical cases related to the outbreak have been reported up to the date of publication.

Conclusion

In order to prevent listeriosis cases before they occur, we should not only carry out outbreak detection, but also set a focus on expansion of the data set available for WGS-matching. One important approach for that purpose would be to intensify regular monitoring in the companies. Preferably, this is largely implemented in the form of company's own checks, so that contaminated food products are detected early enough and do not enter the market. Overall, backtracking and forward checking along the entire food chain must go hand in hand to protect the public from zoonotic pathogens. These terms are inspired by the area of artificial intelligence, where forward checking is used as a look ahead strategy during backtracking [50]. A common database of molecular typing results may solve this problem automatically, since it enables real-time comparison in both directions. Only through this two-sided control strategy, foodborne disease cases can be prevented.

Acknowledgment

We thank the food control and veterinary offices and the public health authorities of the federal states for the productive collaboration during the outbreak. We are grateful for the excellent technical assistance of Anna-Louisa Hauffe, Marett Splett, Ute Strutz and Thomas Fischer throughout the study. We also thank Raskit Lachmann for her feedback on the article.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by a grant of the Federal Ministry of Health (GE 2016 03 26) in the framework of the German Research Platform for Zoonoses, by the German Federal Institute for Risk Assessment (1322-668), and by the Robert Koch-Institute under the grant "Intensified Molecular Surveillance Inititative".

ORCID

Stefanie Lüth http://orcid.org/0000-0002-6321-3869
Sven Halbedel http://orcid.org/0000-0002-5575-8973
Sascha Al Dahouk http://orcid.org/0000-0003-3835-0818

References

[1] EFSA, ECDC. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. EFSA J. 2018;16(12):e05500.
[2] Swaminathan B, Gerner-Smidt P. The epidemiology of human listeriosis. Microbes Infect. 2007;9(10):1236–1243.
[3] COMMISSION REGULATION (EC) No 2073/2005 of 15 November 2005 on microbiological criteria for foodstuffs; 2005.
[4] Allerberger F, Wagner M. Listeriosis: a resurgent foodborne infection. Clin Microbiol Infect. 2010;16(1):16–23.
[5] Robert Koch-Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2017; 2018.
[6] Robert Koch-Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2011; 2012.
[7] Robert Koch-Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2018; 2019.
[8] ECDC. The European Union one health 2018 zoonoses report; 2019.
[9] Kwong JC, Mercoulia K, Tomita T, et al. Prospective whole-genome sequencing enhances national surveillance of Listeria monocytogenes. J Clin Microbiol. 2016;54(2):333–342.
[10] Moura A, Criscuolo A, Pouseele H, et al. Whole genome-based population biology and epidemiological surveillance of Listeria monocytogenes. Nat Microbiol. 2017;2:16185.
[11] Jackson BR, Tarr C, Strain E, et al. Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. Clin Infect Dis. 2016;63(3):380–386.
[12] Pietzka A, Allerberger F, Murer A, et al. Whole genome sequencing based surveillance of *L. monocytogenes* for early detection and investigations of listeriosis outbreaks [review]. Front Public Health. 2019;7(139).
[13] Kleta S, Hammerl JA, Dieckmann R, et al. Molecular tracing to find source of protracted invasive listeriosis outbreak, Southern Germany, 2012–2016. Emerging Infectious Dis J. 2017;23(10):1680.
[14] Gwida M, Lüth S, El-Ashker M, et al. Contamination pathways can be traced along the poultry processing chain by whole genome sequencing of Listeria innocua. Microorganisms. 2020;8(3):414.

[15] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–2120.
[16] Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19 (5):455–477.
[17] Ruppitsch W, Pietzka A, Prior K, et al. Defining and evaluating a core genome multilocus sequence typing scheme for whole-genome sequence-based typing of Listeria monocytogenes. J Clin Microbiol. 2015;53 (9):2869–2876.
[18] Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47(W1):W256–W259.
[19] Seemann T. Snippy - Rapid haploid variant calling and core genome alignment; 2015 [25 January 2019]. Available from: https://github.com/tseemann/snippy
[20] Seemann T. abricate - Mass screening of contigs for antimicrobial and virulence genes; 2018 [09 August 2019]. Available from: https://github.com/tseemann/abricate
[21] Liu B, Zheng D, Jin Q, et al. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 2019;47(D1):D687–d692.
[22] NCBI. NCBI's AMR reference database 2019. Available from: https://www.ncbi.nlm.nih.gov/pathogens/isolates#/refgene/.
[23] Roedel A, Dieckmann R, Brendebach H, et al. Biocide tolerant Listeria monocytogenes isolates from German food production plants do not show cross-resistance to clinically relevant antibiotics. Appl Environ Microbiol. 2019;85(20):AEM.01253-19.
[24] Romanova NA, Wolffs PFG, Brovko LY, et al. Role of efflux pumps in adaptation and resistance of Listeria monocytogenes to benzalkonium chloride. Appl Environ Microbiol. 2006;72(5):3498–3503.
[25] To MS, Favrin S, Romanova N, et al. Postadaptational resistance to benzalkonium chloride and subsequent physicochemical modifications of Listeria monocytogenes. Appl Environ Microbiol. 2002;68(11):5258–5264.
[26] Clayton EM, Daly KM, Guinane CM, et al. Atypical Listeria innocua strains possess an intact LIPI-3. BMC Microbiol. 2014;14:58.
[27] Elhanafi D, Dutta V, Kathariou S. Genetic characterization of plasmid-associated benzalkonium chloride resistance determinants in a Listeria monocytogenes strain from the 1998–1999 outbreak. Appl Environ Microbiol. 2010;76(24):8231–8238.
[28] Halbedel S, Prager R, Fuchs S, et al. Whole genome sequencing of recent Listeria monocytogenes isolates from Germany reveals population structure and disease clusters. J Clin Microbiol. 2018;56(6):00119-18.
[29] Robert Koch-Institute. Epidemiologisches Bulletin, Nr.41; 2019.
[30] Robert Koch-Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2016; 2017.
[31] Robert Koch-Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2015; 2016.
[32] Robert Koch-Institute. Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2014; 2015.
[33] EFSA Panel on Biological Hazards, Ricci A, Allende A, et al. Listeria monocytogenes contamination of ready-to-eat foods and the risk for human health in the EU. EFSA J. 2018;16(1):e05134.
[34] Maury MM, Tsai YH, Charlier C, et al. Uncovering Listeria monocytogenes hypervirulence by harnessing its biodiversity. Nat Genet. 2016;48(3):308–313.
[35] Milohanic E, Jonquieres R, Cossart P, et al. The autolysin Ami contributes to the adhesion of Listeria monocytogenes to eukaryotic cells via its cell wall anchor. Mol Microbiol. 2001;39(5):1212–1224.
[36] Kirchner M, Higgins DE. Inhibition of ROCK activity allows InlF-mediated invasion and increased virulence of Listeria monocytogenes. Mol Microbiol. 2008;68 (3):749–767.
[37] Reis O, Sousa S, Camejo A, et al. Lapb, a novel Listeria monocytogenes LPXTG surface adhesin, required for entry into eukaryotic cells and virulence. J Infect Dis. 2010;202(4):551–562.
[38] Vilchis-Rangel RE, Espinoza-Mellado MDR, Salinas-Jaramillo IJ, et al. Association of Listeria monocytogenes LIPI-1 and LIPI-3 marker llsX with invasiveness. Curr Microbiol. 2019;76(5):637–643.
[39] Painset A, Björkman JT, Kiil K, et al. LiSEQ – whole-genome sequencing of a cross-sectional survey of Listeria monocytogenes in ready-to-eat foods and human clinical cases in Europe. Microb Genomics. 2019;5(2).
[40] Smith A, Hearn J, Taylor C, et al. Listeria monocytogenes isolates from ready to eat plant produce are diverse and have virulence potential. Int J Food Microbiol. 2019;299:23–32.
[41] Cabal A, Allerberger F, Huhulescu S, et al. Listeriosis outbreak likely due to contaminated liver pâté consumed in a tavern, Austria, December 2018. Eurosurveillance. 2019;24(39):1900274.
[42] Thomas J, Govender N, McCarthy KM, et al. Outbreak of listeriosis in South Africa associated with processed meat. N Engl J Med. 2020;382(7):632–643.
[43] ECDC, EFSA. Rapid outbreak assessment: Multi-country outbreak of Listeria monocytogenes sequence type 6 infections linked to ready-to-eat meat products 2019 [05/2020]. Available from: https://www.ecdc.europa.eu/sites/default/files/documents/Listeria-rapid-outbreak-assessment-NL-BE.pdf
[44] Moretro T, Schirmer BCT, Heir E, et al. Tolerance to quaternary ammonium compound disinfectants may enhance growth of Listeria monocytogenes in the food industry. Int J Food Microbiol. 2017;241: 215–224.
[45] Wilson A, Gray J, Chandry PS, et al. Phenotypic and genotypic analysis of antimicrobial resistance among Listeria monocytogenes isolated from Australian food production chains. Genes. 2018;9(2):80.
[46] Fillgrove KL, Pakhomova S, Schaab MR, et al. Structure and mechanism of the genomically encoded fosfomycin resistance protein, FosX, from Listeria monocytogenes. Biochemistry. 2007;46(27):8110–8120.
[47] Troxler R, von Graevenitz A, Funke G, et al. Natural antibiotic susceptibility of Listeria species: L. grayi, L. innocua, L. ivanovii, *L. monocytogenes*, L. seeligeri and L. welshimeri strains. Clin Microbiol Infect. 2000;6(10):525–535.
[48] Gómez D, Iguácel LP, Rota MC, et al. Occurrence of Listeria monocytogenes in ready-to-eat meat products and meat processing plants in Spain. Foods. 2015;4 (3):271–282.
[49] Stessl B, Szakmary-Brändle K, Vorberg U, et al. Temporal analysis of the Listeria monocytogenes population structure in floor drains during reconstruction and expansion of a meat processing plant. Int J Food Microbiol. 2020;314:108360.
[50] Haralick RM, Elliott GL. Increasing tree search efficiency for constraint satisfaction problems. Artif Intell. 1980;14(3):263–313.

Supplementary Table 1

sample_alias	Accession
13-00476	ERS4419154 (SAMEA6659693)
13-01044	ERS4419155 (SAMEA6659694)
13-03570	ERS4419156 (SAMEA6659695)
13-05213	ERS4419157 (SAMEA6659696)
14-02242	ERS2103036 (SAMEA104485094)
14-03077	ERS2103039 (SAMEA104485097)
14-03633	ERS2103040 (SAMEA104485098)
14-03879	ERS4418847 (SAMEA6659385)
14-04042	ERS4418848 (SAMEA6659386)
14-04253	ERS2103045 (SAMEA104485103)
14-04789	ERS2103047 (SAMEA104485105)
14-04961	ERS2103049 (SAMEA104485107)
14-05083	ERS2103050 (SAMEA104485108)
14-05266	ERS4419158 (SAMEA6659697)
14-05359	ERS4419159 (SAMEA6659698)
14-05360	ERS4419160 (SAMEA6659699)
14-05399	ERS4419161 (SAMEA6659700)
14-05646	ERS4419162 (SAMEA6659701)
14-LI00510-0	ERS4418849 (SAMEA6659387)
15-02696	ERS2103126 (SAMEA104485184)
15-02726	ERS4419163 (SAMEA6659702)
15-03239	ERS2103133 (SAMEA104485191)
15-03310	ERS2103135 (SAMEA104485193)
16-00477	ERS2103172 (SAMEA104485230)
16-00805	ERS4419164 (SAMEA6659703)
16-01265	ERS2103199 (SAMEA104485257)
16-01571	ERS2103215 (SAMEA104485273)
16-01618	ERS2103216 (SAMEA104485274)
16-01908	ERS2103226 (SAMEA104485284)
16-02525	ERS2103252 (SAMEA104485310)
16-02555	ERS2103254 (SAMEA104485312)
16-02585	ERS4418850 (SAMEA6659388)
16-02586	ERS4418851 (SAMEA6659389)
16-02725	ERS2103259 (SAMEA104485317)
16-02841	ERS2103262 (SAMEA104485320)
16-02862	ERS2103266 (SAMEA104485324)
16-02902	ERS2103270 (SAMEA104485328)
16-03056	ERS4419165 (SAMEA6659704)
16-03062	ERS2103284 (SAMEA104485342)
16-03341	ERS2103298 (SAMEA104485356)
16-03643	ERS2103308 (SAMEA104485366)
16-03665	ERS2103309 (SAMEA104485367)
16-03703	ERS2103310 (SAMEA104485368)
16-03780	ERS2103314 (SAMEA104485372)
16-03868	ERS2103324 (SAMEA104485382)
16-03885	ERS4419166 (SAMEA6659705)
16-03886	ERS2103325 (SAMEA104485383)
16-03934	ERS2103326 (SAMEA104485384)
16-04180	ERS2103335 (SAMEA104485393)
16-04181	ERS2103336 (SAMEA104485394)
16-04235	ERS2103337 (SAMEA104485395)
16-LI00360-0	ERS4418852 (SAMEA6659390)
17-00472	ERS2103403 (SAMEA104485461)
17-01287	ERS4418853 (SAMEA6659391)
17-01357	ERS4418854 (SAMEA6659392)
17-01478	ERS4418855 (SAMEA6659393)
17-01870	ERS4418856 (SAMEA6659394)
17-02157	ERS4418857 (SAMEA6659395)
17-02392	ERS4418858 (SAMEA6659396)
17-02419	ERS4419167 (SAMEA6659706)
17-02460	ERS4418859 (SAMEA6659397)
17-02677	ERS4418860 (SAMEA6659398)
17-02864	ERS4418861 (SAMEA6659399)

17-03027	ERS4419168 (SAMEA6659707)
17-03524	ERS4419169 (SAMEA6659708)
17-03596	ERS4419170 (SAMEA6659709)
17-03631	ERS4419171 (SAMEA6659710)
17-05073	ERS4419172 (SAMEA6659711)
17-05076	ERS4419173 (SAMEA6659712)
17-05383	ERS4419174 (SAMEA6659713)
17-05799	ERS4419180 (SAMEA6659719)
17-05803	ERS4419181 (SAMEA6659720)
17-06371	ERS4418862 (SAMEA6659400)
17-LI00727-0	ERS4418863 (SAMEA6659401)
17-LI00728-0	ERS4418864 (SAMEA6659402)
17-LI00729-0	ERS4418865 (SAMEA6659403)
17-LI00730-0	ERS4418866 (SAMEA6659404)
17-LI00766-0	ERS4418867 (SAMEA6659405)
17-LI00767-0	ERS4418868 (SAMEA6659406)
17-LI00811-0	ERS4418869 (SAMEA6659407)
17-LI00916-0	ERS4418870 (SAMEA6659408)
17-LI00917-0	ERS4418871 (SAMEA6659409)
17-LI00969-0	ERS4418872 (SAMEA6659410)
17-LI00970-0	ERS4418873 (SAMEA6659411)
17-LI00971-0	ERS4418874 (SAMEA6659412)
17-LI00972-0	ERS4418875 (SAMEA6659413)
17-LI00973-0	ERS4418876 (SAMEA6659414)
17-LI00974-0	ERS4418877 (SAMEA6659415)
17-LI00975-0	ERS4418878 (SAMEA6659416)
17-LI00976-0	ERS4418879 (SAMEA6659417)
17-LI00977-0	ERS4418880 (SAMEA6659418)
17-LI00978-0	ERS4418881 (SAMEA6659419)
17-LI00980-0	ERS4418882 (SAMEA6659420)
17-LI00983-0	ERS4418883 (SAMEA6659421)
17-LI00984-0	ERS4418884 (SAMEA6659422)
17-LI00985-0	ERS4418885 (SAMEA6659423)
17-LI00986-0	ERS4418886 (SAMEA6659424)
17-LI00987-0	ERS4418887 (SAMEA6659425)
17-LI00988-0	ERS4418888 (SAMEA6659426)
17-LI00990-0	ERS4418889 (SAMEA6659427)
17-LI00991-0	ERS4418890 (SAMEA6659428)
17-LI00992-0	ERS4418891 (SAMEA6659429)
17-LI00993-0	ERS4418892 (SAMEA6659430)
17-LI00994-0	ERS4418893 (SAMEA6659431)
17-LI00995-0	ERS4418894 (SAMEA6659432)
17-LI00996-0	ERS4418895 (SAMEA6659433)
17-LI00997-0	ERS4418896 (SAMEA6659434)
17-LI00998-0	ERS4418897 (SAMEA6659435)
17-LI00999-0	ERS4418898 (SAMEA6659436)
17-LI01000-0	ERS4418899 (SAMEA6659437)
17-LI01002-0	ERS4418900 (SAMEA6659438)
17-LI01003-0	ERS4418901 (SAMEA6659439)
17-LI01005-0	ERS4418902 (SAMEA6659440)
17-LI01006-0	ERS4418903 (SAMEA6659441)
17-LI01007-0	ERS4418904 (SAMEA6659442)
17-LI01008-0	ERS4418905 (SAMEA6659443)
17-LI01009-0	ERS4418906 (SAMEA6659444)
17-LI01010-0	ERS4418907 (SAMEA6659445)
17-LI01011-0	ERS4418908 (SAMEA6659446)
17-LI01012-0	ERS4418909 (SAMEA6659447)
17-LI01013-0	ERS4418910 (SAMEA6659448)
17-LI01014-0	ERS4418911 (SAMEA6659449)
17-LI01015-0	ERS4418912 (SAMEA6659450)
17-LI01016-0	ERS4418913 (SAMEA6659451)
17-LI01017-0	ERS4418914 (SAMEA6659452)
17-LI01018-0	ERS4418915 (SAMEA6659453)
17-LI01019-0	ERS4418916 (SAMEA6659454)
17-LI01020-0	ERS4418917 (SAMEA6659455)
17-LI01021-0	ERS4418918 (SAMEA6659456)

17-LI01022-0	ERS4418919 (SAMEA6659457)
17-LI01023-0	ERS4418920 (SAMEA6659458)
17-LI01024-0	ERS4418921 (SAMEA6659459)
17-LI01031-0	ERS4418922 (SAMEA6659460)
17-LI01032-0	ERS4418923 (SAMEA6659461)
17-LI01033-0	ERS4418924 (SAMEA6659462)
18-00200	ERS4419175 (SAMEA6659714)
18-00305	ERS4418925 (SAMEA6659463)
18-03370	ERS4419176 (SAMEA6659715)
18-04057	ERS4419177 (SAMEA6659716)
18-04187	ERS4419178 (SAMEA6659717)
18-05769	ERS4419179 (SAMEA6659718)
18-LI00244-0	ERS4418926 (SAMEA6659464)
18-LI00245-0	ERS4418927 (SAMEA6659465)
18-LI00246-0	ERS4418928 (SAMEA6659466)
18-LI00247-0	ERS4418929 (SAMEA6659467)
18-LI00248-0	ERS4418930 (SAMEA6659468)
18-LI00249-0	ERS4418931 (SAMEA6659469)
18-LI00250-0	ERS4418932 (SAMEA6659470)
18-LI00251-0	ERS4418933 (SAMEA6659471)
18-LI00252-0	ERS4418934 (SAMEA6659472)
18-LI00253-0	ERS4418935 (SAMEA6659473)
18-LI00254-0	ERS4418936 (SAMEA6659474)
18-LI00255-0	ERS4418937 (SAMEA6659475)
18-LI00256-0	ERS4418938 (SAMEA6659476)
18-LI00257-0	ERS4418939 (SAMEA6659477)
18-LI00258-0	ERS4418940 (SAMEA6659478)
18-LI00259-0	ERS4418941 (SAMEA6659479)
18-LI00260-0	ERS4418942 (SAMEA6659480)
18-LI00261-0	ERS4418943 (SAMEA6659481)
18-LI00262-0	ERS4418944 (SAMEA6659482)
18-LI00263-0	ERS4418945 (SAMEA6659483)
18-LI00264-0	ERS4418946 (SAMEA6659484)
18-LI00265-0	ERS4418947 (SAMEA6659485)
18-LI00266-0	ERS4418948 (SAMEA6659486)
18-LI00267-0	ERS4418949 (SAMEA6659487)
18-LI00268-0	ERS4418950 (SAMEA6659488)
18-LI00269-0	ERS4418951 (SAMEA6659489)
18-LI00270-0	ERS4418952 (SAMEA6659490)
18-LI00271-0	ERS4418953 (SAMEA6659491)
18-LI00272-0	ERS4418954 (SAMEA6659492)
18-LI00273-0	ERS4418955 (SAMEA6659493)
18-LI00274-0	ERS4418956 (SAMEA6659494)
18-LI00275-0	ERS4418957 (SAMEA6659495)
18-LI00276-0	ERS4418958 (SAMEA6659496)
18-LI00277-0	ERS4418959 (SAMEA6659497)
18-LI00278-0	ERS4418960 (SAMEA6659498)
18-LI00279-0	ERS4418961 (SAMEA6659499)
18-LI00280-0	ERS4418962 (SAMEA6659500)
18-LI00281-0	ERS4418963 (SAMEA6659501)
18-LI00282-0	ERS4418964 (SAMEA6659502)
18-LI00283-0	ERS4418965 (SAMEA6659503)
18-LI00284-0	ERS4418966 (SAMEA6659504)
18-LI00285-0	ERS4418967 (SAMEA6659505)
18-LI00286-0	ERS4418968 (SAMEA6659506)
18-LI00287-0	ERS4418969 (SAMEA6659507)
18-LI00288-0	ERS4418970 (SAMEA6659508)
18-LI00289-0	ERS4418971 (SAMEA6659509)
18-LI00290-0	ERS4418972 (SAMEA6659510)
18-LI00291-0	ERS4418973 (SAMEA6659511)
18-LI00292-0	ERS4418974 (SAMEA6659512)
18-LI00293-0	ERS4418975 (SAMEA6659513)
18-LI00294-0	ERS4418976 (SAMEA6659514)
18-LI00295-0	ERS4418977 (SAMEA6659515)
18-LI00296-0	ERS4418978 (SAMEA6659516)
18-LI00297-0	ERS4418979 (SAMEA6659517)

18-LI00298-0	ERS4418980 (SAMEA6659518)
18-LI00300-0	ERS4418981 (SAMEA6659519)
18-LI00301-0	ERS4418982 (SAMEA6659520)
18-LI00302-0	ERS4418983 (SAMEA6659521)
18-LI00303-0	ERS4418984 (SAMEA6659522)
18-LI00304-0	ERS4418985 (SAMEA6659523)
18-LI00305-0	ERS4418986 (SAMEA6659524)
18-LI00306-0	ERS4418987 (SAMEA6659525)
18-LI00330-0	ERS4418988 (SAMEA6659526)
18-LI00331-0	ERS4418989 (SAMEA6659527)
18-LI00332-0	ERS4418990 (SAMEA6659528)
18-LI00333-0	ERS4418991 (SAMEA6659529)
18-LI00334-0	ERS4418992 (SAMEA6659530)
18-LI00335-0	ERS4418993 (SAMEA6659531)
18-LI00336-0	ERS4418994 (SAMEA6659532)
18-LI00337-0	ERS4418995 (SAMEA6659533)
18-LI00338-0	ERS4418996 (SAMEA6659534)
18-LI00339-0	ERS4418997 (SAMEA6659535)
18-LI00340-0	ERS4418998 (SAMEA6659536)
18-LI00341-0	ERS4418999 (SAMEA6659537)
18-LI00342-0	ERS4419000 (SAMEA6659538)
18-LI00343-0	ERS4419001 (SAMEA6659539)
18-LI00344-0	ERS4419002 (SAMEA6659540)
18-LI00345-0	ERS4419003 (SAMEA6659541)
18-LI00346-0	ERS4419004 (SAMEA6659542)
18-LI00347-0	ERS4419005 (SAMEA6659543)
18-LI00348-0	ERS4419006 (SAMEA6659544)
18-LI00349-0	ERS4419007 (SAMEA6659545)
18-LI00350-0	ERS4419008 (SAMEA6659546)
18-LI00351-0	ERS4419009 (SAMEA6659547)
18-LI00352-0	ERS4419010 (SAMEA6659548)
18-LI00353-0	ERS4419011 (SAMEA6659549)
18-LI00354-0	ERS4419012 (SAMEA6659550)
18-LI00414-0	ERS4419013 (SAMEA6659551)
18-LI00415-0	ERS4419014 (SAMEA6659552)
18-LI00417-0	ERS4419015 (SAMEA6659553)
18-LI00418-0	ERS4419016 (SAMEA6659554)
18-LI00419-0	ERS4419017 (SAMEA6659555)
18-LI00420-0	ERS4419018 (SAMEA6659556)
18-LI00421-0	ERS4419019 (SAMEA6659557)
18-LI00422-0	ERS4419020 (SAMEA6659558)
18-LI00487-0	ERS4419021 (SAMEA6659559)
18-LI00488-0	ERS4419022 (SAMEA6659560)
18-LI00491-0	ERS4419023 (SAMEA6659561)
18-LI00539-0	ERS4419024 (SAMEA6659562)
18-LI00540-0	ERS4419025 (SAMEA6659563)
18-LI00541-0	ERS4419026 (SAMEA6659564)
18-LI00542-0	ERS4419027 (SAMEA6659565)
18-LI00543-0	ERS4419028 (SAMEA6659566)
18-LI00635-0	ERS4419029 (SAMEA6659567)
18-LI00636-0	ERS4419030 (SAMEA6659568)
18-LI00643-0	ERS4419031 (SAMEA6659569)
18-LI00714-0	ERS4419032 (SAMEA6659570)
18-LI00715-0	ERS4419033 (SAMEA6659571)
18-LI00716-0	ERS4419034 (SAMEA6659572)
18-LI00717-0	ERS4419035 (SAMEA6659573)
18-LI00718-0	ERS4419036 (SAMEA6659574)
18-LI00719-0	ERS4419037 (SAMEA6659575)
18-LI00720-0	ERS4419038 (SAMEA6659576)
18-LI00722-0	ERS4419039 (SAMEA6659577)
18-LI00723-0	ERS4419040 (SAMEA6659578)
18-LI00724-0	ERS4419041 (SAMEA6659579)
18-LI00725-0	ERS4419042 (SAMEA6659580)
18-LI00726-0	ERS4419043 (SAMEA6659581)
18-LI00727-0	ERS4419044 (SAMEA6659582)
18-LI00728-0	ERS4419045 (SAMEA6659583)

18-LI00729-0	ERS4419046 (SAMEA6659584)
18-LI00730-0	ERS4419047 (SAMEA6659585)
18-LI00731-0	ERS4419048 (SAMEA6659586)
18-LI00732-0	ERS4419049 (SAMEA6659587)
18-LI00733-0	ERS4419050 (SAMEA6659588)
18-LI00734-0	ERS4419051 (SAMEA6659589)
18-LI00735-0	ERS4419052 (SAMEA6659590)
18-LI00736-0	ERS4419053 (SAMEA6659591)
18-LI00737-0	ERS4419054 (SAMEA6659592)
18-LI00738-0	ERS4419055 (SAMEA6659593)
18-LI00739-0	ERS4419056 (SAMEA6659594)
18-LI00740-0	ERS4419057 (SAMEA6659595)
18-LI00741-0	ERS4419058 (SAMEA6659596)
18-LI00742-0	ERS4419059 (SAMEA6659597)
18-LI00743-0	ERS4419060 (SAMEA6659598)
18-LI00744-0	ERS4419061 (SAMEA6659599)
18-LI00745-0	ERS4419062 (SAMEA6659600)
18-LI00746-0	ERS4419063 (SAMEA6659601)
18-LI00747-0	ERS4419064 (SAMEA6659602)
18-LI00748-0	ERS4419065 (SAMEA6659603)
18-LI00749-0	ERS4419066 (SAMEA6659604)
18-LI00750-0	ERS4419067 (SAMEA6659605)
18-LI00751-0	ERS4419068 (SAMEA6659606)
18-LI00752-0	ERS4419069 (SAMEA6659607)
18-LI00753-0	ERS4419070 (SAMEA6659608)
18-LI00754-0	ERS4419071 (SAMEA6659609)
18-LI00755-0	ERS4419072 (SAMEA6659610)
18-LI00756-0	ERS4419073 (SAMEA6659611)
18-LI00757-0	ERS4419074 (SAMEA6659612)
18-LI00759-0	ERS4419075 (SAMEA6659613)
18-LI00760-0	ERS4419076 (SAMEA6659614)
18-LI00761-0	ERS4419077 (SAMEA6659615)
18-LI00762-0	ERS4419078 (SAMEA6659616)
18-LI00766-0	ERS4419079 (SAMEA6659617)
18-LI00767-0	ERS4419080 (SAMEA6659618)
18-LI00768-0	ERS4419081 (SAMEA6659619)
18-LI00769-0	ERS4419082 (SAMEA6659620)
18-LI00770-0	ERS4419083 (SAMEA6659621)
18-LI00771-0	ERS4419084 (SAMEA6659622)
18-LI00772-0	ERS4419085 (SAMEA6659623)
18-LI00773-0	ERS4419086 (SAMEA6659624)
18-LI00774-0	ERS4419087 (SAMEA6659625)
18-LI00775-0	ERS4419088 (SAMEA6659626)
18-LI00779-0	ERS4419089 (SAMEA6659627)
18-LI00780-0	ERS4419090 (SAMEA6659628)
18-LI00781-0	ERS4419091 (SAMEA6659629)
18-LI00782-0	ERS4419092 (SAMEA6659630)
18-LI00783-0	ERS4419093 (SAMEA6659631)
18-LI00784-0	ERS4419094 (SAMEA6659632)
18-LI00785-0	ERS4419095 (SAMEA6659633)
18-LI00786-0	ERS4419096 (SAMEA6659634)

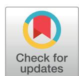
Food Control 96 (2019) 479–487

Contents lists available at ScienceDirect

Food Control

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

Analysis of RASFF notifications on food products contaminated with *Listeria monocytogenes* reveals options for improvement in the rapid alert system for food and feed



Stefanie Lüth[a,b,*], Idesbald Boone[a,1], Sylvia Kleta[a], Sascha Al Dahouk[a,c]

[a] German Federal Institute for Risk Assessment, Diedersdorfer Weg 1, 12277, Berlin, Germany
[b] Freie Universität Berlin, Königin-Luise-Str. 1-3, 14195, Berlin, Germany
[c] RWTH Aachen University Hospital, Pauwelsstraße 30, 52074, Aachen, Germany

ARTICLE INFO

Keywords:
European food safety
RASFF
Reporting
Notification
Listeria monocytogenes
Contamination
Germany

ABSTRACT

Tools for supranational communication of food safety risks like the European Rapid Alert System for Food and Feed (RASFF) play an increasingly important role in consumer protection along global supply chains. They allow for a coordinated response to emerging public health threats such as *Listeria monocytogenes* (*Lm*), the causative agent of the foodborne disease listeriosis. As a result of disease severity and the rising number of human listeriosis cases in Germany since 2011, an effective reporting system on *Lm* contamination in food products has become more crucial than ever to counteract this trend. Therefore, we analysed RASFF notifications on food products contaminated with *Lm* and distributed in Germany, 2001 to 2015, assessed trends in the reported data and addressed options for improvement in the current notification system.

In RASFF *Lm* notifications concerning Germany from 2001 to 2015, there was often a discrepancy between country of origin and notifying country, indicating that the food safety risk was not always recognised at the earliest possible time point of the product's life span. In addition, in our dataset, most *Lm* notifications were driven by official controls when the respective product was already on the market. However, starting in 2005, there was an increasing trend for company's own checks. This trend of making food manufacturers accountable for the detection and notification of contaminated products in the production line is a first step into the right direction as it might help to reduce the number of contaminated food products that enter the market.

Besides its function as a reporting tool, the RASFF may also facilitate the identification of risk factors associated with *Lm* contamination so that the problem can be tackled at its root. Unfortunately, information about packaging and food processing was only mentioned in a minority of *Lm* notifications. Hence, risk factors cannot be easily identified. In the future, a comprehensive database including additional metadata together with the RASFF notification should be established.

Although a solid basis for the surveillance of *Lm*, there is still room for improvement in RASFF to speed-up the flow of information. This might help to identify food safety risks that can be harmful to European consumers much faster, more effectively prevent the spread of risk bearing food products and consequently reduce the burden of listeriosis.

1. Introduction

Globalisation of food trade opens the door for the spread of foodborne infectious diseases. In 2016, a total of 2536 human listeriosis cases were reported in the European Union (EU), corresponding to an incidence rate of 0.47 per 100,000 population (EFSA & ECDC, 2017). Although comparably rare (e. g. incidence rate for salmonellosis 21.2 per 100,000 population), listeriosis has the highest hospitalisation rate of all foodborne zoonoses under EU surveillance (EFSA & ECDC, 2017). Together with its high case fatality rate (16.2% in 2016), it is rightly considered a major EU-wide public health concern (EFSA & ECDC, 2017). The most affected groups are elderly people, immunocompromised patients and pregnant women where the disease is often associated with severe clinical manifestations including

* Corresponding author. German Federal Institute for Risk Assessment, Diedersdorfer Weg 1, 12277, Berlin, Germany.
E-mail address: stefanie.lueth@bfr.bund.de (S. Lüth).
[1] Present address: Robert Koch-Institute, Seestraße 10, 13353, Berlin, Germany.

https://doi.org/10.1016/j.foodcont.2018.09.033
Received 6 July 2018; Received in revised form 25 September 2018; Accepted 26 September 2018
Available online 04 October 2018
0956-7135/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license
(http://creativecommons.org/licenses/BY-NC-ND/4.0/).

septicaemia, meningitis and spontaneous abortion (Schlech & Acheson, 2000).

The causative agent of listeriosis is *Listeria monocytogenes* (*Lm*). The saprophytic bacterium is widely distributed in the environment and can be found in plants, soil, sewage and livestock (Farber & Peterkin, 1991). It is highly adaptable and able to cope with a wide range of environmental stress factors, including temperatures from 0 to 45 °C, a pH range from 4.1 to 9.6 and salt concentrations up to 10% *w/v* NaCl (Lungu, Ricke, & Johnson, 2009; Milillo et al., 2012). These properties allow the pathogen to survive even in preserved foodstuffs stored in cold chains of modern food production and retail systems, thereby constituting a serious problem in food industry.

The vast majority of human listeriosis cases are supposed to be foodborne (Swaminathan & Gerner-Smidt, 2007). There are two ways in which *Lm* can find its way into the food chain: either via primary contamination of raw animal products or via cross-contamination during food processing. In the latter case, niche adaptation of the pathogen, insufficient hygiene measures in food processing plants and as a consequence persistence of *Lm* play an important role (Carpentier & Cerf, 2011). Especially in ready-to-eat (RTE) products, per definition “food intended […] for direct human consumption without the need for cooking or other processing effective to eliminate or reduce to an acceptable level micro-organisms of concern” (European Commission, 2005), *Lm* contamination poses a risk to human health. Therefore, Regulation (EC) 2073/2005 was implemented in 2005 to define microbiological food safety criteria which also apply for *Lm* in RTE foods (European Commission, 2005). Typical examples of such RTE products are deli meat, salads, sandwiches and cheese. If intended for infants or special medical purposes, *Lm* has to be absent in 25 g of the RTE product placed on the market during its shelf life. For all other target groups, during shelf life, a limit of 100 CFU/g applies. However, a further distinction is made between products that are, or are not able to support growth of *Lm*. If the RTE product is able to support growth of *Lm*, at the time before the food has left the immediate control of the food business operator who has produced it, absence of *Lm* in 25 g RTE product is mandatory as well. An exception only applies when the food business operator shows that the limit of 100 CFU/g is not exceeded during shelf life (European Commission, 2005). Despite these well-defined criteria, listeriosis infections still show an alarmingly increasing trend (ECDC, 2016).

In order to manage interconnected food safety issues arising from international trade, the Rapid Alert System for Food and Feed (RASFF) was initiated in 32 countries of the EU and the European Economic Area (EEA) (European Commission, 2017a). Based on Regulation (EC) 178/2002 and Regulation (EC) 16/2011, the RASFF is intended to provide information on food-related, serious, direct or indirect risks to human health to allow an immediate and coordinated response to emerging threats (European Commission, 2002, 2011). RASFF notifications reported by a national food safety authority are verified by the European Commission (EC) as the manager of the system and disseminated to contact points of network members. Affected products can then be traced back and measures can be taken. There are three main types of RASFF notifications: alert, information and border rejection notifications (European Commission, 2017b). Border rejection notifications concern products rejected at the external borders of the EU/EEA, whereas the other two notification types relate to products inside the EU/EEA. Alert and information notifications mainly differ in the speed and type of reaction that is necessary after release of the notification: alert notifications do, whereas information notifications do not require rapid action in other RASFF member countries. Regardless of the notification type, every notification is based on the identification of a risk to human health. Since implementation of Regulation (EC) 16/2011 in 2011, a further subdivision of information notifications into “information notification for follow-up” and “information notification for attention” was introduced (European Commission, 2011). An “information notification for follow-up” is related to a product that is or may be placed on the market in another country and hence similar to an alert notification, although not requiring rapid action. An “information notification for attention” is released if the product is only present in the notifying country, if it is no longer on the market or if it has not even been placed on the market.

In our study, we analysed RASFF notifications on pathogenic microorganisms (PMF) with a special focus on *Lm*, associated with contaminated food products. Based on RASFF *Lm* notifications affecting Germany, 2001 to 2015, we assessed trends in the reports available and tried to identify shortcomings in the current notification system. Although a solid basis for the surveillance of *Lm*, there is still room for improvement in the RASFF to allow for a more detailed risk assessment and earlier reaction to improve consumer safety and finally reduce the burden of listeriosis.

2. Methods

Data were extracted from the RASFF portal (European Commission, 2017b). Search criteria for RASFF PMF notifications with involvement of Germany were “Notified from: 01/01/2001“, “Notified till: 31/12/2015“, “Product type: food”, “Category: pathogenic micro-organisms” and “Country: Germany (DE)” (query from 20/06/2016, last update 12/06/2017). Search criteria for RASFF *Lm* notifications independent of the country involved were “Notified from: 01/01/2001“, “Notified till: 31/12/2015“, “Category: pathogenic micro-organisms” and “Hazard: *Listeria monocytogenes*” (query from 04/04/2016, last update 12/06/2017). As the years 2001–2003 contained incomplete data on origin and distribution of the contaminated food products, the whole datasets for this period were excluded from analysis. Data on RASFF *Lm* notifications in Germany (period 2001 to 2015) were extracted from the German Federal Office of Consumer Protection and Food Safety (BVL) database using an SQL-based algorithm and completed with information from the original RASFF pdf-documents.

Data selected from all three datasets were transferred to Microsoft Excel 2010 (Microsoft Corp., Redmond, USA) to create descriptive statistics, including frequency distributions (Pivot tables, with filtering). The main filter categories took into account a country's role in notification. A country can publish a notification (notifying country), it can be affected by a notification if the product is distributed in the country (affected country) or it can be country of origin of a product. Involvement of a country is stated if at least one of these three conditions is met.

The script for generation of the Chord diagram (Fig. 3) is available at https://github.com/mattflor/chorddiag.

3. Results

3.1. RASFF PMF notifications, involving Germany

Between 2004 and 2015, 1303 PMF notifications with involvement of Germany were published. Of these, 71% (n = 935) were notified by four countries: Germany (29%, n = 381), Denmark (19%, n = 249), France (13%, n = 169) and Italy (10%, n = 136). A total of 719 notifications (55%) were related to products distributed in Germany, while 455 (35%) were related to products exported from Germany without being distributed in Germany. The remaining 129 notifications (10%) were related to products where Germany was involved in the notification process, but neither as country of origin nor as affected country. In 381 PMF notifications (29%) Germany was notifying country, country of origin in 562 notifications (43%) and affected country in 719 notifications (55%). The number of PMF notifications with Germany as notifying country or country of origin was fluctuating from 2004 to 2015 (18–48 notifications per year as notifying country, 31 to 74 notifications per year as country of origin), as well as the share of *Lm* notifications of all PMF notifications (4–31% per year with Germany as notifying country, 6–29% per year with Germany as country of origin).

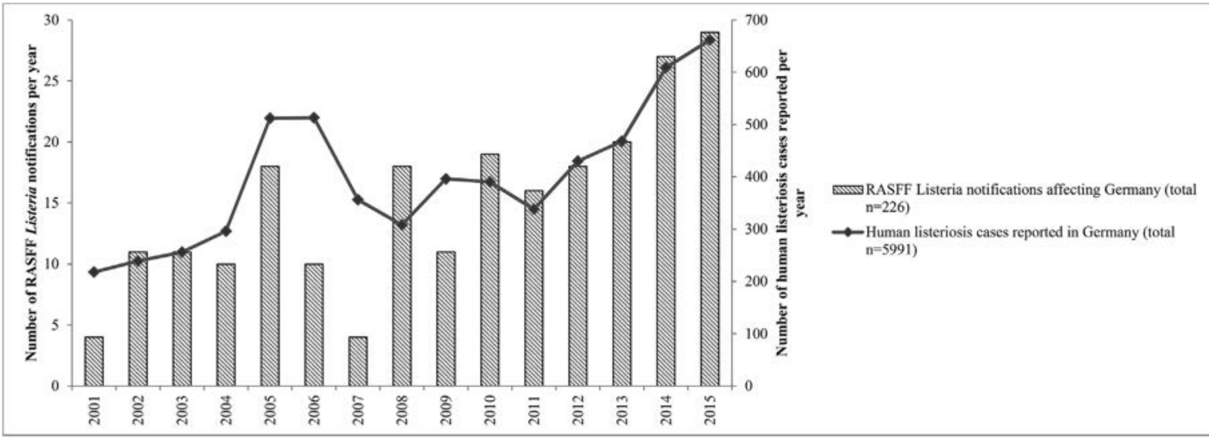


Fig. 1. RASFF *Listeria monocytogenes* notifications for food products affecting Germany (dashed) and human listeriosis cases officially reported in Germany (black), 2001–2015. As a result of the fact that human listeriosis cases cannot be linked to a certain RASFF notification, no reliable causal relationship can be attested. Still, an increasing trend can be seen in both datasets.

In contrast to that, PMF notifications affecting Germany have been increasing since 2011 from 54 to 111 per year in 2015. However, the proportion of *Lm* notifications in these years remained relatively stable between 24 and 31%.

3.2. RASFF Lm notifications

Among all European PMF notifications from 2004 to 2015, 968 concerned *Lm* (*Lm* notifications). France and Italy were notifying countries in almost half of these notifications (France 25%, n = 242; Italy 24%, n = 235), followed by Germany (6%, n = 62), the Netherlands (5%, n = 45), Spain (5%, n = 44) and Poland (4%, n = 43). In 27% (n = 266) of cases, Italy was affected, in 26% (n = 254) of cases, it was France. Germany, Belgium and the Netherlands were among the affected countries in 20% (n = 198), 15% (n = 146) and 11% (n = 106) of notifications, respectively.

3.3. RASFF Lm notifications, involving Germany

From 2001 to 2015, a total of 312 Lm notifications with involvement of Germany were published. The number of *Lm* notifications with Germany as country of origin (total n = 98, 31%) ranged between 2 and 15 notifications per year. Out of these notifications, 73 (74.5%) concerned products of German origin, meant for export only. Products mentioned in 16 notifications (16.3%) were distributed both in Germany and abroad, whereas 9 notifications (9%) were related to

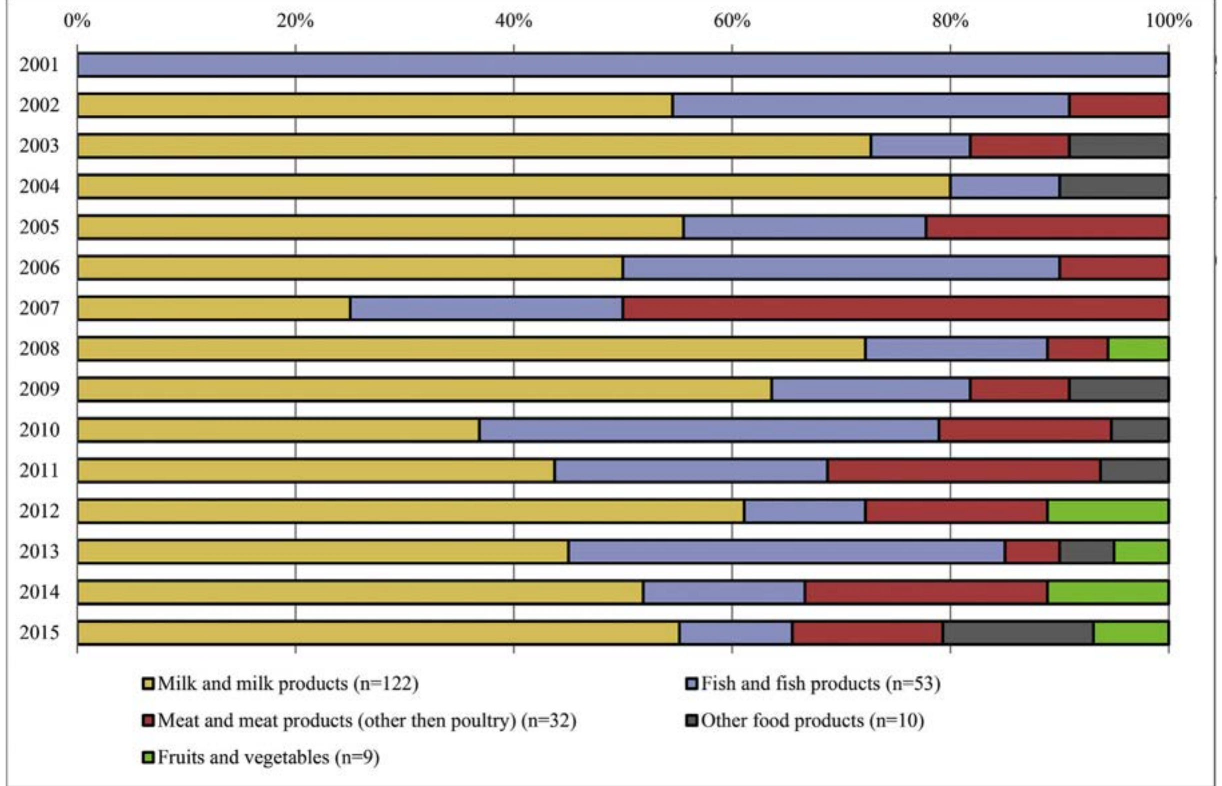


Fig. 2. RASFF *Listeria monocytogenes* notifications for products affecting Germany (n = 226), 2001–2015. Percentages of notifications related to a product category (simplified categories) per year of notification are shown.

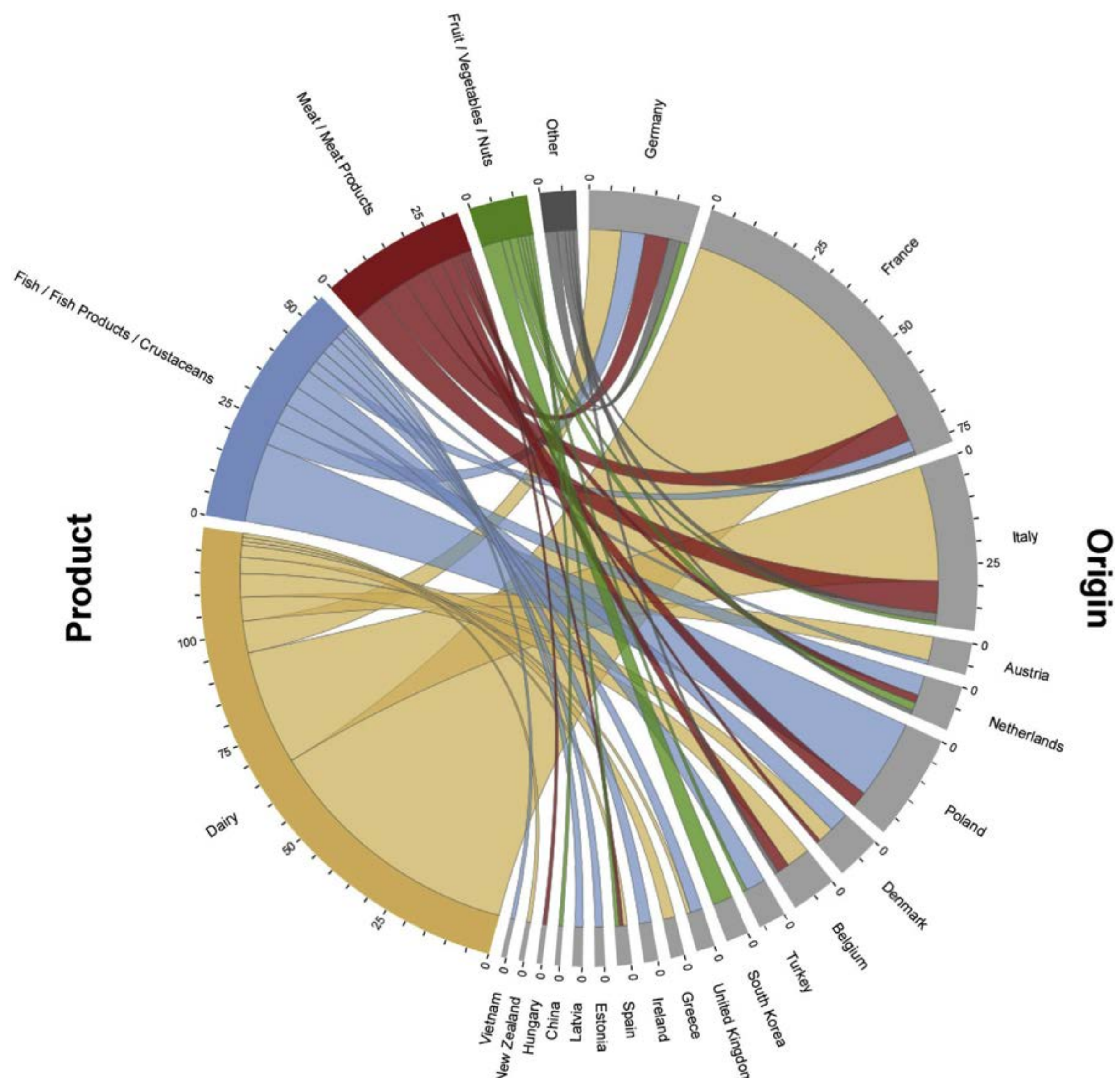


Fig. 3. Product categories of RASFF *Listeria monocytogenes* notifications affecting Germany, by country of origin, 2001–2015. Arc lengths on the outer circle are proportional to the number of notifications of a product category or to the number of notifications with a specific country of origin.

products of German origin, only distributed in Germany. The proportions per year were variable from year to year, but consistently with a main focus on exported products.

3.4. RASFF Lm notifications, affecting Germany

At total of 226 Lm notifications were published between 2001 and 2015 for products affecting Germany. For 33% (n = 75) of these notifications, Germany was the notifying country, followed by France (30%, n = 68), Italy (9%, n = 20) and Austria (6%, n = 13). In Germany, notifications were primarily made in two German Federal States: Baden-Württemberg (33%, n = 25) and Hesse (17%, n = 13). *Lm* notifications affecting Germany have been increasing since 2011, culminating in 29 notifications in 2015. Likewise, the number of human listeriosis cases reported has been increasing in Germany since 2011, reaching a maximum of 662 cases in 2015 (SURVSTAT@RKI 2.0) (Fig. 1).

Products affecting Germany had their origin in 15 EU member states and five non-European countries. Eight notifications were related to multiple countries of origin. Products from four countries accounted for 73% (n = 165) of the notifications: France (34%, n = 77), Italy (17%, n = 39), Germany (11%, n = 25) and Poland (11%, n = 24). *Lm* notifications for products of French origin were mainly notified by France (61 of 77 notifications, 79%). Italian, Polish and German products were mainly notified by Germany (21 of 39, 54%; 14 of 24, 58%; 22 of 25, 88%, respectively) (Table 1).

3.4.1. Notification basis

Most of the 226 *Lm* notifications affecting Germany between 2001 and 2015 were based on official controls on the market (54%, n = 122). From these, most were notified by Germany (45%, n = 55), followed by France (18%, n = 22) and Italy (12%, n = 15). For Germany, Italy and also Austria, the overall percentage of notifications based on official controls on the market clearly exceeded that of notifications based on company's own check (55 based on official controls *versus* 16 based on company's own check, 15 *versus* 4 and 10 *versus* 2, respectively). In contrast to that, notifications from France were mainly due to findings in company's own checks (22 based on official controls *versus* 44 based

Table 1
RASFF *Listeria monocytogenes* notifications affecting Germany, 2001–2015, by country of origin and country of notification.
Note: For eight notifications, there was more than one country of origin (multiple attributions).

Country of origin	Country of notification																Sum
	AT	BE	CZ	DE	DK	EE	ES	FR	GR	IE	IT	LU	NL	PL	UK	NO	
AT	6			1													7
BE		6		1				2					1				10
DK				1	5			1			2					1	10
EE						2											2
FR	1			8				61			2	1	1			3	77
DE	1		1	22				1									25
GR				2					1								3
HU														1			1
IE								1		2							3
IT	3			21			1	2			11					2	40
LV				1							1						2
NL	3	1		4					1				1				10
PL				14				1			1			8			24
ES				2											1		3
UK				1							2				1		4
CN		1															1
NZ													1				1
KR									1				4				5
TR				2									4				6
VN											1						1
Sum	14	8	1	80	5	2	1	69	3	2	20	1	12	9	2	6	

AT-Austria, BE-Belgium, CZ-Czech Republic, DK-Denmark, EE-Estonia, FR-France, DE-Germany, GR-Greece, HU-Hungary, IE-Ireland, IT-Italy, LU-Luxembourg, LV-Latvia, NL-Netherlands, NO-Norway, PL-Poland, ES-Spain, UK-United Kingdom, CN-China, NZ-New Zealand, KR-South Korea, TR-Turkey, VN-Vietnam.

on company's own check). Overall, company's own check contributed to 42% (n = 95) of all *Lm* notifications affecting Germany, 2001 to 2015. While the number of notifications based on official controls on the market has been fluctuating between 3 and 15 per year from 2001 to 2015, notifications based on company's own check first appeared in 2005 from which on they showed an increasing trend (3 in 2005 to 17 in 2015). In five cases (one per year from 2008 to 2011, and another one in 2014), food poisoning was mentioned as notification basis. In three notifications, a RASFF notification was issued as a result of an official control in a non-EU member state (one in 2011, two in 2012).

3.4.2. *Affected food categories*

The majority of *Lm* notifications affecting Germany, 2001 to 2015, concerned milk and milk products (54%, n = 122), followed by fish and fish products (23%, n = 53) and meat and meat products other than poultry (12%, n = 26). Within the milk and milk products, the majority of notifications were related to soft cheese (n = 82, 67% of milk and milk products) (Table 2). While notifications based on company's own check mainly concerned milk and milk products (63%, n = 60), a large number of notifications based on official control on the market were related to fish and fish products (31%, n = 38). From 2001 to 2015, the proportion of food categories complained in *Lm* notifications every year varied, shifting the main focus between milk and milk products, fish and fish products and meat and meat products other than poultry (Fig. 2). In the product category milk and milk products, 54% (n = 66) of *Lm* notifications were related to products from France. For 38% (n = 20) of the notifications concerning fish and fish products Poland was stated as country of origin. For the meat and meat products, the country of origin was very diverse (Fig. 3).

Table 2
RASFF *Listeria monocytogenes* notifications for products affecting Germany, 2001–2015, by food category.

Product category	Total
Milk and milk products	122
soft cheese	82
unspecified	12
fresh cheese	10
semi-soft cheese	8
semi-hard cheese	6
hard cheese	2
sour milk cheese	2
Fish and fish products	53
Meat and meat products (other than poultry)	26
Fruits and vegetables	9
Poultry meat and poultry products	6
Prepared dishes and snacks	4
Cereals and bakery products	2
Crustaceans and products thereof	1
Nuts, nut products and seeds	1
Soups, broths, sauces and condiments	1
Other food product/mixed	1

3.4.3. *Involved companies*

Products of 30 companies (from a total of 176 companies) were the cause of two or more notifications between 2001 and 2015. Twelve out of these companies were mentioned in notifications in two separate years, two companies caused notifications in three different years and one company was associated with notifications in six different years. With respect to milk and milk products, France had the highest number of companies (n = 10) involved in two or more notifications, totalling 25 notifications. Concerning the category of fish and fish products, 17 notifications were associated with two Polish companies.

3.4.4. *Notification types*

Lm notifications affecting Germany were classified as "alert notification" in 175 of 226 notifications (77%), with the majority of alerts issued by France (37%, n = 65) and Germany (23%, n = 41). The remaining 51 notifications (23%) were information notifications. Due to a lack of further subdivision before 2011, 31 of them were only classified as "information notification". After the subdivision of the category in 2011, 13 notifications were classified as "information for attention" and 7 "information for follow-up". Germany was responsible for the majority of the three kinds of information notifications (between 57 and 71%).

3.4.5. *Additional metadata*

3.4.5.1. Packaging type and slicing category. Information for example

Table 3
RASFF *Listeria monocytogenes* notifications for products affecting Germany, 2001–2015, by food category and stabilisation category or raw/non raw category; for other product categories data are missing (unspecified).

Food category	Total	Metadata category
Milk and milk products	122	
raw	49	raw/non-raw
non-raw	12	
unspecified	61	
from raw milk	49	stabilisation
from pasteurised milk	10	
salted	5	
from sour milk	2	
from raw and pasteurised milk	1	
unspecified	55	
Fish and fish products	53	
raw	48	raw/non-raw
unspecified	5	
smoked	39	stabilisation
graved	2	
salted	2	
graved, marinated	1	
marinated	1	
smoked, salted	1	
unspecified	7	
Meat and meat products (other than poultry)	26	
raw	9	raw/non-raw
non-raw	3	
unspecified	14	
cooked	3	stabilisation
smoked	2	
dried	1	
fermented	1	
salted	1	
unspecified	18	
Poultry meat and poultry products	6	
non-raw	1	raw/non-raw
unspecified	5	
cooked	1	stabilisation
unspecified	5	
Crustaceans and products thereof	1	
non-raw	1	raw/non-raw
cooked	1	stabilisation

about packaging or food processing was only provided for a minority of *Lm* notifications affecting Germany. Data on packaging type was available for 16% (n = 37) of notifications, information about the slicing category in 23% (n = 51) of notifications. In the category fish and fish products at least for 20 of 52 notifications (38%), the packaging type was specified. Among these, 15 products (75%) were packaged under vacuum conditions. For the same category, slicing information was available for 34 (65%) notifications. These products were mainly filleted (50%, n = 17) or sliced (38%, n = 13). For meat and meat products other than poultry, information on slicing category was available for only 5 out of 26 notifications (19%), with four products sliced (80%) and one cut (20%).

3.4.5.2. Raw/non-raw status and stabilisation category. The majority of *Lm* notifications affecting Germany concerned raw products (47%, n = 107), but a large proportion of notifications lacked information (45%, n = 101) (Table 3). The raw/non-raw status can be differentiated from the stabilisation category of a product. As in some notifications, not both of the corresponding metadata fields were filled, information could not be combined. *Lm* notifications for milk and milk products mainly concerned the stabilisation category "from raw milk" (42%, n = 49). However, in 45% (n = 55) of notifications in this product category, the stabilisation method was not specified. Within the fish and fish products, smoking was the most reported stabilisation category (73%, n = 38). In this category, only for 14% of notifications (n = 7), no specification was made.

3.4.5.3. Storage temperature, best before date and microbiological threshold values. The majority of food products for which information on the storage temperature was available (specifications made in 66% of the 226 *Lm* notifications, n = 149) were chilled. For about 70% of the chilled products of each category, additional quantitative information on *Lm* contamination was available. Among these, 57–61% exceeded the microbiological limit value of 100 CFU/g. Only in five products, sampling was carried out after the best before date whereas 154 (68%) contaminated products were sampled before. For 73% (n = 113) of the products sampled before the best before date, quantitative information on *Lm* contamination was available. In 83% out of these cases (n = 94), the microbiological threshold value of 100 CFU/g was exceeded. For those products that were sampled after the best before date, quantitative information was available for three of the five samples. Two exceeded the limit, whereas one was below.

4. Discussion

4.1. RASFF notifications - Main players and development over time

Four countries were the main players in releasing PMF notifications involving Germany between 2004 and 2015: Germany, Denmark, France and Italy. Germany, France and Italy are among the most populated EU countries and also among those with the highest gross domestic product at market prices (German Federal Statistical Office, 2017). This could be part of the explanation of their RASFF notification activities, but trade relations and also country-specific differences in awareness and resulting efforts in the national food surveillance systems might play an important role. Furthermore, the market shares of RTE products for the different countries are likely to influence their notification rates.

PMF notifications affecting Germany have been increasing since 2011. Interestingly, the proportion of *Lm* notifications remained relatively stable. Nevertheless, due to the increasing number of PMF notifications, also more *Lm* notifications affecting Germany were published from 2011 to 2015. Although possibly caused by a real increase in microbial food contamination, in times of overall advances in microbiological food surveillance, this trend could also be a result of an increase in awareness, efforts and reporting.

4.2. Notifying country versus country of origin - Discrepancies and possible solutions

Independent of the country affected, France and Italy were the notifying countries in about half of the *Lm* notifications between 2004 and 2015. France and Germany almost covered the total number of notifications dealing with products of French and German origin, respectively (France notifying in 61 of 77 notifications, Germany notifying in 22 of 25 notifications; Table 1). For several other countries, a greater imbalance was observed. Italy and Poland were country of origin in more notifications than notifying country with respect to all notifications on products of Italian and Polish origin, respectively (Italy notifying in 10 of 39 notifications, Poland notifying in 8 of 24 notifications; Table 1). A satisfactory RASFF activity is reflected by the fact that a country does not only recognise contamination in products produced in the respective country but also in products shipped from other EU/EEA countries. If country of origin and notifying country match to a large extent, hazards could possibly be published at an earlier time of the product's life-cycle, thereby more effectively preventing the spread of risk-bearing foodstuffs.

Out of all *Lm* notifications affecting Germany, 2001 to 2015, only in 33% of notifications, Germany itself was the notifying country. In 30%, the notification was released by France and in 9% of cases by Italy. This could be related to the fact that products mentioned in the corresponding *Lm* notifications in 34% of cases had their origin in France and in 17% of cases in Italy. In contrast to that, only a small proportion of

notifications (11%) were related to products of German origin. While *Lm* notifications for products of French origin were mainly reported by France; Italian, Polish and German products were mainly notified by Germany. This probably explains the divergence between the proportion of notifications notified by Germany and the proportion of affected products of German origin. Again, this shows the interconnectedness of EU countries in means of food safety issues through trade and underlines the value of a comprehensive early-warning system like the RASFF. However, it also demonstrates that notification activities within the EU could benefit from harmonisation. According to Regulation (EC) 882/2004, it is up to the EU member states to ensure that official controls are carried out regularly, on a risk basis and with appropriate frequency (European Commission, 2004). Hence, surveillance activity is regulated by individual country legislation and thereby subject to variations. To achieve higher food safety standards, harmonised and sometimes enhanced national surveillance activity might be needed.

A further limitation to harmonisation lies in the Regulation (EC) 2073/2005 on microbiological criteria for foodstuffs. This regulation defines criteria for *Lm* in RTE food products (European Commission, 2005). These criteria differentiate between food products that are unable or able to support growth of *Lm*. If the food product is unable to support growth of *Lm*, bacterial concentration in products placed on the market must not exceed 100 CFU/g during shelf life. For food products that are able to support growth, this applies as well. However, in this latter case, an additional, stricter criterion is also valid, namely absence of *Lm* in 25 g before the food product has left the final control of the food business operator who has produced it. These specifications do not leave room for interpretation. However, there is no strict consensus on the classification of foodstuffs as “able to support growth” or “unable to support growth”. Either a pH < 4.4 and $a_w < 0.92$ or a combination of pH < 5.0, $a_w < 0.94$ and NaCl $> 16\%$ are generally considered as intrinsic food conditions that do not support growth (Buchanan, Gorris, Hayman, Jackson, & Whiting, 2017). For all other categories, however, scientific proof is needed to verify that the microbial limit of 100 CFU/g will not be exceeded during shelf life (European Commission, 2005). Unfortunately, limited data exist on growth rates of *Lm* in different foodstuffs, making it sometimes difficult to provide evidence for a very specific product without performing a time-consuming shelf life study (U. S. Food and Drug Administration, 2014). Therefore, in the case of sufficient justification, data for comparable foodstuffs can be used as a reference. In this way, however, a little scope for interpretation may be introduced for the decision whether a very specific product does or does not enable *Lm* growth and as a consequence whether contamination is considered as hazardous or not. Accordingly, different country specific interpretations may lead to different RASFF notification activities. The most comprehensive way to deal with this limitation could be an overall zero tolerance limit in RTE food products like already active in the USA (Chen, Ross, Scott, & Gombas, 2003). However, this strategy has its weaknesses as well. Besides the considerable expense associated with its implementation, its advantages over the EU-wide 100 CFU/g limit are controversial (Chen et al., 2003; Tompkin, 2002). In a microbial risk assessment performed by Chen and colleagues for example, the risk reduction potential of a non-zero strategy outperformed that of the zero tolerance one (Chen et al., 2003). Overall, a compromise has to be pursued to further standardise recognition of *Lm*-related food safety issues and to achieve the highest possible reduction of foodborne listeriosis infections.

4.3. Notification basis - “The sooner, the better.”

Most of *Lm* notifications affecting Germany, 2001 to 2015, were based on official controls on the market, followed by company's own checks. In our dataset, notifications following company's own checks did not appear before 2005. This is in line with implementation of Regulation (EC) 2073/2005 in 2005 as a consequence of the White Paper on Food Safety published by the European Commission in 2000 (European Commission, 2000, 2005). As a reaction to several food safety crises in the 1990s, this White Paper aimed to revolutionise food safety. One key vision was a complete recast of the different control requirements in order to ensure that all links in the food production chain are covered by effective controls. As a result, more responsibility was assigned to food business operators (European Commission, 2005). Among other things, they were obliged to appropriately test against the defined microbiological criteria, inducing the sudden appearance of company's own checks in 2005. Since then, they showed an increasing trend. This is very laudable as, in contrast to official controls on the market, company's own checks usually detect a microbiological contamination before the product is sold on the market or even earlier in the product's life-cycle. As a result, the risk of a consumer to eat foodstuff concerned is far lower as when a control happens when the product is already on the market. In many cases where quantitative information was available on *Lm* contamination in a RASFF notification, the threshold value of 100 CFU/g (European Commission, 2005) was exceeded even in chilled products and also before the best before date showing the importance and also the value of the early-warning system. If timely intervention measures like product recalls are taken, public health risks can efficiently be decreased. While for example in Germany, Italy and Austria official controls on the market prevail, the majority of notifications from France are a result of company's own checks. Overall, a promising trend can be observed as, since 2013, the number of company's own check based notifications per year exceeds that based on official controls on the market. Hopefully, this development will help in the long run to ameliorate the quality of food products that enter the market and thus help to prevent foodborne infections.

4.4. RASFF notifications types for risk assessment

A large majority of *Lm* notifications affecting Germany between 2001 and 2015 were alert notifications where rapid action by other RASFF members was required. This highlights the importance of an EU-wide, fast communication system like the RASFF to quickly and comprehensively manage the risk posed by a contaminated food product. However, information notifications also play an important role. “Information notifications for follow-up” for example are similar to alert notifications. The only difference is the status of reaction of all countries involved at the time of publication of the RASFF notification. In their case, the report was released after measures had been taken. Even though not requiring rapid action, they report about a risk for the consumer that occurred and concerned multiple EU countries, hence providing valuable information on the overall prevalence of *Lm* contamination. The same is true for “information notification for attention”. This type of notification is probably the weakest concerning the need for rapid action as distribution of the concerned product was restricted to one country, the product is no longer on the market or has not even been placed on the market. Still, a potential risk for consumers was identified. In conclusion, although only alert notifications might use the entire power of the EU-wide communication tool RASFF, also information notifications essentially contribute to a comprehensive risk assessment and are therefore an indispensable and valuable part of RASFF notification activities.

4.5. Identification of risk factors for Lm contamination

The majority of *Lm* notifications affecting Germany, 2001 to 2015, concerned milk and milk products followed by fish and fish products and meat and meat products other than poultry. This goes along with the fact that products from these three categories were also reported as causative in 59% of the foodborne outbreaks caused by *Lm* in the EU/EEA between 2008 and 2015 (EFSA Panel on Biological Hazards, 2017). Probably as a result from the risk-based character of the RASFF notification system, these categories were associated with most of the *Lm* notifications.

In the category of milk and milk products, the majority of notifications were related to products from France, whereas a large share of concerned fish and fish products originated from Poland. This is also reflected in the involvement of specific companies from the two countries into notifications. Regarding the market shares of France and Poland for products of these two categories, this is not surprising. While France had the highest number of companies involved in notifications for milk and milk products, two Polish companies were involved in 17 notifications in the category of fish and fish products. Together with the fact that some companies were associated with notifications over several years, this proves the stability of *Lm* contamination in various production plants. Undetected sources of the bacterium may persist for years and can consequently lead to repeated re-contamination of foodstuffs (Carpentier & Cerf, 2011). Although identification and notification of contaminated food products are relevant to take adequate countermeasures and to protect consumer health, this approach will be not sufficient to improve food safety in the long run. Reporting alone will not suffice to contain *Lm* contamination of foodstuffs and thus prevent human infections. In order to really address the problem, its root has to be addressed. On the one hand, this could be achieved through improved hygiene measures, for example using the seek-and-destroy strategy (Butts, 2003; Malley, Butts, & Wiedmann, 2015). Persistent *Lm* strains in food processing plants have been identified as the most common post processing contaminants (Tompkin, 2002). Hence, the seek-and-destroy strategy aims to identify harbourage sites and niche locations, where bacterial strains withstand cleaning and sanitation measures. Shortly, it combines different disassembly stages of equipment with repeated sampling and sanitation measures (usually flood or heat) until proven elimination of contamination (for detailed description see (Malley et al., 2015)). On the other hand, also overall more strict internal controls could help to tackle the problem of *Lm* contamination at its root. In this context, the increasing quality of company's own checks is a first step in the right direction.

Cross-contamination during the processing is one of the most important reasons for *Lm* contamination of foodstuffs. Hence, in order to identify specific entry routes of *Lm* into the food chain, additional information on product properties and additives would be desirable. Unfortunately, information for example about packaging or food processing was only provided for a minority of RASFF notifications. For instance, the packaging category of a product could be interesting to know, because different atmospheres (vacuum, modified, normal etc.) might selectively support bacterial growth thereby giving an advantage to *Lm* proliferation (Tsigarida, Skandamis, & Nychas, 2000). Furthermore, the slicing category could be a valuable type of information, as instruments used in this processing step are suspected to be a common source of contamination (Lin et al., 2006). In the future, it would thus be useful to provide as much information concerning a product in the RASFF notification as possible to establish a more comprehensive database and to better identify risk factors.

A first promising step towards simplification and harmonisation of reporting was made through introduction of the interactive RASFF (iRASFF) in 2011 which replaced Microsoft Word-templates for notification by an online IT application (European Commission, 2015). Dropdown menus are available in all official EU languages for the key data (product category etc.) that can later be found in the RASFF portal. Companies or hazards are stored in a database, but it is still possible to integrate a free-text description. However, concerning additional metadata, no uniform rule exists. While for example information on storage temperature can be selected in a dropdown menu, no such possibility exists concerning packaging type, which is a free-text field. In general, all of these metadata fields are not mandatory which is why information might not be available in some notifications. On the one hand, this appears necessary as in the case of missing information, the notification would otherwise not be publishable. However, an explicit "not specified" option in a mandatory field could help to handle this problem and on the other hand promote the provision of crucial additional information to improve risk assessment and thus European food safety.

5. Conclusion

The number of RASFF *Lm* notifications in food products distributed in Germany and the number of human listeriosis cases reported have been increasing in parallel from 2011 on. As a result of the fact that human listeriosis cases can usually not be linked to a certain RASFF notification, no proven causal relationship can be attested. Nevertheless, the common trend in both notifications is alarming. Due to the fact that food contamination is not a "one country" problem, communication between countries maintaining close trade relations, a European surveillance system like the RASFF and hence timely reaction to food safety issues are of crucial importance. Furthermore, EU-wide identification and communication of potential consumer risks provide a major contribution to risk assessment. However, improvements in the current system and better exchanging and linking between food safety and public health authorities will be indispensable in order to further promote this development.

First, a key performance indicator could be a largest possible match between country of origin and notifying country in RASFF notifications, thus enabling risk communication and interventional measures at a very early time point of the product's life cycle. This can be achieved by further extension of company's own checks. However, also overall enhanced national surveillance activities would be beneficial. Secondly, although a powerful tool for communication, reporting alone via the RASFF system will not be sufficient to contain *Lm* contamination of foodstuffs. To be able to address the root of the problem, more metadata should be made available together with the RASFF notifications to allow identification of risk factors. For that purpose, adjustment of the reporting system setting more value on information about product properties would be desirable.

Funding

This work was supported by a grant of the Federal Ministry of Health (GE 2016 03 26) in the framework of the German Research Platform for Zoonoses.

Declaration of interests

None.

Acknowledgement

We would like to express special thanks to the Department 1 (Food, Feed and Commodities) from the German Federal Office of Consumer Protection and Food Safety (BVL) for providing us with RASFF notifications from their database and accompanying documents.

References

Buchanan, R. L., Gorris, L. G. M., Hayman, M. M., Jackson, T. C., & Whiting, R. C. (2017). A review of Listeria monocytogenes: An update on outbreaks, virulence, dose-response, ecology, and risk assessments. *Food Control, 75*, 1–13.

Butts, J. (2003). Seek & destroy: Identifying and controlling Listeria monocytogenes growth niches. *Food Safety Magazine, 58*, 24–29.

Carpentier, B., & Cerf, O. (2011). Persistence of Listeria monocytogenes in food industry equipment and premises. *International Journal of Food Microbiology, 145*(1), 1–8.

Chen, Y., Ross, W. H., Scott, V. N., & Gombas, D. E. (2003). Listeria monocytogenes: Low levels equal low risk. *Journal of Food Protection, 66*(4), 570–577.

Commission Regulation (EC) (2005). *No 2073/2005 of 15 November 2005 on microbiological criteria for foodstuffs*.

Commission Regulation (EU) (2011). *No 16/2011 of 10 January 2011 laying down implementing measures for the Rapid alert system for food and feed*.

ECDC (2016). Annual epidemiological report 2016 – listeriosis. Retrieved from https://ecdc.europa.eu/en/publications-data/listeriosis-annual-epidemiological-report-2016-2014-data, Accessed date: 13 December 2017.

EFSA Panel on Biological Hazards (2017). DRAFT SCIENTIFIC OPINION: Listeria monocytogenes contamination of ready-to-eat foods and the risk for human health in the EU. Retrieved from https://www.efsa.europa.eu/sites/default/files/engage/170724-0.pdf, Accessed date: 13 December 2017.

EFSA, & ECDC (2017). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. *EFSA Journal, 15*(12).

European Commission (2000). *White Paper on food safety*.

European Commission (2015). *B. How to use RASFF WINDOW 2.* IRASFF - Europa EU. Retrieved from https://ec.europa.eu/food/sites/food/files/safety/docs/rasff_reg-guid_sops_wi-7-1_en.pdf, Accessed date: 12 December 2017.

European Commission (2017a). *RASFF - food and feed safety alerts*. Retrieved from https://ec.europa.eu/food/safety/rasff_en, Accessed date: 12 December 2017.

European Commission (2017b). *RASFF portal.* Retrieved from https://webgate.ec.europa.eu/rasff-window/portal/?event=SearchForm&cleanSearch=1, Accessed date: 14 December 2017.

Farber, J., & Peterkin, P. (1991). Listeria monocytogenes, a food-borne pathogen. *Microbiological Reviews, 55*(3), 476–511.

German Federal Statistical Office (2017). *EU comparison 2017: Germany and the other member states.* Retrieved from https://www.destatis.de/Europa/EN/Country/Comparison/GER_EU_Compared.html;jsessionid=BD3E120921532E6F4E49E85A040E1042.InternetLive1?nn=478000, Accessed date: 17 January 2018.

Lin, C. M., Takeuchi, K., Zhang, L., Dohm, C. B., Meyer, J. D., Hall, P. A., et al. (2006). Cross-contamination between processing equipment and deli meats by Listeria monocytogenes. *Journal of Food Protection, 69*(1), 71–79.

Lungu, B., Ricke, S., & Johnson, M. (2009). Growth, survival, proliferation and pathogenesis of Listeria monocytogenes under low oxygen or anaerobic conditions: A review. *Anaerobe, 15*(1), 7–17.

Malley, T. J., Butts, J., & Wiedmann, M. (2015). Seek and destroy process: Listeria monocytogenes process controls in the ready-to-eat meat and poultry industry. *Journal of Food Protection, 78*(2), 436–445.

Milillo, S. R., Friedly, E. C., Saldivar, J. C., Muthaiyan, A., O'Bryan, C., Crandall, P. G., et al. (2012). A review of the ecology, genomics, and stress response of Listeria innocua and Listeria monocytogenes. *Critical Reviews in Food Science and Nutrition, 52*(8), 712–725.

Regulation (EC) (2002). *No 178/2002 OF the european parliament and OF the council of 28 january 2002*.

Regulation (EC) (2004). *No 882/2004 OF the european parliament and OF the council of 29 April 2004*.

Schlech, W. F., & Acheson, D. (2000). Foodborne listeriosis. *Clinical Infectious Diseases, 31*(3), 770–775.

SURVSTAT@RKI 2.0. SURVSTAT@RKI 2.0. Retrieved from https://survstat.rki.de/. Accessed 9 January 2018..

Swaminathan, B., & Gerner-Smidt, P. (2007). The epidemiology of human listeriosis. *Microbes and Infection, 9*(10), 1236–1243.

Tompkin, R. (2002). Control of Listeria monocytogenes in the food-processing environment. *Journal of Food Protection, 65*(4), 709–725.

Tsigarida, E., Skandamis, P., & Nychas, G. J. E. (2000). Behaviour of Listeria monocytogenes and autochthonous flora on meat stored under aerobic, vacuum and modified atmosphere packaging conditions with or without the presence of oregano essential oil at 5 °C. *Journal of Applied Microbiology, 89*(6), 901–909.

U. S. Food and Drug Administration (2014). *Appendix 8: Growth of Listeria monocytogenes in foods.* Retrieved from https://www.fda.gov/downloads/Food/ScienceResearch/ResearchAreas/RiskAssessmentSafetyAssessment/UCM197321.pdf, Accessed date: 28 February 2018.