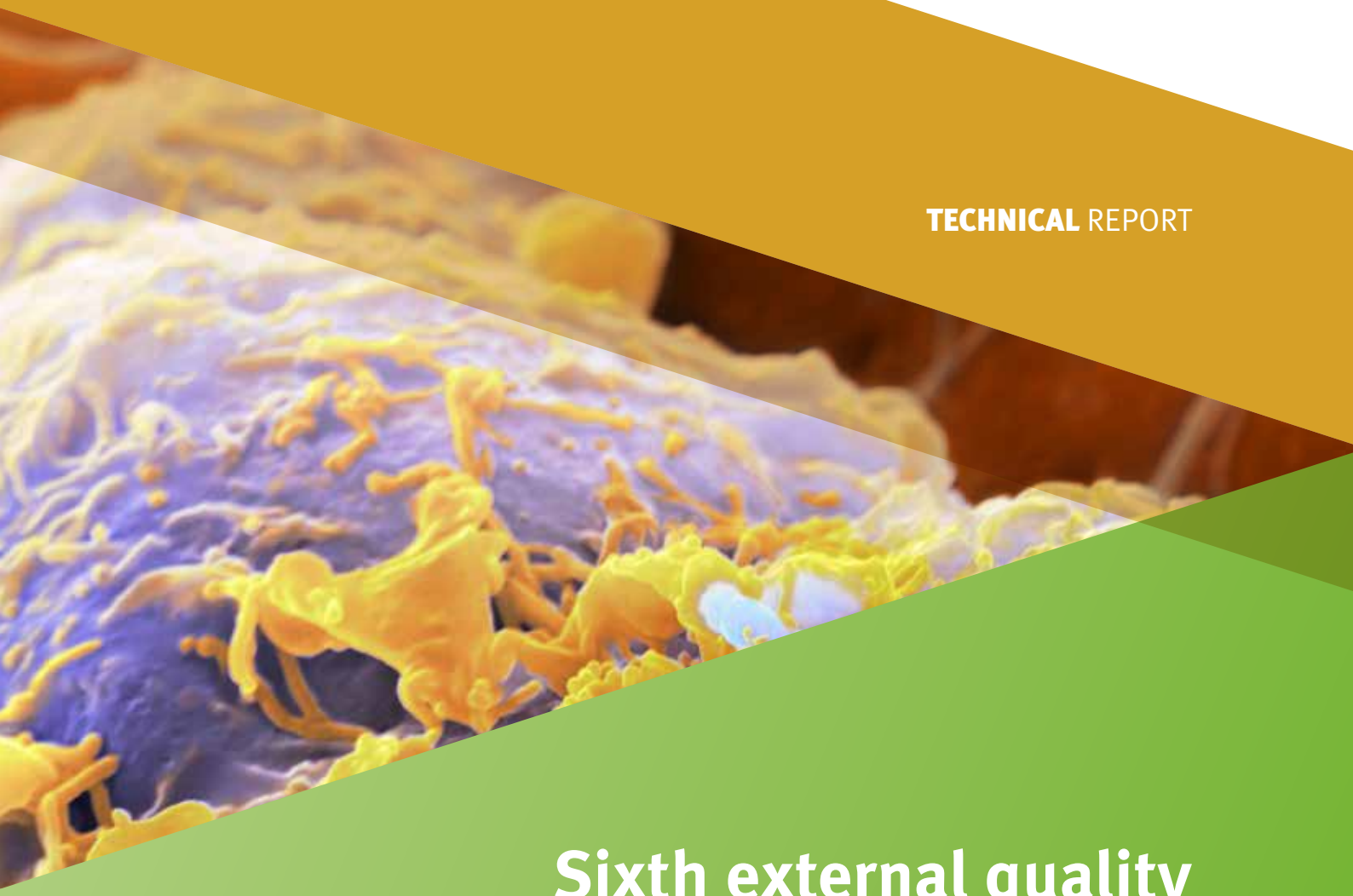
ecdc
EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL



TECHNICAL REPORT

Sixth external quality assessment scheme for *Listeria monocytogenes* typing

www.ecdc.europa.eu

ECDC TECHNICAL REPORT

Sixth external quality assessment scheme for *Listeria monocytogenes* typing

This report was commissioned by the European Centre for Disease Prevention and Control (ECDC), coordinated by Taina Niskanen (ECDC Food- and Waterborne Diseases and Zoonoses Programme) and produced by Susanne Schjørring, Gitte Sørensen, Kristoffer Kiil, Louise Dahl, Malgorzata Ligowska-Marzeta and Eva Møller Nielsen of the Section for Foodborne Infections at Statens Serum Institut, Copenhagen, Denmark.

Suggested citation: European Centre for Disease Prevention and Control. Sixth external quality assessment scheme for *Listeria monocytogenes* typing. Stockholm: ECDC, 2019.

Stockholm, August 2019

ISBN: 978-92-9498-376-3
DOI: 10.2900/990270
Catalogue number: TQ-01-19-701-EN-N

Contents

Figures

Tables

Annexes

Abbreviations

BN	BioNumerics
cgMLST	Core genome multilocus sequence type
wgMLST	Whole genome multilocus sequence type
EFSA	European Food Safety Authority
EQA	External quality assessment
EU/EEA	European Union/European Economic Area
EURL	European Union Reference Laboratory
FWD	Food- and waterborne diseases and zoonoses
FWD-Net	Food- and Waterborne Diseases and Zoonoses Network
PFGE	Pulsed-field gel electrophoresis
QC	Qualitative control
SNP	Single nucleotide polymorphism
SSI	Statens Serum Institut
ST	Sequence type
TESSy	The European Surveillance System
WGS	Whole-genome sequencing

Executive summary

This report presents the results of the sixth round of the external quality assessment (EQA-6) scheme for *Listeria monocytogenes (L. monocytogenes)* typing organised for laboratories providing data to the Food- and Waterborne Diseases and Zoonoses Network (FWD-Net) managed by ECDC. Since 2012, the Section for Foodborne Infections at Statens Serum Institut (SSI) in Denmark has arranged the EQA under a framework contract with ECDC. EQA-6 contains serotyping and molecular typing-based cluster analysis.

Human listeriosis is a relatively rare but serious zoonotic disease with an EU notification rate of 0.48 cases per 100 000 population in 2017 [3]. The number of human listeriosis cases in the EU has increased since 2008, with a peak in the annual number in 2016 of 2 509.

Since 2007, ECDC's FWD Programme has been responsible for the EU-wide surveillance of listeriosis, including facilitating, detecting and investigating foodborne outbreaks. Surveillance data, including basic typing parameters for the isolated pathogen, are reported by Member States to The European Surveillance System (TESSy), including molecular typing data. This molecular surveillance system relies on the capacity of laboratories providing data to FWD-Net to produce comparable typing results. In order to ensure the EQA is linked to the development of surveillance methods used by public health national reference laboratories in Europe, EQA-6 contains a molecular typing-based cluster analysis using either pulsed-field gel electrophoresis (PFGE) and/or whole-genome sequencing (WGS)-derived data.

The objectives of the EQA are to assess the quality and comparability of typing data reported by public health national reference laboratories participating in FWD-Net. Test isolates for the EQA were selected to cover isolates currently relevant to public health in Europe and represent a broad range of clinically relevant types for invasive listeriosis. Two separate sets of 11 test isolates were selected for serotyping and molecular typing-based cluster analysis. Eighteen laboratories signed up and 17 completed the exercise, representing a decrease in participation of 15% from the previous assessment (EQA-5) and an even larger decrease from EQA-4 of 26%. The decrease in the number of participants may have been caused by adding WGS or removing PFGE as an independent part. The majority (65%) of participants completed the full EQA scheme. In total, 13 (77%) participants participated in the serotyping part and 15 (88%) in the molecular typing-based cluster analysis.

Molecular serotyping results were provided by all 13 (100%) participants. Only four participants performed both conventional and molecular serotyping. The performance of conventional serotyping was highest, with 100% correct results. A total of 77% of the respective participants correctly serotyped all test isolates by molecular method. Since the first EQA in 2012, a trend towards substituting conventional serotyping with molecular serotyping has been observed as none of the participants exclusively performed conventional serotyping in EQA-6.

Of the 17 laboratories participating in EQA-6, 15 (88%) performed molecular typing-based cluster analysis using any method. The idea of the cluster analysis part of the EQA was to assess the public health national reference laboratories' ability to identify a cluster of genetically closely related isolates given that a multitude of different laboratory methods and analytical methods are used as the primary cluster detection approach in Member States. This part of the EQA was atypical in the sense that the aim was to assess the participants' ability to reach the correct conclusion, i.e. to correctly categorise the cluster test isolates, not to follow a specific procedure.

The cluster of closely related isolates contained seven ST399 isolates that could be identified by PFGE- (by either *Apa*I or *Asc*I profiles). The correct WGS cluster contained five ST399 isolates. The expected cluster was based on a predefined categorisation by the organiser. Seven laboratories used PFGE for cluster analysis and four of them also reported cluster analysis based on WGS data. Only one laboratory did not identify the correct cluster using PFGE.

Twelve laboratories performed cluster analysis using WGS-derived data. Performance was high, with 11/12 (92%) of the participants correctly identifying the cluster of closely related isolates. An allele-based method was preferred since 75% (9/12) used core genome multilocus sequence type (cgMLST), compared with 25% (3/12) using single nucleotide polymorphism (SNP). Only one laboratory did not identify the correct cluster, but included additional two isolates. The results were based on a SNP analysis.

In the EQA, participants were free to choose their preferred analytical method for the WGS-based cluster identification. The conclusion from EQA-5 showed cgMLST to have higher consistency compared to SNP analysis. The conclusion was not as obvious this year since only three SNP analyses were reported in EQA-6 compared with six in EQA-5. Comparing cgMLST result of Illumina and Ion Torrent data revealed differences, which suggest lower inter-laboratory comparability across sequencing platforms.

1 Introduction

1.1 Background

ECDC is an EU agency with a mandate to operate dedicated surveillance networks. The mission of ECDC is to identify, assess and communicate current and emerging threats to human health from communicable diseases. ECDC fosters the development of sufficient capacity within the European Community's network for diagnosis, detection, identification and characterisation of infectious agents that may threaten public health. ECDC maintains and extends such cooperation and supports the implementation of quality assurance schemes [1].

External quality assessments (EQAs) are an essential part of quality management and use an external evaluator to assess the performance of participating laboratories on test samples supplied specifically for the purpose.

ECDC's disease networks organise a series of EQAs for EU/European Economic Area (EEA) countries. The aim of EQAs is to identify areas of improvement in the laboratory diagnostic capacities relevant to epidemiological surveillance of communicable diseases as in Decision No 1082/2013/EU [2] and ensure the reliability and comparability of results generated by laboratories across all EU/EEA countries.

The main purposes of EQA schemes are to:

- assess the general standard of performance ('state of the art')
- assess the effects of analytical procedures (method principle, instruments, reagents, calibration)
- evaluate individual laboratory performance
- identify and justify of problem areas
- provide continuing education; and
- identify needs for training activities.

Since 2012, the Section for Foodborne Infections at Statens Serum Institut (SSI) in Denmark has been the EQA provider for the three lots covering *Salmonella enterica* ssp. *enterica*, Shiga toxin/verocytotoxin-producing *Escherichia coli (E. coli)* (STEC/VTEC) and *L. monocytogenes* typing. In 2016, SSI was also granted the new round of tenders (2017–2020) for all three lots. For lot 3 (*L. monocytogenes*) from 2017, the EQA scheme no longer covers assessment of the PFGE quality. However, it still covers serotyping and includes a new part for cluster identification of *L. monocytogenes.* The present report presents the results of the sixth EQA scheme (*Listeria* EQA-6).

1.2 Surveillance of listeriosis

Human listeriosis is a relatively rare but serious foodborne disease, with high rates of morbidity, hospitalisation and mortality in vulnerable populations. The number of human listeriosis cases in the EU has increased since 2008, with a peak in the annual number of cases in 2016 of 2 509. In the EU, 2 480 confirmed human listeriosis cases were reported in 2017, corresponding to a notification rate of 0.48 cases per 100 000 population, which is at a similar level as that observed in 2016 [3].

ECDC's FWD Programme is responsible for EU-wide surveillance of listeriosis and facilitating detecting and investigating foodborne outbreaks since 2007. One of the key objectives for the FWD programme is to improve and harmonise the surveillance system in the EU to increase scientific knowledge of aetiology, risk factors and burden of food- and waterborne diseases and zoonoses. Surveillance data, including some basic typing parameters for the isolated pathogen, are reported by Member States to TESSy. In addition to the basic characterisation of the pathogens isolated from infections, there is a public health value to use more discriminatory typing techniques in the surveillance of foodborne infections. Since 2012, ECDC has enhanced surveillance incorporating molecular typing data. Three selected FWD pathogens were included: *Salmonella enterica* ssp. *enterica*, *L. monocytogenes* and STEC/VTEC. The overall aims of integrating molecular typing into EU level surveillance are to:

- foster rapid detection of dispersed international clusters/outbreaks
- facilitate the detection and investigation of transmission chains and relatedness of isolates across Member States and contribute to global investigations
- detect emergence of new evolving pathogenic isolates
- support investigations to trace back the source of an outbreak and identify new risk factors; and
- aid the study of a particular pathogen's characteristics and behaviour in a community of hosts.

Molecular typing-enhanced surveillance gives Member State users access to EU-wide molecular typing data for the included pathogens. It also gives users the opportunity to perform cluster searches and cross-sector comparability of EU-level data to determine whether isolates characterised by molecular typing at the national level are part of a multinational cluster that may require cross-border response collaboration.

The EQA schemes target public health national reference laboratories already expected to be performing molecular surveillance at the national level.

1.3 Objectives

1.3.1 Serotyping

The EQA-6 scheme assessed serotype determination by either conventional antigen-based typing of somatic 'O' antigens and flagellar 'H' antigens and/or PCR-based molecular serotyping.

1.3.2 Molecular typing-based cluster analysis

The objective of *L. monocytogenes* EQA-6 was to assess the ability to detect a cluster of closely related isolates. Laboratories could perform analysis using PFGE and/or derived data from WGS.

2 Study design

2.1 Organisation

Listeria EQA-6 was funded by ECDC and arranged by SSI following ISO/IEC 17043:2010 [4]. EQA-6 included serotyping and molecular typing-based cluster analysis and was carried out from June–December 2018.

Invitations were emailed to ECDC contact points in FWD-Net (26 countries nominated laboratories to participate on the EQA rounds from 2017–2020) by 31 May 2018, with a deadline to respond by 8 June 2018. In addition, invitations were sent to EU candidate and potential candidate countries Turkey and Kosovo[i], which signed up for the *Salmonella* EQA rounds from 2017-2020. Each laboratory was asked to fill in the reason for participating or not participating.

Eighteen public health national reference laboratories in EU/EEA and EU candidate countries accepted the invitation to participate and 17 submitted results (Annex 1). In Annex 2, details of participation in EQA-5 and EQA-6 are listed to give an overview of the trend in the number of participants. The EQA test isolates were sent to participants from 26–29 June 2018. Participants were asked to submit their results by 3 September 2018 using the online form (Annex 14). If WGS was performed, submission of the raw reads (FASTQ files) to http://sikkerftp.ssi.dk was also requested. The EQA submission protocol was distributed by email and available on the online site.

2.2 Selection of test isolates

Twenty-two *L. monocytogenes* test isolates were selected to fulfil the following criteria:

- cover a broad range of the common clinically relevant types for invasive listeriosis
- include closely related isolates; and
- remain stable during the preliminary test period at the organising laboratory.

Forty-four candidate isolates were analysed by the methods used in the EQA before and after re-culturing. All candidate isolates remained stable using these methods and the final selection of 22 test isolates, including technical duplicates (same isolate culture twice), was made. Eleven test isolates for serotyping were selected to cover different serotypes relevant for the current epidemiological situation in Europe. Isolates within serotypes 1/2a/IIa, 1/2b/IIb, 1/2c/IIc, and 4b/IVb were selected. Among the serotyping test isolates, the three repeat isolates from EQA-1 to 5 were included to evaluate the performance development of the participants. Two sets of technical duplicates were also included this year (Annex 4). Eleven test isolates for cluster analysis were selected to include isolates with different or varying relatedness isolates and different multilocus sequence types (ST 6, 11, 14 and 399). Using either PFGE or WGS-derived data, the cluster of closely related isolates consisted of seven or five *L. monocytogenes* ST399 isolates (one technical duplicate) respectively. The characteristics of all the *L. monocytogenes* test isolates are listed as Sero/REF in Annexes 4–11.

2.3 Carriage of isolates

At the end of June 2018, all test isolates were blinded and shipped from 26–29 June. The protocol for the EQA exercise and a letter stating the unique isolate IDs were included in the packages and distributed individually to the participants by email on 26 June 2018 as an extra precaution. Thirteen participants received their dispatched isolates within one day and five within three days after shipment. The parcels were shipped from SSI labelled as UN 3373 Biological Substance. No participants reported damage to the shipment or errors in the unique isolates IDs.

On 5 July 2018, instructions to the submission of results procedure were emailed to the participants. This included the links to the online uploading site and online submission form.

2.4 Testing

In the serotyping part, 11 *L. monocytogenes* isolates were tested to assess the participants' ability to obtain the correct serotype. Participants could choose to perform conventional serological methods and/or PCR-based molecular serotyping (multiplex PCR according to the protocol suggested by Doumith et al. [5]). The serotypes were submitted in the online form.

[i] This designation is without prejudice to positions on status, and in line with UNSCR 1244/99 and the ICJ Opinion on the Kosovo Declaration of Independence.

In the cluster analysis part, participants could choose to perform the laboratory part using PFGE (*Apa*I and *Asc*I profiles) and/or WGS-derived data. The participants were instructed to report the IDs of the isolates included in the cluster of closely related isolates by method. If PFGE analysis was conducted, the participant reported the total number of bands and number of shared bands with a selected cluster representative isolate for both *Apa*I and *Asc*I.

Laboratories performing WGS could use their own analysis pipeline for cluster analysis, e.g. single nucleotide polymorphism analysis (SNP-based) or whole genome multilocus sequence typing (wgMLST)/cgMLST (allele-based) and was asked to submit the isolates identified as cluster of closely related isolates based on the analysis used. The laboratories could report results from up to three analyses (one main and one to two additional), but the detected cluster had to be based on results from the main analysis. Laboratories reported SNP distance or allelic differences between a selected cluster isolate and each test isolate and uploaded the raw reads (FASTQ files).

2.5 Data analysis

As participating laboratories submitted their results, serotyping and cluster analysis results, as well as the participants' uploaded raw reads, were imported to a dedicated *Listeria* EQA-6 BioNumerics (BN) database. The EQA provider accepted two participants to submit data 14 days after deadline.

Serotyping results were evaluated according to the percentage of correct results, generating a score from 0-100%. Molecular typing-based cluster analysis was evaluated according to correct or incorrect identification of the expected cluster of closely related isolates based on a pre-defined categorisation by the organiser. The EQA provider's PFGE results were based on combined *Apa*I and *Asc*I profiles. Cluster analysis based on WGS-derived data was derived on allele-based (cgMLST [6] and SNP analysis (NASP, [7]). The correct number of closely related *L. monocytogenes* isolates could be identified by both PFGE- and WGS-derived data including 7 or 5 respectively. The PFGE cluster contained seven ST399 isolates: REF1, REF3, REF4, REF6, REF7, REF8 and REF11 (REF1 and REF6 were technical duplicates). The WGS cluster contained five ST399 isolates: REF1, REF3, REF6, REF8 and REF11 (REF1 and REF6 were technical duplicates). The EQA provider found at most two allele differences or four SNPs between any two isolates in the cluster. The rest of the cluster test isolates were an additional three ST399s, one ST6, one ST11 and one ST14.

Individual evaluation reports were distributed to participants in December 2018 and certificates of attendance in March 2019. If WGS data were used, the evaluation report included a quality assessment made by the EQA provider's in-house quality control pipeline (e.g. coverage, N50, sequence length and number of contigs). The evaluation report did not include an evaluation based on quality thresholds.

3 Results

3.1 Participation

Laboratories could participate either in the full EQA scheme or one part only (serotyping or molecular typing-based cluster analysis). Of the 18 participants who signed up, 17 completed and submitted their results. Two laboratories gave a reason for not participating: a combination of lack of capacity and not performing WGS and low PFGE quality, lack of laboratory capacity and only a few cases of listeriosis per year. The majority of the participants (65%, 11/17) completed the full EQA scheme. In total, 13 (76%) participants participated in serotyping and 15 (88%) in cluster analysis. Conventional serotyping results were provided by 31% (4/13) participants and all four also performed molecular serotyping. Molecular serotyping was provided by 13 (76%) participants. Of the 13 participants who completed the serotyping part, eight reported the reason for participating: laboratory policy to enhance the typing quality in a combination of accreditation needs, institute policy and/or national policy (Annex 3).

Most participants (80%: 12/15) reported cluster analysis using WGS-derived data, while three (20%) reported only using PFGE data. Four participants (27%) submitted cluster data based on both PFGE and WGS (Table 1). Of the 15 laboratories that completed the cluster analysis, 13 reported the reason for participating: laboratory policy to enhance the typing quality in a combination of accreditation needs, institute policy and/or national policy (Annex 3).

Table 1. Number and percentage of laboratories submitting results for each method

	Serotyping				Cluster analysis			
	Conventional only	Molecular only	Both	Total	PFGE-only	WGS-only	Both	Total
Number of participants	0	9	4	13	3	8	4	15
Percentage of participants	0%	69%	31%	76%*	20%	53%	27%	88%*

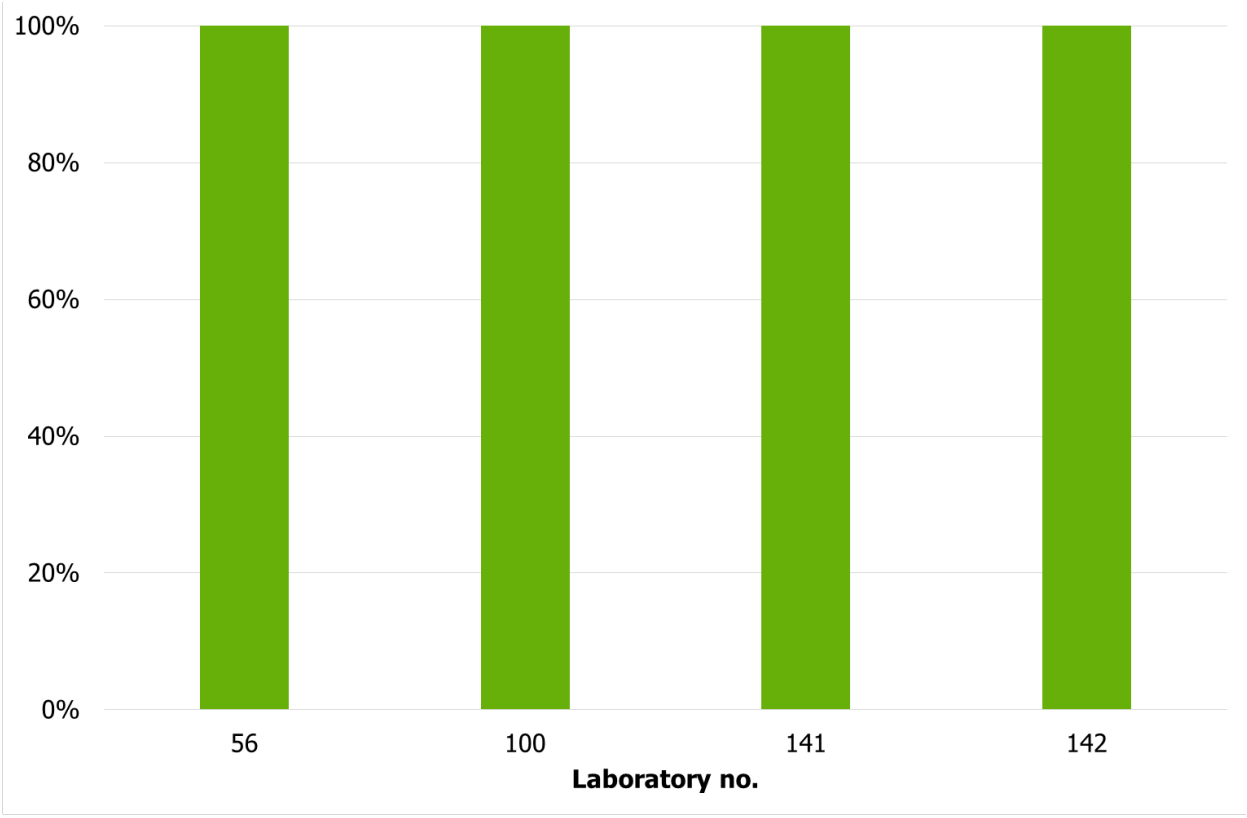
Eleven of the 17 participants (65%) completed both parts (serotyping and cluster analysis) of the EQA.
**: percentage of total number of participating laboratories (17).*

3.3 Serotyping

3.3.1 Conventional serotyping

Only four participants performed conventional serotyping of *L. monocytogenes* (Figure 1). Performance was high, with all four (100%) participants correctly serotyping all 10 test isolates. One isolate, Sero8, was excluded as retesting of the isolate showed unstable results.

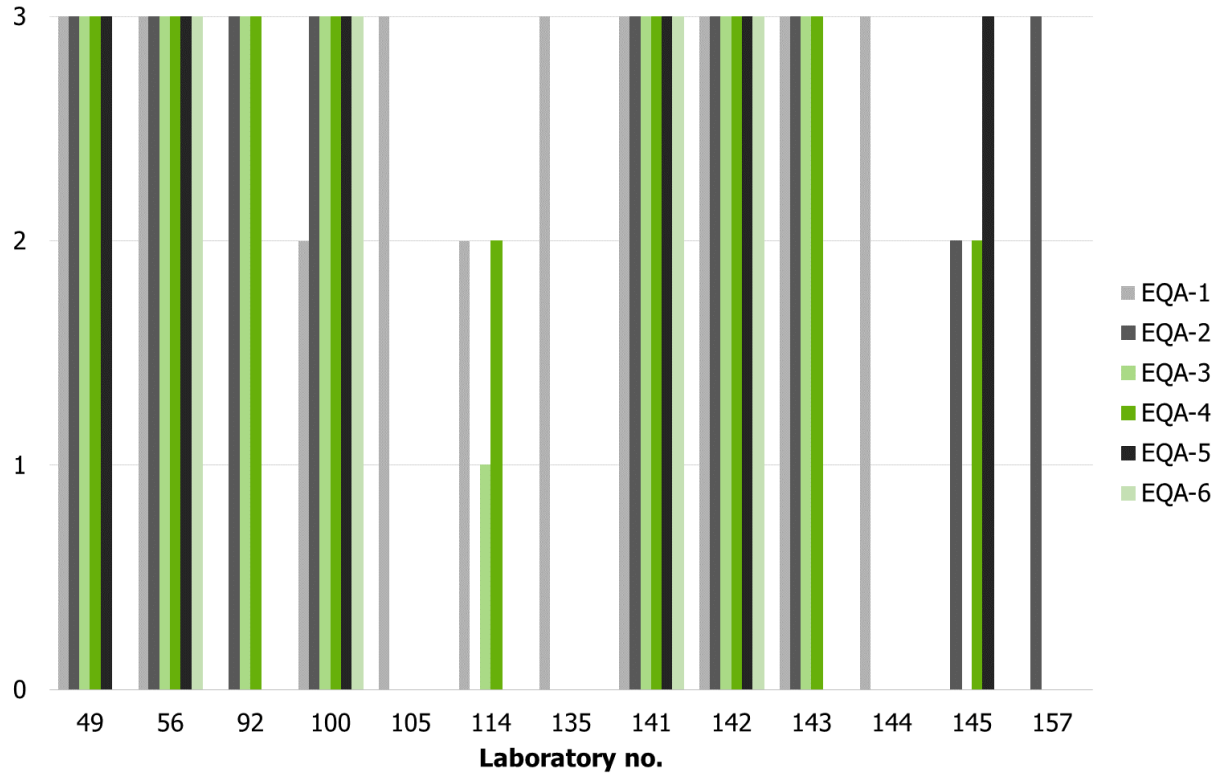
Figure 1. Participant scores for conventional serotyping of 10 test isolates



Arbitrary numbers represent participating laboratories. Bars represent the percentage of correctly assigned serotypes for the 10 test isolates (Sero1–7 and Sero9–11). Sero8 was excluded.

To follow the development of each laboratory's performance, three isolates of different serotypes were included in EQA-1 to 5: Sero9 (1/2a - IIa), Sero2 (technical duplet with isolates Sero11) (4b - IVb) and Sero6 (1/2c - IIc). Figure 2 shows the individual participants' performances on conventional serotyping of these three repeated isolates during the six EQAs. Conventional serotyping results on the repeated isolates show stability and high performance among the participants. All participants serotyped the three isolates correctly in EQA-6.

Figure 2. Correct conventional serotyping of three repeated isolates through EQA-1 to 6

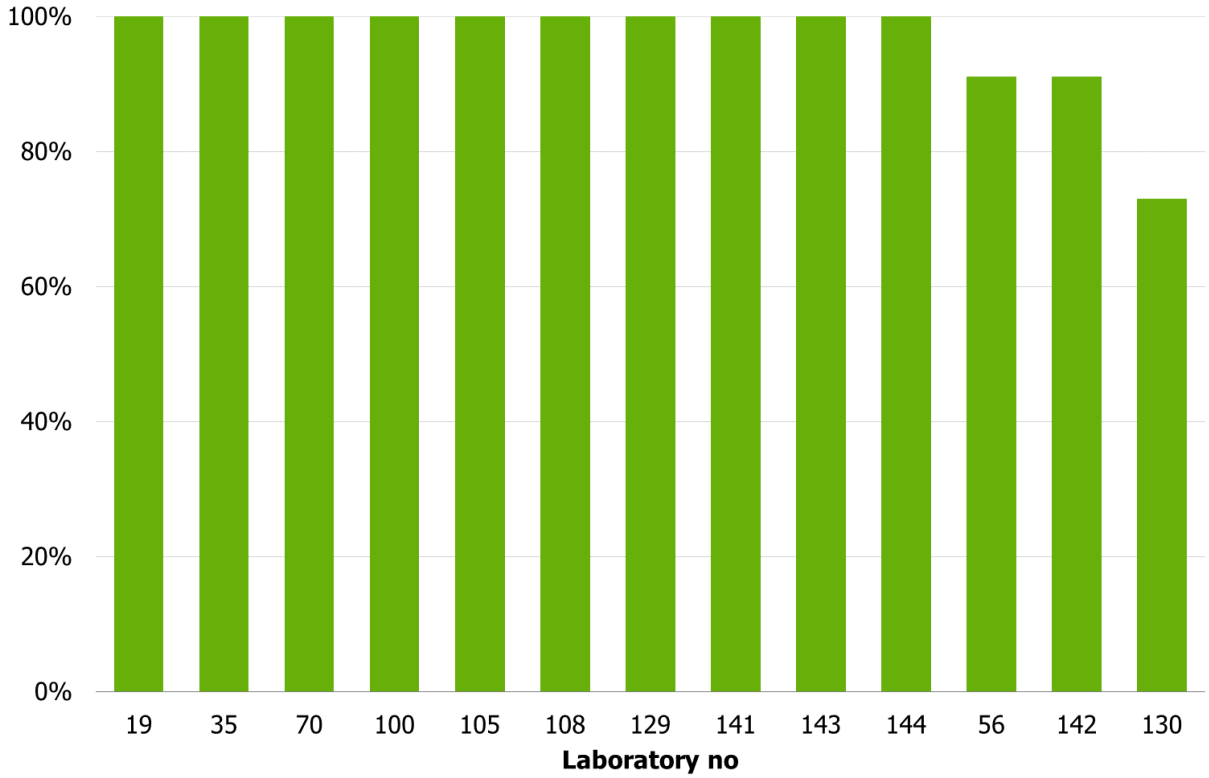


Arbitrary numbers represent the participating laboratories. Bars represent the number of correctly assigned serotypes for the three repeated isolates (Sero2, 6 and 9).

3.3.2 Molecular serotyping

Thirteen participants performed molecular serotyping of *L. monocytogenes* (Figure 3). In 2017 (EQA-5), two new laboratories (96 and 130) participated in molecular serotyping, but in EQA-6, only laboratory 130 participated again. Molecular serotyping was carried out in accordance with guidelines in Doumith et al. [5] and nomenclature from Doumith et al. [8] was used. Ten (77%) of the 13 participants were able to correctly serotype all 11 EQA test isolates. Laboratory 130 had only eight isolates correct and two other laboratories (56 and 142) both had one incorrect result. Three of the 13 participants reported using WGS-based analysis (*in silico* PCR) for molecular serotyping.

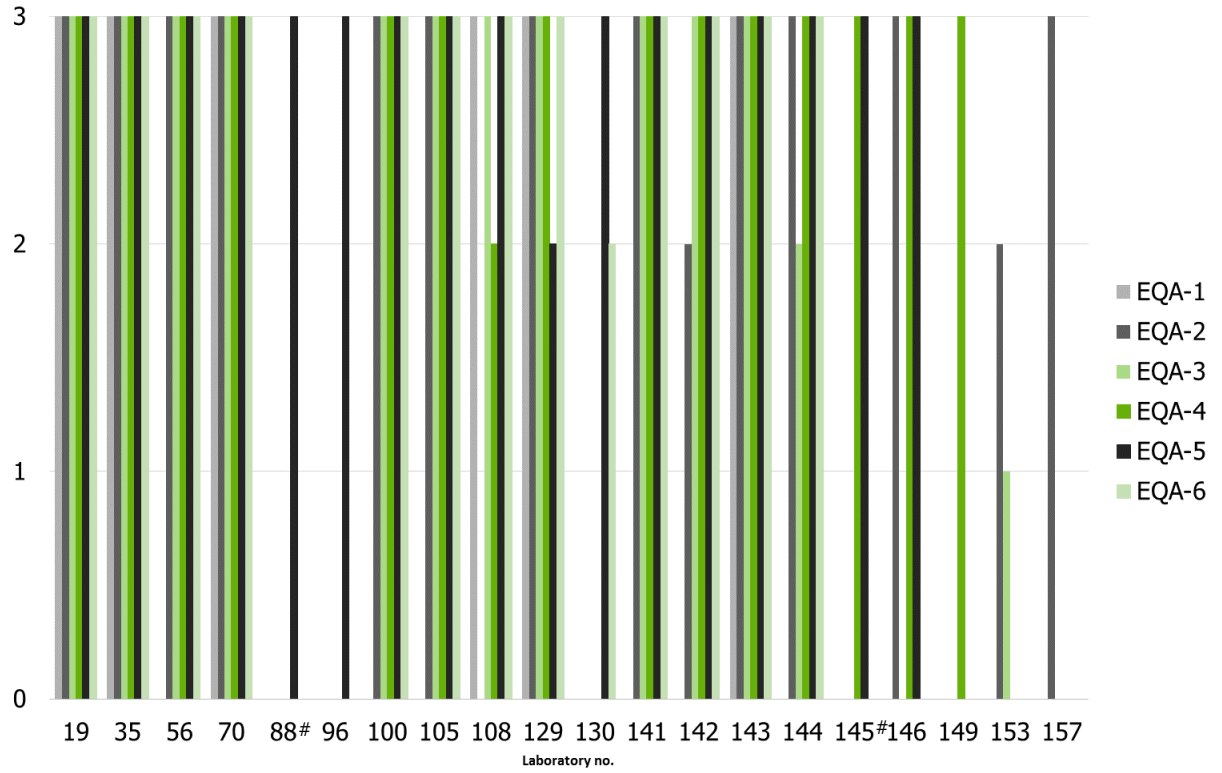
Figure 3. Participant scores for molecular serotyping of 11 test isolates



Arbitrary numbers represent the participating laboratories. Bars represent the percentage of correctly assigned serotypes for the 11 test isolates Sero1–11.

Figure 4 shows the individual participants' performances on molecular serotyping of the three repeated isolates during the five EQAs. As for conventional serotyping, the general performance among participating laboratories was high and stable. The majority of participants (89%: 71/80) correctly serotyped all three repeated isolates when participating.

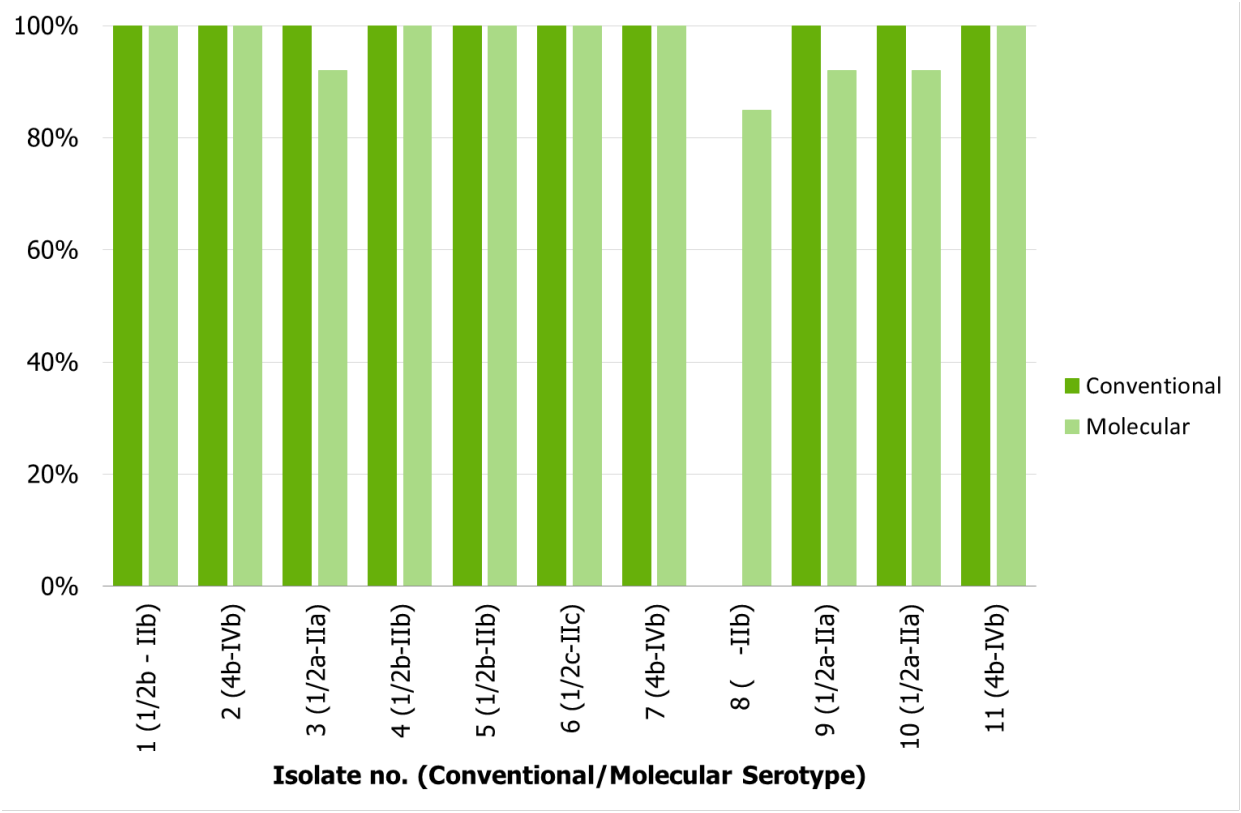
Figure 4. Correct molecular serotyping of three repeated isolates through EQA-1 to 6



Arbitrary numbers represents the participating laboratories. Bars represent the number of correctly assigned serotypes for the three repeated isolates (Sero2, 6 and 9).
#: laboratory did not correctly identify any of the three repeated isolates.

Seven (64%) of the 11 test isolates were correctly serotyped by all participants in the molecular serotyping and all 10 of the test isolates (isolate Sero8 was excluded) were correctly serotyped by all participants in conventional serotyping (Figure 5). All participants using conventional and/or molecular serotyping correctly serotyped seven isolates, but errors were reported in isolates Sero3, Sero8, Sero9 and Sero10. These isolates belonged to serotypes 1/2a – IIa and IIb. Again this year, serotype 4b -IVb was correctly assigned by all participating laboratories and the isolate with 1/2c - IIc was also serotyped 100% correctly.

Figure 5. Average score of 11 test isolates



Bars represent the percentage of correctly assigned serotypes by the participants.

3.4 Molecular typing-based cluster analysis

Participants were tested on their ability to correctly identify the cluster of closely related isolates defined by pre-categorisation from the EQA provider among the 11 cluster test isolates using either PFGE and/or WGS-derived data.

The EQA provider's PFGE results were based on combined *Apa*I and *Asc*I profiles. The EQA provider's cluster analysis of WGS-derived data was based on an allele-based (cgMLST [6]) and SNP analysis (NASP [7]).

The correct PFGE cluster contained seven ST399 isolates: REF1, REF3, REF4, REF6, REF7, REF8 and REF11 (REF1 and REF6 were technical duplicates). The correct WGS cluster contained five ST399 isolates: REF1, REF3, REF6, REF8 and REF11 (REF1 and REF6 were technical duplicates). The EQA provider found at most two allele differences or four SNPs between any two isolates in the cluster. The rest of the cluster test isolates were an additional three ST399s, one ST6, one ST11 and one ST14. (Annexes 5, 9).

3.4.1 PFGE-derived data

Seven (7/17, 41%) participants performed cluster analysis using PFGE-derived data. Performance was high, with six (6/7, 86%) of the participants correctly identifying the cluster of closely related isolates defined by a pre-categorisation from the EQA provider among the 11 cluster test isolates. Table 2 shows the overview of the isolate each participant included or excluded in their cluster identification.

Table 2. Results of cluster identification based on PFGE-derived data

		Laboratory ID						
Isolate number	**ST**	**19**	**56**	**130**	**138**	**141**	**142**	**144**
REF1[#‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF2	14	No	No	No	No	No	No	Yes
REF3[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF4[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF5	11	No	No	No	No	No	No	No
REF6[#‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF7[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF8[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF9	399	No	No	No	No	No	No	Yes
REF10	6	No	No	No	No	No	No	No
REF11[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster-identified conclusion		**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**No**

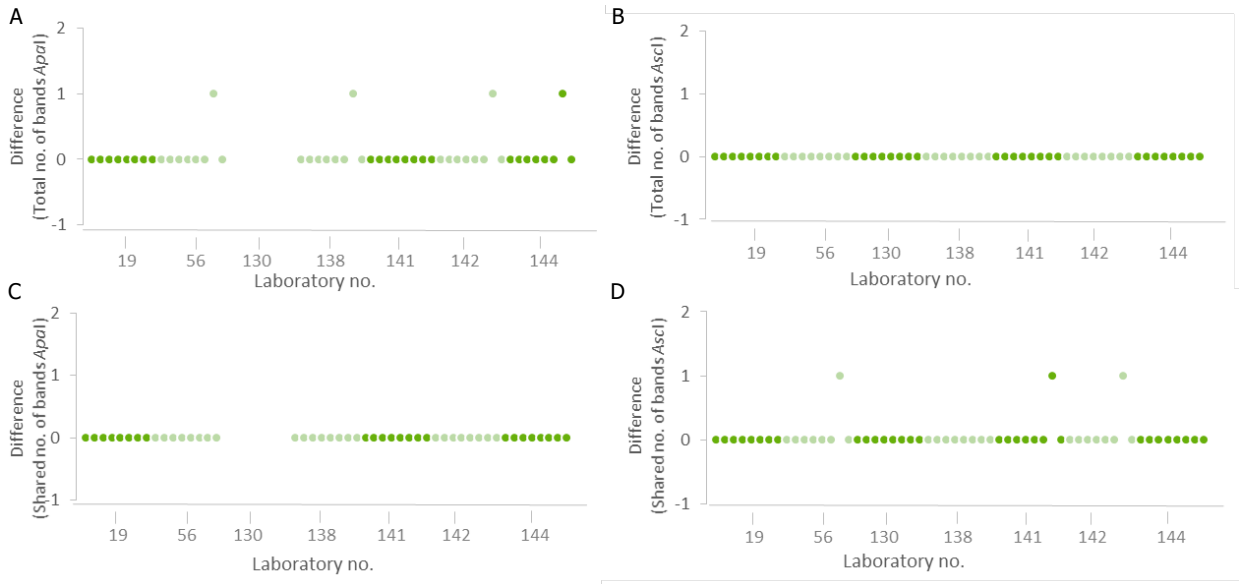
‡: closely related isolates
#: technical duplicate isolates (Annex 6).

For each isolate, participants were instructed to report the total number of bands in the *Apa*I and *Asc*I profiles separately. The number of bands shared between each test isolate and the selected cluster representative was reported for each enzyme (Figure 6). Data from laboratory 130 only contained *Asc*I profiles, but the correct cluster was identified (Annexes 6–7).

In Figure 6, A and B show the difference between the number of bands reported by the participants and the number observed by the EQA provider for *ApaI* and *AscI* respectively. Only one band difference was observed by four laboratories in the *Apa*I profiles in isolate RFE9. In the *AscI* profiles, one band is close to 33kb, so two results for total number of bands and shared bands were accepted as correct and are not shown as band differences in Figure 6B.

C and D show the difference between the participants' reported number of shared bands with a selected cluster representative and the number observed by the EQA provider for *Apa*I and *Asc*I, respectively. No band difference was observed in the *Apa*I profiles. The three differences were reported by laboratories 56, 141 and 142, which recorded a one band difference in shared bands using *Asc*I. Laboratory 144 did not identify the correct cluster of closely related isolates as they included REF2 (ST14) probably accepting several band differences (Annexes 5, 7) and also included REF9 (ST399) probably accepting one band difference of a larger band (Annexes 5, 7). The results of REF2 (ST14) are not included in Figure 9, which only includes the results of ST399.

Figure 6. Difference between reported total number of bands (A and B) and shared bands (C and D) for each isolate to selected isolates



Data from all eight ST399 isolates: REF1, REF3, REF4, REF6, REF7, REF8, REF9 and REF11.
Laboratory 130 only reported data for AscI. *In the* AscI *profiles, one band is close to 33kb, so two results (for total number of bands and shared bands) were accepted as correct (Annex 6).*

3.4.2 WGS-derived data

Reported results from participants

Twelve participants (12/17, 71%) performed cluster analysis using WGS-derived data. Only one laboratory reported using external assistance for sequencing. Different sequencing platforms were listed among the participants: one MiniSeq, five MiSeq, two HiSeq, two NextSeq, one Ion GeneStudio S5 System and one Ion Torrent. All reported using commercial kits for library preparation. Of the 12 participants, nine (75%) used Illumina's Nextera kit. One participant reported volume changes from the manufactory protocol and one laboratory listed less time for shearing and volume changes from the manufactory protocol (Annex 8).

Performance was high in cluster analysis with WGS-derived data. Eleven participants (92%) correctly identified the cluster of closely related isolates defined by pre-categorisation from the EQA provider among the 11 test isolates (Table 3). Laboratory 56 included two additional ST399 isolates (REF4 and REF7) as being in the cluster of closely related isolates.

Table 3. Results of cluster identification based on WGS-derived data

		Laboratory ID											
Isolate number	**ST**	**19**	**35**	**56**	**70**	**105**	**108**	**129**	**135**	**141**	**142**	**146**	**149**
REF1[#‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF2	14	No	No	No	No	No	No	No	No	No	No	No	No
REF3[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF4	399	No	No	Yes	No	No	No	No	No	No	No	No	No
REF5	11	No	No	No	No	No	No	No	No	No	No	No	No
REF6[#‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF7	399	No	No	Yes	No	No	No	No	No	No	No	No	No
REF8[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF9	399	No	No	No	No	No	No	No	No	No	No	No	No
REF10	6	No	No	No	No	No	No	No	No	No	No	No	No
REF11[‡]	399	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main analysis		Allele	Allele	SNP	Allele	Allele	SNP	Allele	Allele	Allele	Allele	SNP	Allele
Cluster-identified conclusion		**Yes**	**Yes**	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**

‡: closely related isolates
#: technical duplicate isolates
ST: 7 multilocus sequence type
Allele: allele-based analysis
SNP: single-nucleotide polymorphism (Annex 10).

Laboratories were instructed to report the data analysis used for cluster identification and select a representative isolate in the cluster for reporting SNP distance or allelic differences between the selected isolate and each test isolate included in analysis. Laboratories could report results from up to three analyses (1 main and 1–2 additional), but the detected cluster had to be based on results from the main analysis. In EQA-6, only one analysis per laboratory was reported.

Of the 12 participants, three (25%) used SNP, one in-house pipeline and two published pipelines. Two used a reference-based approach with different ST399 isolates as reference. One used CLC for both read mapper and variant caller and the other used Burrows-Wheeler Aligner (BWA) as the read mapper and a different variant caller, Genome Analysis Toolkit (GATK; Tables 4–5).

Table 4. Results of SNP-based cluster analysis

Laboratory		SNP-based						
	SNP pipeline	Approach	Reference	Read mapper	Variant caller	Assembler	Distance within cluster	Distance outside cluster¤
Provider	NASP	Reference-based	REF11 (ST399)	BWA	GATK		0–3	34-51 (7953–115961)
56	kSNP3	Assembly-based	-	-	kSNP3	SPAdes	0–1299	662 (10177–40660)
108	In-house pipeline	Reference-based	In-house strain resp ST	CLC assembly cell v4.4.2	CLC assembly cell v4.4.2	-	0–4	32–831
146	SnapperDB + PHEnix - in house/ publicly available	Reference-based	ST399, FR733650	BWA via PHEnix	GATK via PHEnix	-	0–4	Not reported

¤: reported distance to ST399 isolates (non-ST399; Annex 11).

Nine of the 12 participants that used allele-based analysis selected this method as the main analysis for cluster detection (Table 3). Seven (78%) used an assembly-based allele calling method and two used both mapping and assembly-based allele calling (Table 5). All reported using cgMLST, six (67%) used cgMLST Ruppitsch (1701 loci) [9], two cgMLST Pasteur (1748 loci) [6] and one an in-house cgMLST scheme with only 1503 loci.

Table 5. Results of allele-based cluster analysis

Laboratory	Allele-based analysis						
	Approach	Allelic calling method	Assembler	Scheme	Number of loci	Difference within cluster	Difference outside cluster¤
Provider	BioNumerics	Assembly- and mapping-based	SPAdes	Applied Math (cgMLST/Pasteur)	1748	0–2	17–23 (511–1690)
19	BioNumerics	Assembly- and mapping-based	SPAdes	Applied Math (cgMLST/Pasteur)	1748	0–3	15–22 (513–1705)
35	SeqPhere	Only assembly based	Velvet	Ruppitsch (cgMLST)	1701	0–1	14–26 (489–1669)
70	SeqPhere	Only assembly based	Velvet	Ruppitsch (cgMLST)	1701	0–1	Not reported
105	SeqPhere	Only assembly based	SPAdes	Ruppitsch (cgMLST)	1701	0–1	14–26 (491–1672)
129	SeqPhere	Only assembly based	Velvet	Other (Ridom SeqSphere+ software's Target Definer, cgMLST)	1503	0-2	10–22 (413–1468)
135	SeqPhere	Assembly- and mapping-based	CLC Genomics Workbench	Ruppitsch (cgMLST)#	1701	0-1	14–51 (537–1672)
141	SeqPhere	Only assembly based	SPAdes 3.11.1	Ruppitsch (cgMLST)	1701	0-1	15–26 (492–1675)
142	Other	Only assembly based	SPAdes	Pasteur (cgMLST)	1748	0-2	10–24 (514–1736)
149	SeqPhere	Only assembly based	Velvet	Ruppitsch (cgMLST)	1701	0-1	14–26 (493–1679)

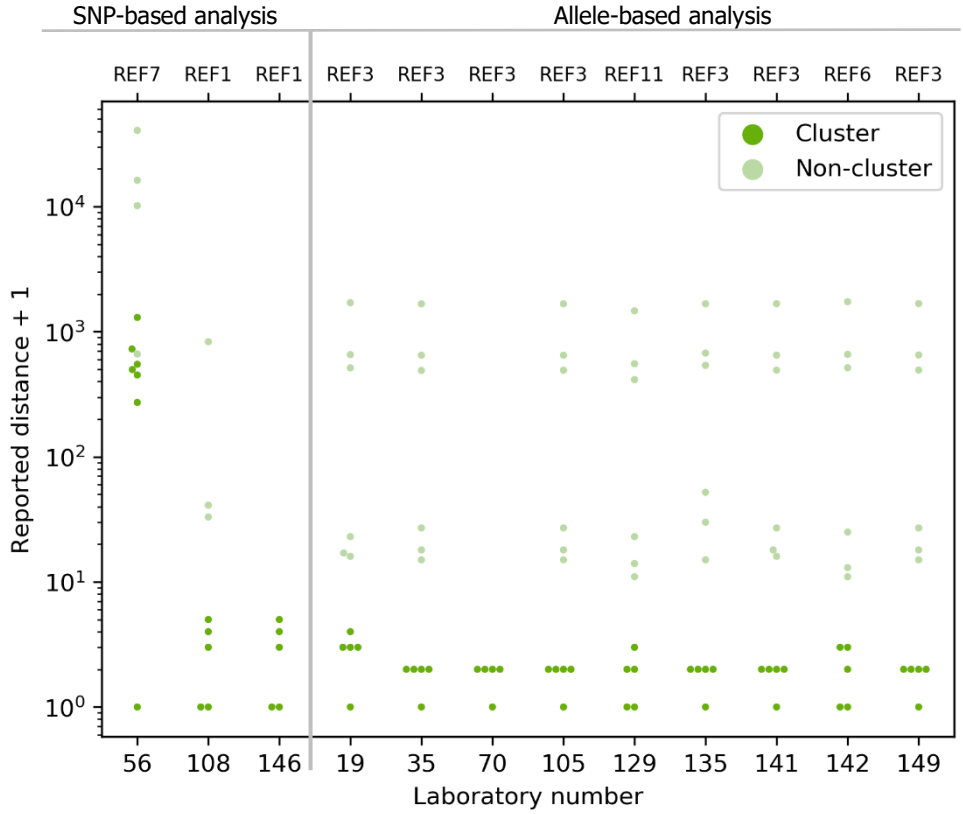
¤: reported distance to ST399 isolates (non-ST399)
#: reported as standard cgMLST scheme available in Seqsphere (Annex 11).

All nine laboratories performing cgMLST identified the correct cluster of closely related isolates (Figure 7). All nine laboratories reported allele differences of 0–3 within the cluster (Table 5). Seven laboratories selected REF3, one used REF6 and one REF11 (Figure 7).

Three other test isolates (REF4, REF7 and REF9) were also ST399, but not predefined by the EQA provider as part of the cluster. Based on cgMLST, the nine laboratories reported allele differences to the selected cluster isolate at 10–51 for this group of isolates. Three test isolates (REF2, REF5 and REF10) were not ST399 and allele differences to the selected cluster isolate at 413–1736 were reported (Annex 11).

Laboratories 19 and 142 used the same cgMLST scheme as the EQA provider (cgMLST/Pasteur) [6] and all but one laboratory (129) used the Ruppitsch scheme [9].

Figure 7. Reported SNP distances or allelic differences for each test isolate to selected cluster representative isolate



SNP: single nucleotide polymorphism
Selected cluster representative marked as REF.
Dark green: reported cluster of closely related isolates
Light green: not reported as part of cluster.

Of the three laboratories performing SNP analysis, laboratories 108 and 146 identified the correct cluster of closely related isolates (Figure 7). Laboratory 56 performed only SNP analysis and could not identify the correct cluster.

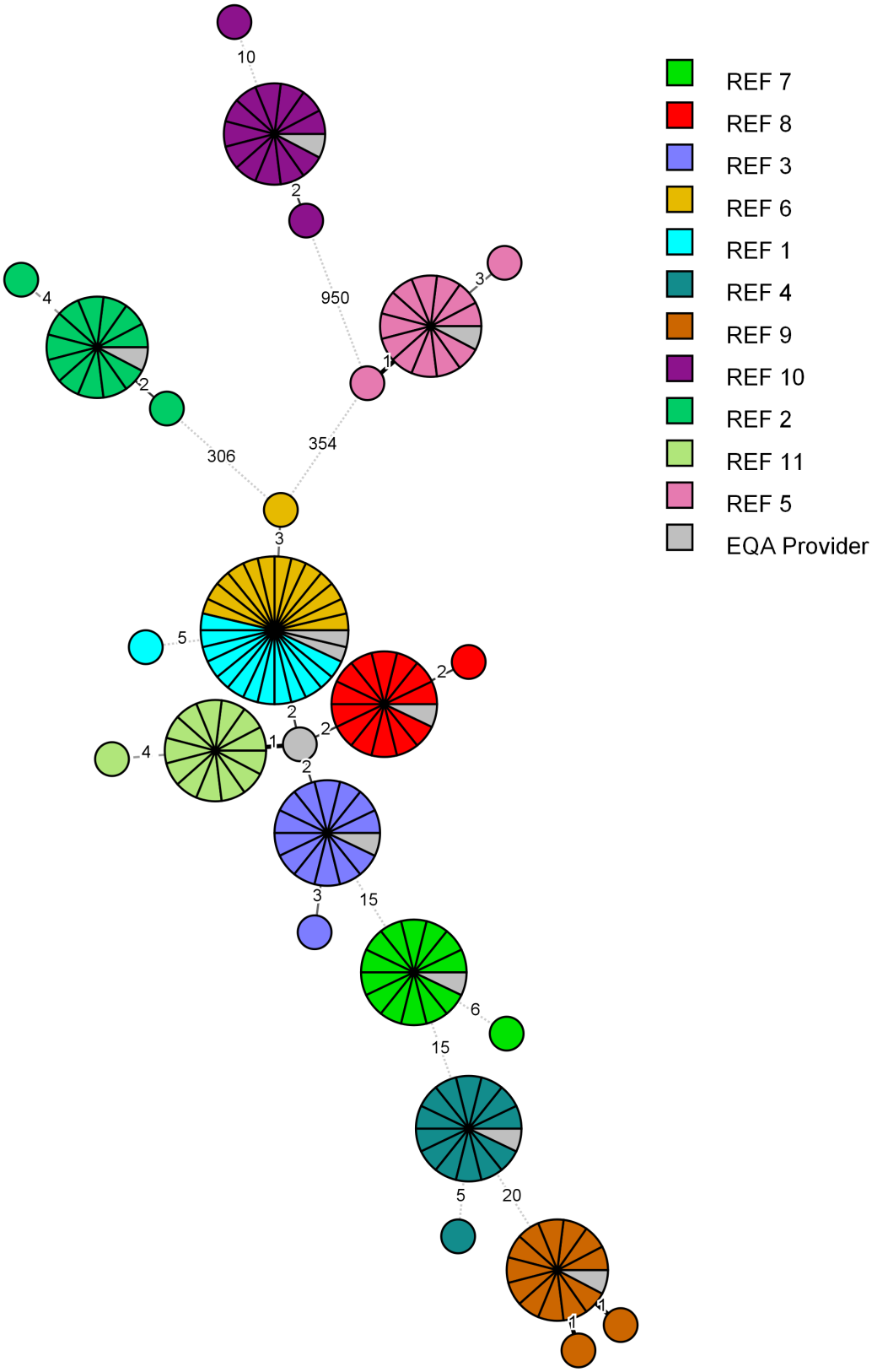
The reported SNP differences within the cluster varied from 0–4 (laboratories 108 and 146) to 0–1299 (laboratory 56). If the cut-off for cluster definition used by laboratory 56 was used on data from laboratory 108, all eight ST399 isolates would have been included in the cluster. Laboratory 56 had selected REF7, which is not a part of the correct cluster, as the cluster representative. Laboratories 108 and 146 selected REF1 as the cluster representative isolate.

Analysis of raw reads uploaded by participants

In addition to the reported cluster identification, participants submitted their FASTQ files to be evaluated by the EQA provider. The FASTQ files were uploaded to an Applied Maths calculation engine for allele calling (Institut Pasteur) [6] and evaluated by the EQA provider's in-house quality control (QC) pipeline [10].

The overall cgMLST analysis, shown in the minimum spanning tree (MST) based on submitted raw reads (FASTQ files) from 12 laboratories reveals clear clustering of the results for each test isolate (Figure 8).

Figure 8. Minimum spanning tree of core genome multilocus sequence typing participant FASTQ files

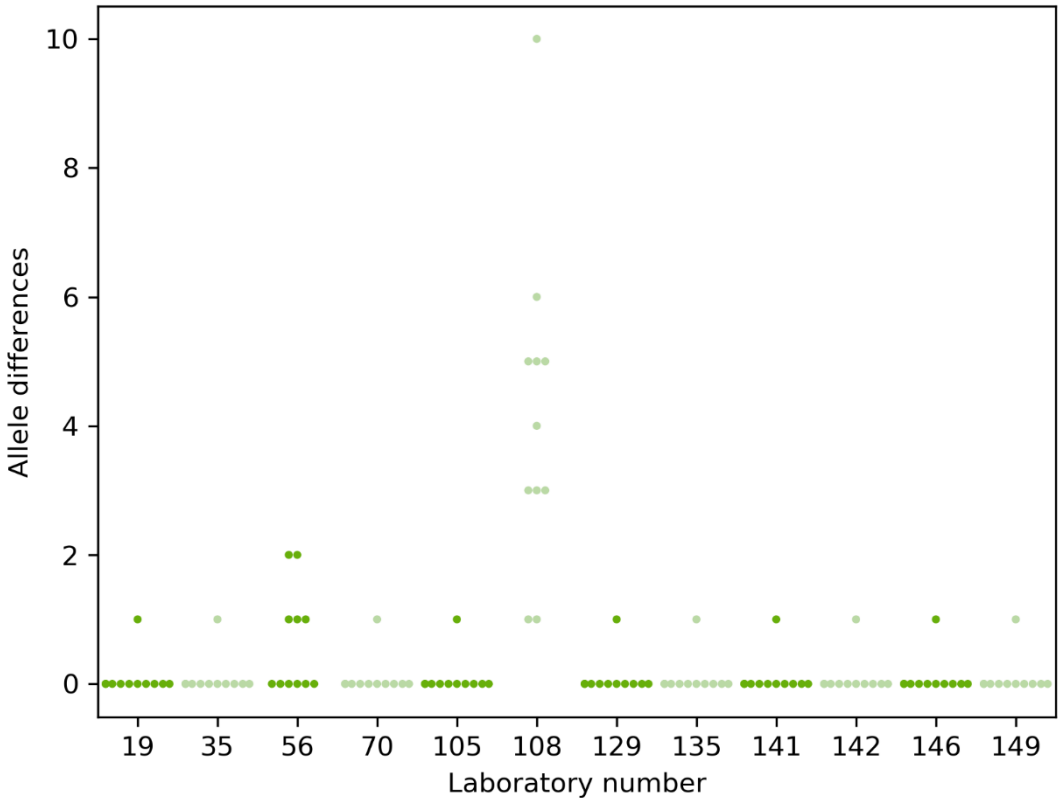


Minimum spanning tree (MST) in log scale of core genome multi-locus sequence typing (cgMLST) [6] based on submitted raw reads (FASTQ files). Each of the REF1–11 test isolates have a different colour. REF results from the EQA-provider are in grey. Of the 12 laboratories.

The allele differences in Figure 8 do not exactly match those illustrated in the individual reports and consequently those in Figures 9, where the same data are used. This discrepancy is caused by loci being dropped if they do not pass QC in all isolates in the analysis. Joint analysis thus contains fewer loci. As seen in Figure 8, the provider isolate REF11 is one allele removed from all participant isolates.

For each laboratory, cgMLST was performed on the submitted raw reads (FASTQ files), applying Applied Maths allele calling with the Pasteur scheme [6]. For each laboratory, a hierarchical single linkage clustering was performed on the submitted data along with the EQA provider's reference isolates. Figure 9 shows the allele differences between each submitted sequence and the corresponding reference.

Figure 9. Participant allele difference from reference result (EQA provider) for each test isolate



Allele difference of participant isolates from corresponding REF isolates (EQA provider) based on submitted raw reads (FASTQ files).

For 106 of 132 results (80%), no difference was identified. For 17 results (13%), a difference of 1–2 alleles from the REF isolate was calculated and a difference of 3–10 alleles were seen for nine results (7%). The provider result for REF11 was one allele removed from most of the participants (Figure 8) and leads to the majority of the one-allele differences in Figure 9. Laboratory 56 had additionally allele differences ranging from 1–2 and results from laboratory 108 showed allele difference for all 11 isolates, 8 isolates with difference of 3–6 alleles and one isolate with a difference of 10 alleles.

Separately, the laboratories responded to QC parameters used to evaluate their data. Both coverage and confirmation of genus was the most widely used QC parameter, with 83% of the laboratories using this parameter (Table 6). Participants used different thresholds of coverage ranging from 20–60x coverage. Different programs used for the contamination check of genus were reported. The number of good cgMLST loci was used as a QC parameter by 75% of the participants, with thresholds ranging from 89–99%. Q score and genomic size were used by 42% and 50% of the participants respectively.A few laboratories reported additional parameters (Annex 12). For the full QC evaluation of all isolates, see Annex 13.

Table 6. Summary of selected QC parameters reported by participants

Laboratory	Confirmation of genus	Coverage	Q score (Phred)	Genome size	Number of good cgMLST loci
70	Specific task in SeqSphere targets prs gene	Coverage > 50	No	Length of contigs assembled < ref genome + 10%	cgMLST alleles found and called >95%
149	KRAKEN	No	No	No	>95%
142	No	40	28	No	0.97
146	Kmer ID	MOST – coverage for MLST locus + SnapperDB global coverage	30	No	No
35	cgMLST	No	No	No	Not yet validated
129	Presence of prfA gene (LIP)	>29	No	No	>89%
135	No	Coverage threshold >30	Phred score threshold >30	No	Standard procedure in SeqSphere
56	No mismatches in alignment with 7 housekeeping genes of MLST panel	30	22	2.8 Mb	No
19	Kraken-correct genus and <5% contamination with others	Minimum x25	No	Genome size: 2.800000bp–3.100000bp	Core percent, all 99% or above. Also, max of 15 loci with multiple consensus. No actual threshold employed on regular basis for either.
105	Assembled genomes aligned against *L. monocytogenes* reference genome (threshold: >90% nucleotide identity)	Depth of coverage >45X	Trimming performed with Trimmomatic, removing 3 nucleotides with Phred <10 or an average Phred <15 in a sliding window of 4 nucleotides. Sequences with a length of <70 bases were also removed.	<=3.3Mb	>=95%
108	Blast	20 x	No	+/- 20%	No
141	JSpecies	60 x	No	Approximate size of *Listeria* genome: 2.8 million	Minimum 99% good targets
% of laboratories using QC parameter	**83%**	**83%**	**42%**	**50%**	**75%**

For each laboratory, the submitted raw reads (FASTQ files) were evaluated by the EQA provider's in-house quality control pipeline [10]. Table 7 shows the QC parameters and range of QC values per laboratory. For the full QC evaluation of all isolates, see Annex 13.

According to the QC parameters, sequencing quality was uniformly good. Coverage was high overall. Laboratory 35 used single-end sequencing on the Illumina platform without apparent negative consequences for any QC parameters or the cluster analysis.

Table 7. Results of raw reads submitted by participants evaluated by EQA provider QC pipeline summarised by laboratory

		Laboratory ID											
Parameters	Ranges*	19	35	56	70	105	108	129	135	141	142	146	149
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		2.6–4.3	1.2–2.1	1.9–3.1	1.3–2.2	1.7–10.4	1.5–4.2	0.6–1.2	2.1–3.2	1.3–2.1	0.5–1.7	0.7–2.4	1.8–3.0
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	2.9–3.0	2.9–3.0	0.5–2.5	2.9–3.0	2.9–5.8	2.5–3.0	2.9–3.0	2.9–3.0	2.6–3.0	2.9–3	2.9–3.0	2.9–3.1
Length [0-25] x min. coverage (Mbp)	{<0.25}	0	0	0.0–0.8	0	0.0–1.2	0.0–0.4	0	0	0.0–0.3	0	0	0
Number of contigs at 25x min. coverage	{>0}	17–40	31–62	277–1 099#	19–100	16–151	282–638#	16–35	16–35	25–119	12–31	17–60	11–499
Number of contigs [0–25] x min. coverage#	{<1000}	0-1	0–17	0–541#	0–20	0–82	0–183#	0	0	0–58	0–5	0–6	0
Average coverage	{>50}	117–178	49–98	21–81	35–124	96–284	29–91	119–225	150–338	29–65	44–71	35–159	134–201
Number of reads (x1000)		1339–2091	641–1684	242–1131	365–1314	2068–3733	441–954	1253–2400	2037–4665	187–434	294–454	580–2556	1108–1508
Number of trimmed reads (x1000)		1325–2074	640–1682	226–1063	361–1300	2044–3674	393–854	1231–2369	2013–4 601	177–424	286–444	559–2 480	1056–1436
Maximum read length		151	151	280	151	151	194–324	151	126	301	251	101	301
Mean read length		133–139	143-148	143-178	144–147	116–144	147–216	145–148	116–117	231–251	217–340	98–99	195–215
Read insert size		235–287	NA	NA	303–345	195–285	NA	369–430	259–282	316–423	425–608	287–367	269–302
Insert size StdDev		109–127	NA	NA	100–122	102–122	NA	145–164	165–182	128–158	144–204	143–193	115–136
N50 (kbp)		292–456	109–327	1–2	50–456	80–510	5–17	295–456	331–456	30–349	223–542	90–456	9–520
N75 (kbp)		110–252	69–150	1	30–332	30–332	3–9	125–332	165–331	18–129	109–456	53–332	5–456

**: indicative QC range*
Lm: L. monocytogenes
NA: not available
#: QC values unreliable due to assembly issues for Ion Torrent data (Annex 13).

4 Discussion

4.1 Serotyping

Thirteen (76%) laboratories participated in the serotyping part of EQA-6 and all 13 (100%) provided molecular serotyping results. Almost half of the EU laboratories have developed the capacity to perform molecular serotyping, but the high number of missing laboratories indicates that transition is uncompleted.

4.1.1 Conventional serotyping

The number of participants decreased from 10 in EQA-3 to four in EQA-6, highlighting the transition towards the use of molecular serotyping. Comparing the conventional serotyping results from EQA-1 to 6, showed stable high performances among participants during the EQAs (EQA-1 to 6: 94%; 87%; 91%; 97%; 98%; 100%). The four participants in EQA-6 performed at the same level (100%) or better than the year before. One isolate (Sero8) was excluded from analysis due to unstable results.

4.1.2 Molecular serotyping

A higher number of participants was observed in EQA-6 than in EQA-3, which shows the increased use of molecular serotyping,. and three laboratories in EQA-6 even reported the use of *in silico* PCR serotyping. The performance of the laboratories with regard to molecular serotyping was high, with 77% of participants obtaining a score of 100% correct. The general performance among the participating laboratories has been high over the years from EQA-1 to 6: 98%; 94%; 94%; 94%; 99%; 97%. Only one laboratory did not correctly serotype all three repeated isolates.

4.2 Molecular typing-based cluster analysis

In EQA-5 and 6, PFGE is no longer an independent part, but by adding cluster identification using either PFGE and/or WGS-derived data, this EQA is contemporary with the development of surveillance methods used by public health national reference laboratories in Europe. This adjustment of the EQA still appears to be well accepted by the Member States, as 15 of the 17 laboratories (88%) participated. Only seven participated in cluster identification using PFGE-derived data, while four also participated in cluster identification using WGS-derived data.

4.2.1 PFGE-derived data

Of the 17 laboratories, seven (41%) performed cluster analysis using PFGE-derived data. Performance was high, with six participants (86%) correctly identifying the cluster of closely related isolates. The present cluster designed by the EQA provider allowed participants to detect the cluster of closely related isolates by either using the *Apa*I or *Asc*I enzymes. Only one laboratory did not identify the cluster. The laboratory had a minor difference in the total number of observed bands in an *Apa*I profile and in the number of observed shared bands in an *Asc*I profile compared with the EQA provider.

In EQA-5, seven laboratories also participated in the cluster part using PFGE, but only four were the same as in EQA-6. However, two of the three laboratories who stopped using PFGE did not participate in the WGS-based cluster analysis.

4.2.3 WGS-derived data

Twelve of the 17 laboratories (71%) performed cluster analysis using WGS-derived data. In EQA-5, 12 laboratories also participated in WGS-based cluster analysis, but in EQA-6, one laboratory stopped and one new started. Performance was again very high, with 11 (92%) correctly identifying the cluster of closely related isolates. As in EQA-5, only one laboratory reported the use of external assistance for sequencing and the majority (9/12) also reported using an Illumina platform. All reported using commercial kits for preparing the library.

Eleven of the 12 participating laboratories were the same as in EQA-5 and out of these, one changed from SNP analysis in EQA-5 to allele-based analysis in EQA-6. The one laboratory that did not identify the correct cluster had used SNP analysis.

The two laboratories identifying the correct cluster by SNP analysis reported distances comparable to those reported using allele-based methods (0–4 SNPs). The remaining laboratories reported SNP distances that were several orders of magnitude higher than for the correct cluster (271–1299) as they selected a cluster representative outside the expected cluster. The laboratories identified a very large number of SNPs and could not separate the cluster from the non-cluster isolates using the submitted SNP distances. Laboratory 56 used assembly-based SNP analysis from Ion Torrent data. Assembly of Ion Torrent data has been reported to be challenging [11]. Laboratory 108 identified the correct cluster using Ion Torrent data mapping to a reference.

Inflated distances were also reported by laboratory 56 in EQA-5, so this may indicate a problem with the used method.

Reported cgMLST results were almost identical, with only 0–3 allelic differences within the cluster isolates and 10–51 for ST399 outside the cluster despite not being analysed using the same scheme or assembly tool.

Analysing all participants' raw reads in the same scheme [6], the maximum differences between any two cluster isolates were three alleles (data not shown) using Illumina data, which are similar to the distances reported by the participants. For Ion Torrent data, the maximum difference between any two cluster isolates were eight alleles when data were analysed by EQA provider.

The main reported QC parameters were coverage, cgMLST allele calls and genus/species confirmation, which are all essential for the end use of data.

In order to compare the quality of the raw data, the EQA provider analysed the submitted raw reads to obtain selected QC parameters. Using single-end sequencing on the Illumina platform, laboratory 35 reduced the run time without apparent negative consequences for any QC parameters or the cluster analysis.

The EQA provider's analysis of the submitted raw data showed that when using a standardised cgMLST analysis, a very high concordance was obtained (Figure 9). The majority of the 'one-allele differences' in Figure 9 were caused by the EQA provider result of REF11. Only laboratories 56 and 108 had additional allele differences ranging from 1–2 and 1–10 respectively. These laboratories both provided Ion Torrent data, for which the EQA provider's analysis is not optimised, making correct assembly difficult. Accordingly, the observed allele differences may be method artefacts.

5 Conclusions

Seventeen laboratories participated in the EQA-6 scheme, with 13 (76%) performing serotyping and 15 (88%) cluster identification. In EQA-5, a change was made from including quality assessment of PFGE to include molecular typing-based cluster analysis using either PFGE and/or WGS-derived data. Incorporating molecular typing-based cluster analysis in the EQA is up to date with the development of surveillance methods used by public health national reference laboratories in Europe. This adjustment of the EQA still seemed to be well accepted by Member States, as the same number of laboratories participated in cluster identification in EQA-6, but the level of participation in serotyping decreased.

In the present EQA, none of the serotyping participating laboratories solely performed conventional serotyping. Most laboratories (69%, 9/13) performed only molecular serotyping and 31% molecular serotyping in combination with the conventional serotyping. In general, the trend towards substituting conventional serotyping with molecular was observed through the five EQAs.

The quality of serotyping was high. The performance of conventional serotyping was highest, with 100% and 77% of the participants correctly serotyping all test isolates by conventional and molecular methods respectively.

Seven laboratories participated using PFGE for cluster analysis. Three participated solely using PFGE-derived data for analysis. Only one did not identify the correct cluster using PFGE, accepting too many band differences.

Twelve laboratories performed cluster analysis using WGS-derived data. The performance was very high, 11 (92%) of the participants correctly identified the cluster of closely related isolates.

An allele-based method was preferred, as 75% (9/12) used cgMLST compared to 25% (3/12) using SNP as the main reported cluster analysis. In general, the reported cgMLST results were at a comparable level of allelic difference (0–3) within the cluster isolates despite analysis with different schemes. Similar results were obtained by the two laboratories using SNP analysis and reporting the correct cluster (0–4 SNP distance inside the cluster).

Laboratory 56 not identifying the correct cluster used an assembly-based SNP analysis on Ion Torrent data. Nevertheless, laboratory 108 identified the correct cluster using Ion Torrent data by mapping to a reference. In the proficiency test for *Listeria monocytogenes* whole genome assembly conducted by ECDC in 2018, the assembly of Ion Torrent data was shown to be challenging. The issue with the inflated distances was also reported by laboratory 56 in EQA-5. This may indicate a problem with the used method as the laboratory identified a very large number of SNPs and using the submitted SNP distances, the laboratory could not separate the cluster from the non-cluster isolates.

In EQA-5, it was concluded that cgMLST had higher consistency than SNP analyses, resulting in better inter-laboratory comparability. This was not as clear in EQA-6 since only three laboratories reported SNP distances and only one laboratory could not identify the correct cluster. Still, standardized cgMLST analyses leave little room for error, resulting in good performance. The current EQA also showed a difference comparing Illumina and Ion Torrent data using cgMLST, leading potentially to lower inter-laboratory comparability across sequencing platforms.

The current EQA scheme for *L. monocytogenes* typing is the sixth organised for laboratories in FWD-Net. The molecular typing-enhanced surveillance system implemented as part of TESSy relies on the capacity of FWD-Net laboratories to produce analysable and comparable typing results in a central database. In March 2019, ECDC launched the possibility of submitting WGS variables for *L. monocytogenes* to TESSy to be used for EU-wide surveillance and cross-sector comparison.

6 Recommendations

6.1 Laboratories

Laboratories are recommended to use EQA-provided data and isolates to validate their analysis methods when incorrect results (e.g. EQA) are obtained or when implementing new methods and procedures.

6.2 ECDC and FWD-Net

ECDC works actively with FWD-Net to improve the quality of sequence data generation and analysis for *L. monocytogenes* through appropriate means like EQA schemes, expert exchange visits and workshops.

6.3 EQA provider

For the following EQA round, the EQA provider will evaluate the possibility of modifying the cluster analysis to mimic a more realistic microbiological investigation by providing genome sequences that can be included in the WGS analysis. This part is designed to be a simulation of an outbreak situation in the country and compare the original cluster with genomes produced in other laboratories which might be using different procedures and equipment.

7 References

1. European Parliament, European Council. Regulation (EC) No 851/2004 of the European Parliament and of the Council of 21 April 2004 establishing a European Centre for Disease Prevention and Control, Article 5.3. Off J Eur U. 2004 Apr 30;L142:1-11. Available from: http://ecdc.europa.eu/en/aboutus/Key%20Documents/0404_KD_Regulation_establishing_ECDC.pdf
2. European Parliament, European Council. Decision No 1082/2013/EU of the European Parliament and the Council of 22 October 2013 on serious cross-border threats to health and repealing Decision No 2119/98/EC (Text with EEA relevance). Off JEur U. 2013 Nov 5;L293:1-15. Available from: http://ec.europa.eu/health/preparedness_response/docs/decision_serious_crossborder_threats_22102013_en.pdf
3. European Food Safety Authority, European Centre for Disease Prevention and Control. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. EFSA J. 2018 Nov 19;16(12):5500. Available from: http://ecdc.europa.eu/publications-data/european-union-summary-report-trends-and-sources-zoonoses-zoonotic-agents-and-10
4. International Organization for Standardization. ISO/IEC 17043:2010 – Conformity assessment -- General requirements for proficiency testing. Geneva: ISO; 2015. Available from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=29366
5. Doumith M, Buchrieser C, Glaser P, Jacquet C, Martin P. Differentiation of the Major *Listeria monocytogenes* Serovars by Multiplex PCR. J Clin Microbiol. 2004 Aug;42(8):3819-22.
6. 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*. Nat Microbiol. 2016 Oct 10;2:16185. Available from: http://www.nature.com/articles/nmicrobiol2016185
7. Doumith M, Jacquet C, Gerner-Smidt P, Graves LM, Loncarevic S, Mathisen T, et al. Multicenter validation of a multiplex PCR assay for differentiating the major *Listeria monocytogenes* serovars 1/2a, 1/2b, 1/2c, and 4b: toward an international standard. J Food Prot. 2005 Dec;68(12):2648-50.
8. Sahl JW, Lemmer D, Travis J, Schupp JM, Gillece JD, Aziz M, et al. NASP: an accurate, rapid method for the identification of SNPs in WGS datasets that supports flexible input and output formats. Microb Genom. 2016 Aug 25;2(8):e000074. Available from: http://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.000074
9. 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*. J Clin Microbiol. 2015 Sep;53(9):2869-76.
10. Statens Serum Institut. SerumQC [Internet, software package]. Copenhagen: SSI; 2019 [cited 20 June 2018]. Available from: http://www.github.com/ssi-dk/SerumQC
11. European Centre for Disease Prevention and Control. Proficiency test for *Listeria monocytogenes* whole genome assembly – 2018. Stockholm: ECDC; 2019. Available from: http://ecdc.europa.eu/publications-data/proficiency-test-listeria-monocytogenes-whole-genome-assembly-2018

Annex 1. List of participants

Country	Laboratory	National institute
Austria	National Research Laboratory *Listeria* Austria	AGES – Austrian Agency for Health and Food Safety
Belgium	National Reference Centre *Listeria* Belgium	Sciensano
Denmark	Diagnostic and Typing of Gastrointestinal Bacteria	Statens Serum Institut
Finland	Expert Microbiology	National Institute for Health and Welfare
France	National Reference Centre and WHO Collaborating Centre for *Listeria*	Institut Pasteur
Germany	Binational German-Austrian Consiliary Laboratory for *Listeria*	Robert Koch Institute
Greece	National Reference Centre for *Salmonella* and Other Enteropathogenes	National School of Public Health
Hungary	Department of Phage-Typing and Molecular Epidemiology	National Public Health Institute
Italy	Department of Food Safety, Nutrition and Veterinary Public Health	Istituto Superiore di Sanità
Latvia	National Microbiology Reference laboratory	Infectology Centre of Latvia
Luxembourg	Epidemiology and Microbial Genomics	Laboratoire National de Sante
The Netherlands	Centre for Infectious Research, Diagnostics and Screening	National Institute for Public Health and the Environment
Norway	National Reference Laboratory for Enteropathogenic Bacteria	Norwegian Institute of Public Health
Slovenia	Department for Public Health Microbiology	National Laboratory of Health, Environment and Food
Spain	*Neisseria*, *Listeria* and *Bordetella* Unit	National Centre for Microbiology, Instituto de Salud Carlos III
Sweden	Microbiology	Folkhälsomyndigheten
UK	Gastrointestinal Bacteria Reference Unit	Public Health England

Annex 2. Participation overview EQA-5 and 6

Laboratory	2017 to 2018 (EQA-5)					2018 to 2019 (EQA-6)				
	All[#]	Serotyping		Cluster		All[#]	Serotyping		Cluster	
		Conventional	Molecular	PFGE	WGS		Conventional	Molecular	PFGE	WGS
19	X		X	X	X	X		X	X	X
35	X		X		X	X		X		X
49	X	X								
56	X	X	X		X	X	X	X	X	X
70	X		X		X	X		X		X
88	X		X							
96	X		X							
100	X	X	X	X		X	X	X		
105	X		X	X	X	X		X		X
108	X		X		X	X		X		X
129	X		X		X	X		X		X
130	X		X			X		X	X	
135*	X				X	X				X
138	X			X		X			X	
141	X	X	X	X	X	X	X	X	X	X
142	X	X	X	X	X	X	X	X	X	X
143	X		X			X		X		
144	X		X		X	X		X	X	
145	X	X	X	X						
146	X		X		X	X				X
149						X				X
Number of participants	**20**	6	17	7	12	**17**	4	13	7	12

[#]: participating in at least one part
**: previously laboratory 77.*

Annex 3. Reason(s) for participating in EQA

	Serotyping				Cluster			
Lab ID	**Accreditation needs**	**Institute policy**	**National policy**	**Enhance typing quality**	**Accreditation needs**	**Institute policy**	**National policy**	**Enhanced typing quality**
141	X	X			X	X		X
142	X				X			
19		X		X				
129		X		X		X		X
143	X				*No need for accreditation			
35		X		X		X		X
130		X	X	X		X	X	X
144	X	X	X	X	X	X	X	X
56				X				X
138	*Lack of laboratory capacity							X
70				X				X
149	*Lack of financial means				X	X		X
100		X		X	*Lack of laboratory capacity			
105	X			X	X			X
108	X							X
135	*							X
146	*Serotyping inferred from WGS				X			X
Number of participants	**6**	**6**	**2**	**8**	**6**	**6**	**2**	**13**

1: selected as reason for participating
**: reasons given when not participating.*

Annex 4. Serotyping result scores

Conventional serotyping

Isolate number	Provider	Laboratory ID				Total score
		56	100	141	142	
Sero1[#1]	1/2b	1/2b	1/2b	1/2b	1/2b	100
Sero2[#2]	4b	4b	4b	4b	4b	100
Sero3	1/2a	1/2a	1/2a	1/2a	1/2a	100
Sero4	1/2b	1/2b	1/2b	1/2b	1/2b	100
Sero5[#1]	1/2b	1/2b	1/2b	1/2b	1/2b	100
Sero6	1/2c	1/2c	1/2c	1/2c	1/2c	100
Sero7	4b	4b	4b	4b	4b	100
Sero8						
Sero9	1/2a	1/2a	1/2a	1/2a	1/2a	100
Sero10	1/2a	1/2a	1/2a	1/2a	1/2a	100
Sero11[#2]	4b	4b	4b	4b	4b	100
Total score		100	100	100	100	

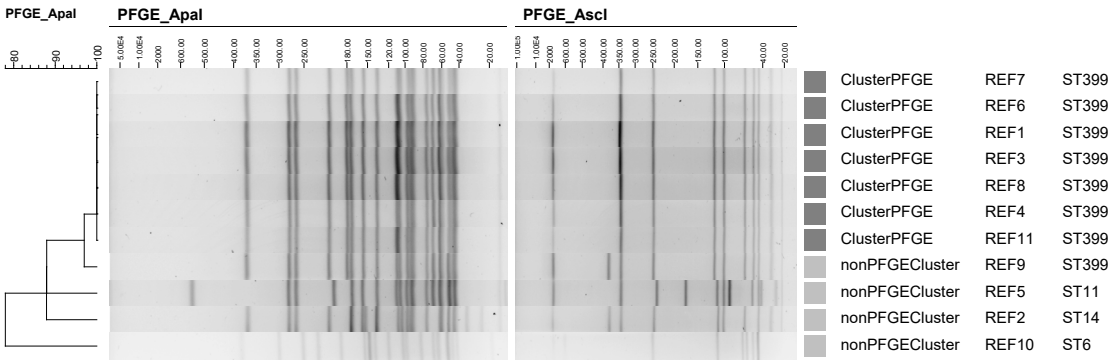
Isolate Sero8 were excluded as the were unstable.

Molecular serotyping

Isolate number	Provider	Laboratory ID													Total
		19	35	56	70	100	105	108	129	130	141	142	143	144	
Sero1[#1]	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	100
Sero2[#2]	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	100
Sero3	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIc	IIa	IIa	IIa	IIa	92
Sero4	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	100
Sero5[#1]	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	100
Sero6	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	100
Sero7	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	100
Sero8	IIb	IIb	IIb	IVb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	Untypable	IIb	IIb	85
Sero9	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIc	IIa	IIa	IIa	IIa	92
Sero10	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIc	IIa	IIa	IIa	IIa	92
Sero11[#2]	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	100
Total score		100	100	91	100	100	100	100	100	73	100	91	100	100	

Pink: incorrect
Purple: repeat isolates in EQA-1 to 6
#: set of technical duplicates 1 and 2.

Annex 5. EQA provider cluster analysis based on PFGE-derived data



Cluster of closely related isolates: REF1, REF3, REF4, REF6, REF7, REF8 and REF11

Annex 6. Reported cluster of closely related isolates based on PFGE-derived data

Lab ID	Reported cluster	Corresponding REF isolates	Correct
Provider		REF1, REF3, REF4, REF6, REF7, REF8 and REF11 (1 and 6 technical duplicates)	
19	2065, 2100, 2108, 2093, 2491, 2702	REF7, REF6, REF4, REF1, REF3, REF11, REF8	Yes
56	2377, 2463, 2488, 2553, 2575, 2726	REF8, REF3, REF7, REF1, REF6, REF4, REF11	Yes
130	2115, 2161, 2284, 2423, 2673, 2715	REF7, REF3, REF1, REF8, REF4, REF11, REF6	Yes
138	2353, 2433, 2487, 2684, 2938, 2941	REF4, REF11, REF3, REF7, REF1, REF8, REF6	Yes
141	2216, 2300, 2397, 2424, 2567, 2606	REF4, REF1, REF6, REF8, REF3, REF11, REF7	Yes
142	2262, 2290, 2412, 2845, 2902, 2942	REF7, REF4, REF8, REF3, REF6, REF1, REF11	Yes
144	2249, 2328, 2486, 2540, 2685, 2806, 2953, 2395	REF3, REF7, REF4, REF2, REF6, REF9, REF8, REF11, REF1	No

Annex 7. Reported band differences

			Laboratory ID						
Isolate number	ST	Expected *ApaI* bands	19	56	130	138	141	142	144
REF1#‡	399	16	16	16	9999	16	16	16	16
REF2	14	Clearly unrelated profile	16	14	9999	15	16	14	16
REF3‡	399	16	16	16	9999	16	16	16	16
REF4‡	399	16	16	16	9999	16	16	16	16
REF5	11	Clearly unrelated profile	16	14	9999	16	16	16	15
REF6#‡	399	16	16	16	9999	16	16	16	16
REF7‡	399	16	16	16	9999	16	16	16	16
REF8‡	399	16	16	16	9999	16	16	16	16
REF9	399	17	17	16	9999	16	17	16	16
REF10	6	Clearly unrelated profile	17	17	9999	14	17	16	14
REF11‡	399	16	16	16	9999	16	16	16	16

			Laboratory ID						
Isolate number	ST	Shared *ApaI* bands	19	56	130	138	141	142	144
REF1#‡	399	16	16	16	9999	16	16	16	16
REF2	14	Clearly unrelated profile	14	14	9999	14	9999	13	16
REF3‡	399	16	16	16	9999	16	16	16	16
REF4‡	399	16	16	16	9999	16	16	16	16
REF5	11	Clearly unrelated profile	13	13	9999	13	9999	11	4
REF6#‡	399	16	16	16	9999	16	16	16	16
REF7‡	399	16	16	16	9999	16	16	16	16
REF8‡	399	16	16	16	9999	16	16	16	16
REF9	399	16	16	16	9999	16	16	16	16
REF10	6	Clearly unrelated profile	4	6	9999	6	9999	2	4
REF11‡	399	16	16	16	9999	16	16	16	16

			Laboratory ID						
Isolate number	ST	Expected *AscI* bands	19	56	130	138	141	142	144
REF1#‡	399	9	9	8	9	9	8	8	9
REF2	14	Clearly unrelated profile	9	8	9	9	8	8	9
REF3‡	399	9	9	8	9	9	8	8	9
REF4‡	399	9	9	8	9	9	8	8	9
REF5	11	Clearly unrelated profile	10	10	12	9	10	10	0
REF6#‡	399	9	9	8	9	9	8	8	9
REF7‡	399	9	9	8	9	9	8	8	9
REF8‡	399	9	9	8	9	9	8	8	9
REF9	399	10	10	9	10	10	9	9	10
REF10	6	Clearly unrelated profile	8	8	7	8	8	8	8
REF11‡	399	9	9	8	9	9	8	8	9

			Laboratory ID						
Isolate number	ST	Shared *AscI* bands	19	56	130	138	141	142	144
REF1#‡	399	9	9	8	9	9	8	8	9
REF2	14	Clearly unrelated profile	7	5	7	7	9999	6	7
REF3‡	399	9	9	8	9	9	8	8	9
REF4‡	399	9	9	8	9	9	8	8	9
REF5	11	Clearly unrelated profile	4	6	6	5	9999	5	0
REF6#‡	399	9	9	8	9	9	8	8	9
REF7‡	399	9	9	8	9	9	8	8	9
REF8‡	399	9	9	8	9	9	8	8	9
REF9	399	9	9	8	9	9	8	8	9
REF10	6	Clearly unrelated profile	3	6	4	5	9999	5	3
REF11‡	399	9	9	8	9	9	8	8	9

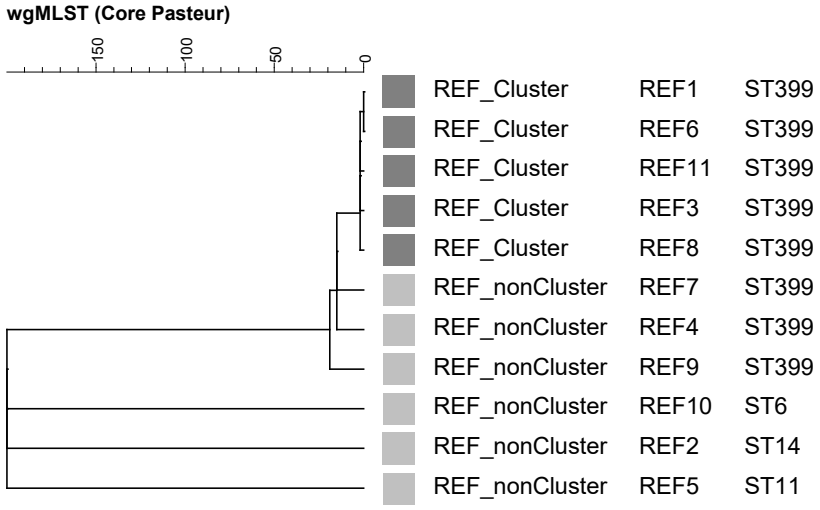
‡: cluster identification of closely related isolates (based on PFGE-derived data)
#: technical duplicate
ST: sequence type
9999: not reported by laboratory.

Annex 8. Reported sequencing details

Sequencing performed	Protocol (library prep)	Commercial kit	Sequencing platform
In own laboratory	Commercial kits	Nextera XT kit - Illumina*	MiniSeq - Illumina
In own laboratory	Commercial kits	Kapa HyperPlus (Kapa Biosystems)	MiSeq
In own laboratory	Commercial kits	Nextera	MiSeq
In own laboratory	Commercial kits	Illumina Nextera XT	HiSeq 2500
In own laboratory	Commercial kits	Nextera XT DNA Library Prep Kit	MiSeq
In own laboratory	Commercial kits	Nextera XT	MiSeq
Externally	Commercial kits	Nextera xt	HiSeq 2500
In own laboratory	Commercial kits	NEBNext® Fast DNA Fragmentation & Library Prep Set for Ion Torrent, New England Biolabs**	Ion GeneStudio S5 System
In own laboratory	Commercial kits	NexteraXT	NextSeq
In own laboratory	Commercial kits	Illumina Nextera XT library Prep Kit	NextSeq
In own laboratory	Commercial kits	Ion Xpress TM Plus Fragment Library Kit for AB Library Builder TM System	IonTorrent S5XL
In own laboratory	Commercial kits	Nextera XT DNA Library Preparation Kit, Illumina	MiSeq

** We use half the volume of reagents for each step of the protocol*
***The shearing was carried out for 15 minutes at 25 degrees instead than from 20 minutes. The reason is that we used a 400bp sequencing protocol. The reaction was performed in half of the volume suggested by the manufacturer's instructions, starting from 100 ng of DNA.*

Annex 9. EQA provider cluster analysis based on WGS-derived data



Single linked dendrogram of core genome multi-locus sequence typing (cgMLST) profiles of Listeria *EQA-6 isolates (cgMLST, Pasteur, Moura et al 2016).*
Analysed in BioNumerics: maximum distance of 200 exceeded; results clipped.
Dark grey: cluster isolates
Light grey: outside cluster isolates.

Annex 10. Reported cluster of closely related isolates based on WGS-derived data

Lab ID	Reported cluster	Corresponding to REF isolates	Correct
Provider		REF1, REF3, REF6, REF8 and REF11 (1 and 6 technical duplicates)	
19	2108, 2065, 2093, 2702, 2491	REF1, REF6, REF3, REF8, REF11	Yes
35	2098, 2562, 2716, 2788, 2844	REF1, REF3, REF6, REF11, REF8	Yes
56	2197, 2377, 2463, 2488, 2553, 2575, 2726	REF8, REF3, REF7, REF1, REF6, REF4, REF11	No
70	2102, 2891, 2548, 2474, 2743	REF3, REF11, REF8, REF6, REF1	Yes
105	2128, 2202, 2130, 2557, 2859	REF6, REF1, REF3, REF11, REF8	Yes
108	2464, 2022, 2177, 2092, 2509	REF6, REF1, REF8, REF11, REF3	Yes
129	2047, 2070, 2071, 2280, 2400	REF11, REF1, REF8, REF6, REF3	Yes
135	2077, 2330, 2794, 2837, 2977	REF6, REF3, REF11, REF1, REF8	Yes
141	2216, 2300, 2397, 2424, 2567	REF1, REF6, REF8, REF3, REF11	Yes
142	2290, 2412, 2845, 2902, 2942	REF8, REF3, REF6, REF1, REF11	Yes
146	2079, 2204, 2912, 2635, 2979	REF1, REF6, REF11, REF8, REF3	Yes
149	2346, 2850, 2856, 2865, 2974	REF8, REF3, REF11, REF1, REF6	Yes

Annex 11. Reported SNP distance and allelic differences

SNP distances

			Laboratory ID		
Isolate number	ST	Provider	56	108	146
REF1#‡	399	2	548	0¤	0¤
REF2	14	7953	10177	9999	9999
REF3‡	399	3	497	3	3
REF4	399	38	271	40	9999
REF5	11	13525	16239	9999	9999
REF6#‡	399	0	727	0	0
REF7	399	34	0¤	32	9999
REF8‡	399	2	1299	2	2
REF9	399	51	662	831	9999
REF10	6	115961	40660	9999	9999
REF11‡	399	0¤	450	4	4

Allelic distances

			Laboratory ID								
Isolates number	ST	Provider	19	35	70	105	129	135	141	142	149
REF1#‡	399	0	2	1	1	1	1	1	1	0	1
REF2	14	511	513	489	9999	491	413	537	492	514	493
REF3‡	399	2	0¤	0¤	0¤	0¤	0	0¤	0¤	2	0¤
REF4	399	18	16	17	9999	17	13	29	17	12	17
REF5	11	653	655	648	9999	648	553	674	649	658	651
REF6#‡	399	0¤	2	1	1	1	2	1	1	0¤	1
REF7	399	17	15	14	9999	14	10	14	15	10	14
REF8‡	399	2	2	1	1	1	1	1	1	1	1
REF9	399	23	22	26	9999	26	22	51	26	24	26
REF10	6	1690	1705	1669	9999	1672	1468	1672	1675	1736	1679
REF11‡	399	2	3	1	1	1	0¤	1	1	2	1

**: additional analysis*
‡: closely related isolates
#: technical duplicate isolate
¤: isolate used as cluster representative by participant
9999: isolates not included in analysis by participant
ST: sequence type.

Annex 12. Additional reported QC parameters

Lab ID	1		2		3	
	Parameter	Threshold	Parameter	Threshold	Parameter	Threshold
146	% non consensus bases(MLST alleles)	<15%	Min consensus depth (MLST alleles)	>5 reads	-	-
19	# contigs	Available from QC evaluation, but no threshold	N50	Available from QC evaluation, but no threshold	Read length	Available from QC evaluation, but not employed as such (usually not a problem for Illumina)
141	Number of contigs	200 bases (contigs shorter than 200 bases have to be ignored)	Fast QC	Per base sequence quality, per base sequence content, per sequence quality score	-	-

Annex 13. Calculated qualitative/quantitative parameters

		Laboratory 19										
Quali-/Quantitative	Ranges*	2100	2702	2322	2065	2264	2692	2491	2802	2937	2108	2093
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		3.14	3.97	4.13	3.15	4.33	2.92	3.75	2.59	4.01	2.76	3.57
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	2.9	3.0	3.0	3.0	2.9	3.0	3.0	2.9	3.0	3.0	3.0
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	23	25	24	28	23	40	21	17	25	21	22
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	0	1	0	0	0	0	0
Average coverage	{>50}	121.1	176.9	144.3	131.7	149.2	116.5	159.2	137.2	132.0	129.2	178.4
Number of reads (x1000)		1353.9	2090.5	1696.3	1498.4	1728.0	1338.5	1861.2	1 96.2	1534.5	1450.7	2086.8
Number of trimmed reads (x1000)		1341.6	2074.1	1682.8	1484.6	1713.8	1325.4	1847.1	1483.1	1522.1	1437.7	2069.2
Maximum read length		151	151	151	151	151	151	151	151	151	151	151
Mean read length		137.6	133.7	133.4	137.4	133.3	138.1	134.4	138.1	134.4	139.3	135.1
Read insert size		285.1	234.9	253.2	275.6	261.0	286.8	250.4	279.4	259.0	282.0	239.2
Insert size StdDev		126.9	108.8	121.1	121.4	123.6	124.2	120.1	123.9	119.4	119.2	108.9
N50 (kbp)		342.8	318.7	331.3	331.5	344.6	344.4	331.5	291.5	456.1	344.0	361.1
N75 (kbp)		251.3	124.7	251.5	110.4	251.3	124.7	144.0	223.0	169.9	179.4	142.9

		Laboratory 35										
Quali-/Quantitative	Ranges*	2098	2086	2118	2297	2422	2562	2446	2716	2582	2788	2844
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		1.92	1.74	1.23	1.67	1.71	2.08	2.06	1.72	1.83	1.72	1.72
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	2.9	2.9	3.0	2.9	3.0	2.9	3.0	3.0	3.0	3.0
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	34	38	31	62	32	53	33	42	36	41	47
Number of contigs [0–25] x min. coverage	{<1000}	0	0	17	0	0	2	0	0	0	0	0
Average coverage	{>50}	57.8	97.6	62.6	60.9	86.1	58.2	88.3	48.5	70.4	70.9	54.4
Number of reads (x1000)		822.6	1683.6	926.7	894.1	1415.1	830.7	1456.4	641.0	1070.2	1061.0	745.7
Number of trimmed reads (x1000)		821.7	1681.6	925.8	893.2	1413.7	829.8	1454.6	640.3	1069.0	1059.8	744.9
Maximum read length		151	151	151	151	151	151	151	151	151	151	151
Mean read length		146.4	142.9	147.0	146.3	146.7	145.1	144.9	147.8	146.8	147.9	147.9
Read insert size		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Insert size StdDev		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
N50 (kbp)		192.6	254.9	327.1	124.7	210.3	108.7	178.7	124.7	210.3	150.3	108.7
N75 (kbp)		100.9	150.0	107.7	79.7	116.3	69.3	122.9	83.3	94.9	94.9	86.7

		Laboratory 56										
Quali-/Quantitative	**Ranges***	**2068**	**2160**	**2197**	**2488**	**2377**	**2463**	**2553**	**2536**	**2575**	**2726**	**2655**
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		2.7	3	3.08	2.72	2.77	2.6	2.89	2.14	1.89	2	1.93
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	0.5	1	1.3	1.9	2.3	1.8	1.4	1.5	2.5	1.8	1.6
Length [0-25] x min. coverage (Mbp)	{<0.25}	0.8	0.8	0.4	0.5	0.2	0.3	0.5	0.3	0	0	0
Number of contigs at 25x min. coverage	{>0}	277	448	720	717	912	876	738	831	1 099	1 050	962
Number of contigs [0–25] x min. coverage	{<1000}	541	468	248	301	104	191	313	197	0	0	0
Average coverage	{>50}	21	24.7	28.4	29.1	35.5	30.8	26.5	27.8	80.8	60.9	59.9
Number of reads (x1000)		241.8	286.7	377	357	470.4	399.6	355.7	382.8	1 130.8	948.9	987.5
Number of trimmed reads (x1000)		225.5	265.9	350.5	333.7	439.7	375	331.6	359.4	1 062.6	897.1	931.9
Maximum read length		280	280	280	280	280	280	280	280	280	280	280
Mean read length		147.7	142.8	143.7	154	152.1	155	150.7	158.9	172.2	174.5	177.5
Read insert size		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Insert size StdDev		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
N50 (kbp)		1.2	1.6	1.3	2.4	2.3	1.6	1.4	1.3	2.1	1.3	1.2
N75 (kbp)		0.8	1	0,9	1.4	1.3	1	0.9	0.8	1.3	0.9	0.8

		Laboratory 70										
Quali-/Quantitative	**Ranges***	**2317**	**2102**	**2429**	**2247**	**2431**	**2474**	**2513**	**2548**	**2891**	**2588**	**2743**
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		1.6	1.94	1.28	1.92	1.96	1.79	2.02	2.08	1.5	2.23	2.01
Length at 25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	2.9	2.9	3.0	3.0	3.0	3.0	3.0	2.9	3.0
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	100	35	32	37	28	19	21	24	20	26	29
Number of contigs [0–25] x min. coverage	{<1000}	20	1	1	2	0	0	0	0	0	0	1
Average coverage	{>50}	34.6	104.8	78.6	95.7	76.3	105.1	95.5	103.9	71.4	123.6	119.0
Number of reads (x1000)		365.4	1122.5	801.7	1008.6	808.7	1117.7	1019.3	1111.8	745.9	1313.9	1275.4
Number of trimmed reads (x1000)		361.1	1109.6	796.4	997.6	799.7	1107.3	1007.5	1097.4	742.8	1299.5	1261.8
Maximum read length		151	151	151	151	151	151	151	151	151	151	151
Mean read length		146.8	145.2	144.9	145.3	145.3	145.7	145.1	145.3	146.1	143.8	145.0
Read insert size		343.4	321.7	302.5	330.6	345.3	320.6	341.5	321.5	344.0	309.3	316.0
Insert size StdDev		105.6	111.0	100.1	113.3	117.8	107.4	115.4	107.1	122.2	112.3	110.2
N50 (kbp)		50.1	125.0	232.4	228.7	224.2	429.1	436.1	245.4	456.1	264.4	241.3
N75 (kbp)		30.2	108.0	108.8	107.2	129.1	331.5	231.3	129.1	331.5	129.1	109.5

		Laboratory 105										
Quali-/Quantitative	**Ranges***	**2130**	**2202**	**2128**	**2557**	**2512**	**2299**	**2713**	**2859**	**2893**	**2956**	**2854**
Number of genera detected	{1}	1	1	1	1	2	1	1	1	1	1	1
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		8.68	7.05	5.5	4.81	10.38	7.46	2.02	1.68	3.38	4.25	2.8
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	5.8	2.9	3.0	3.0	2.9	3.0	2.9
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	22	22	21	20	151	18	22	19	16	22	20
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	82	0	0	0	0	0	0
Average coverage	{>50}	195.2	205.9	284.0	226.6	95.7	215.3	229.8	185.4	230.4	213.7	217.2
Number of reads (x1000)		2772.8	2782.5	3732.7	2844.0	2788.5	2903.9	2606.9	2067.6	2697.9	2618.2	2501.1
Number of trimmed reads (x1000)		2744.0	2752.7	3674.4	2814.9	2759.0	2872.7	2577.5	2043.7	2668.5	2590.6	2471.0
Maximum read length		151	151	151	151	151	151	151	151	151	151	151
Mean read length		115.7	120.7	125.2	129.6	132.7	119.2	141.9	143.5	133.4	132.1	140.0
Read insert size		195.7	201.1	205.4	227.3	231.7	194.8	263.9	285.1	245.0	234.6	268.1
Insert size StdDev		109.1	108.8	101.7	117.9	116.6	102.9	104.8	114.1	122.3	117.8	122.2
N50 (kbp)		424.7	424.7	424.7	360.6	80.4	456.1	436.1	429.1	510.3	361.2	360.6
N75 (kbp)		143.0	143.0	143.0	179.4	29.8	250.3	251.7	331.5	257.3	169.9	214.7

		Laboratory 108										
Quali-/Quantitative	Ranges*	2022	2050	2092	2177	2194	2464	2207	2509	2510	2872	2535
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		2.17	1.98	2.09	2.0	1.47	4.24	1.86	2.29	1.74	2.46	1.88
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	2.9	2.9	2.9	3.0	2.9	2.5	3.0	2.9	2.9	3.0	2.9
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	392#	419#	410#	291#	282#	638#	394#	367#	386#	305#	375#
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	0	183	0	0	0	0	0
Average coverage	{>50}	65.5	62.0	63.0	90.6	89.1	29.3	65.2	52.1	55.1	52.4	62.9
Number of reads (x1000)		687.8	642.8	687.0	954.0	879.6	441.1	670.6	518.3	589.3	508.6	631.4
Number of trimmed reads (x1000)		617.2	572.2	610.3	853.8	790.2	392.9	606.7	466.7	520.0	451.6	567.0
Maximum read length		324	324	321	321	321	194	319	320	324	318	321
Mean read length		209.1	214.3	205.4	207.3	209.4	146.9	215.5	214.3	205.0	216.3	215.4
Read insert size		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Insert size StdDev		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
N50 (kbp)		10.8	11.1	10.9	15.2	15.9	4.5	11.6	12.6	11.9	16.8	11.9
N75 (kbp)		6.4	5.8	6.0	8.8	8.3	2.5	6.4	7.0	6.3	8.4	6.9

		Laboratory 129										
Quali-/Quantitative	Ranges*	2071	2047	2070	2254	2569	2280	2733	2144	2885	2911	2400
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		0.73	0.92	1.17	0.77	0.57	0.69	0.59	0.6	0.58	0.96	0.71
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	2.9	3.0	3.0	3.0	3.0	2.9	2.9	3.0
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	18	17	18	18	35	22	24	22	16	17	19
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	0	0	0	0	0	0	0
Average coverage	{>50}	137.3	225.1	198.8	119.3	126.3	118.5	133.2	133.3	133.6	134.5	159.8
Number of reads (x1000)		1523.3	2399.7	2134.0	1324.1	1346.9	1253.2	1384.3	1490.4	1389.6	1411.6	1673.1
Number of trimmed reads (x1000)		1443.1	2369.1	2086.4	1238.4	1330.1	1230.6	1373.4	1404.0	1339.1	1374.8	1659.1
Maximum read length		151	151	151	151	151	151	151	151	151	151	151
Mean read length		145.4	147.2	147.0	144.5	147.1	146.4	147.8	145.0	146.4	146.9	147.7
Read insert size		388.7	369.1	385.2	424.5	430.1	398.4	419.4	389.3	400.8	400.0	401.8
Insert size StdDev		145.0	150.5	149.6	157.6	164.1	149.0	156.4	145.9	151.1	151.9	155.3
N50 (kbp)		331.5	456.1	456.1	361.3	441.8	331.5	330.7	436.1	295.0	318.7	330.8
N75 (kbp)		210.8	331.5	331.5	242.9	124.7	129.1	125.0	251.6	257.3	242.9	210.8

		Laboratory 135										
Quali-/Quantitative	**Ranges***	**2040**	**2330**	**2077**	**2385**	**2408**	**2441**	**2529**	**2977**	**2794**	**2837**	**2670**
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		2.71	3.02	3.04	3.21	2.12	2.75	2.82	2.92	2.91	3.01	2.92
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	2.9	3.0	2.9	3.0	3.0	3.0	2.9
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	35	19	16	21	16	21	20	18	19	19	21
Number of contigs [0–25] x min. coverage	{<1 000}	0	0	0	0	0	0	0	0	0	0	0
Average coverage	{>50}	176.6	150.2	262.5	259.3	286.5	300.3	337.8	167.0	190.6	297.5	307.1
Number of reads (x1000)		2453.7	2036.5	3660.5	3596.3	3806.4	4196.7	4665.4	2260.2	2600.5	4161.4	4236.3
Number of trimmed reads (x1000)		2417.5	2013.4	3614.8	3556.1	3755.3	4150.9	4600.8	2232.8	2563.9	4108.7	4177.0
Maximum read length		126	126	126	126	126	126	126	126	126	126	126
Mean read length		116.5	116.2	116.2	115.7	116.9	116.1	116.7	116.9	116.7	116.3	116.4
Read insert size		278.7	267.3	268.7	258.9	282.0	258.7	274.7	279.6	279.4	269.0	275.2
Insert size StdDev		182.3	173.0	175.1	167.8	181.8	164.8	176.8	181.4	181.9	174.4	179.2
N50 (kbp)		420.3	424.1	456.0	456.0	440.8	435.7	425.8	331.4	331.4	456.0	331.4
N75 (kbp)		164.9	215.2	331.4	169.6	257.3	251.6	215.3	245.8	215.2	330.8	210.3

		Laboratory 141										
Quali-/Quantitative	**Ranges***	**2397**	**2695**	**2216**	**2300**	**2767**	**2219**	**2903**	**2424**	**2104**	**2567**	**2606**
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		1.56	1.78	1.57	1.51	1.76	1.53	1.46	1.25	1.51	1.68	2.1
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	2.6	2.9	3.0	3.0	3.0	3.0	3.0	3.0	2.9	3.0	2.9
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	119	47	74	30	44	44	50	29	25	28	58
Number of contigs [0–25] x min. coverage	{<1000}	58	3	5	1	2	1	1	0	0	0	8
Average coverage	{>50}	29.3	58.6	43.1	61.3	48.3	64.6	64.5	58.5	53.3	62.7	54.6
Number of reads (x1000)		186.6	367.5	272.9	386.0	323.4	422.6	433.6	363.2	335.1	396.0	352.7
Number of trimmed reads (x1000)		177.2	360.9	265.4	379.7	314.7	414.3	424.4	352.3	329.4	385.0	339.9
Maximum read length		301	301	301	301	301	301	301	301	301	301	301
Mean read length		245.7	235.9	244.6	244.2	231.4	236.0	234.6	250.9	240.8	246.0	238.5
Read insert size		423.4	335.1	352.2	340.3	316.4	321.8	317.3	377.4	331.9	371.2	378.0
Insert size StdDev		158.3	132.2	139.5	135.9	130.0	131.7	128.4	147.6	132.5	147.9	149.5
N50 (kbp)		29.7	143.3	62.1	204.5	145.9	143.7	183.0	348.6	267.4	242.8	77.9
N75 (kbp)		17.5	65.4	36.3	129.3	90.4	88.4	92.3	88.4	99.2	107.0	48.9

		Laboratory 142										
Quali-/Quantitative	**Ranges***	**2054**	**2262**	**2076**	**2170**	**2290**	**2845**	**2078**	**2366**	**2412**	**2902**	**2942**
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		0.47	0.77	0.88	1.7	0.85	0.76	0.66	0.87	0.68	0.76	1.01
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	2.9	2.9	3.0	2.9	3.0	3.0	3.0	3.0	3.0	3.0	3.0
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	23	15	24	23	12	21	16	12	13	13	31
Number of contigs [0–25] x min. coverage	{<1000}	3	0	0	4	0	3	0	1	0	0	5
Average coverage	{>50}	53.4	65.1	46.3	43.7	45.3	53.7	66.5	59.9	70.6	63.2	45.9
Number of reads (x1000)		331.7	410.5	313.7	316.5	294.3	345.6	429.3	386.5	454.4	413.3	310.6
Number of trimmed reads (x1000)		323.5	403.2	302.1	298.3	286.0	337.5	420.8	378.4	443.8	400.2	295.7
Maximum read length		251	251	251	251	251	251	251	251	251	251	251
Mean read length		239.5	240.2	235.1	217.4	238.6	239.5	238.4	239.2	240.1	238.0	233.1
Read insert size		521.1	470.6	489.4	424.7	540.1	517.7	444.8	459.3	492.8	518.1	608.4
Insert size StdDev		169.4	168.4	161.2	143.9	177.8	171.2	170.1	163.9	172.1	168.8	204.0
N50 (kbp)		223.4	542.1	541.9	360.6	542.2	456.3	539.8	541.3	512.2	542.1	243.7
N75 (kbp)		158.2	448.7	442.9	180.8	456.3	110.6	436.2	456.3	456.3	456.3	109.4

		Laboratory 146										
Quali-/Quantitative	Ranges*	2635	2079	2518	2204	2320	2640	2646	2912	2979	2951	2950
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		1.44	2.4	0.89	1.34	0.92	1.06	0.71	1.47	2.06	1.18	1.51
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	2.9	2.9	2.9	3.0	3.0	3.0	2.9
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	19	20	31	21	45	60	17	21	21	37	19
Number of contigs [0–25] x min. coverage	{<1000}	0	0	1	0	6	4	0	0	0	0	0
Average coverage	{>50}	35.3	106.8	101.8	125.0	114.9	138.0	140.0	158.9	96.8	118.5	158.5
Number of reads (x1000)		580.1	1716.9	1622.1	2029.4	1824.5	2203.0	2208.5	2556.2	1551.3	1929.7	2516.5
Number of trimmed reads (x1000)		559.2	1677.9	1572.3	1949.1	1768.5	2145.6	2105.9	2480.0	1510.3	1874.1	2445.4
Maximum read length		101	101	101	101	101	101	101	101	101	101	101
Mean read length		99.1	97.8	99.3	99.1	99.3	99.1	99.0	99.1	98.3	99.1	99.0
Read insert size		351.2	286.7	329.4	321.0	334.1	308.3	367.4	338.3	308.6	330.2	327.3
Insert size StdDev		192.6	173.9	150.3	144.9	148.1	142.9	165.2	173.5	183.1	172.2	177.3
N50 (kbp)		331.4	456.0	267.7	331.4	118.8	90.2	440.8	331.4	331.6	331.6	331.6
N75 (kbp)		245.7	331.6	108.4	215.2	63.9	53.3	254.8	215.1	215.3	142.7	215.1

		Laboratory 149										
Quali-/Quantitative	Ranges*	2261	2346	2627	2220	2666	2850	2856	2865	2877	2974	2998
Number of genera detected	{1}	1	1	1	1	1	1	1	1	1	1	1
Detected species	*{Lm}*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
Unclassified reads (%)		2.38	2.65	1.78	2.32	2.31	2.34	2.28	2.35	2.4	2.57	3.03
Length at 25x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	2.9	3.0	3.0	3.0	3.0	3.0	3.1	3.0	3.0
Length [0–25] x min. coverage (Mbp)	{<0.25}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25x min. coverage	{>0}	346	364	11	364	345	18	16	323	427	499	388
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	0	0	0	0	0	0	0
Average coverage	{>50}	159.9	134.0	160.2	173.9	155.2	179.3	200.9	177.5	171.5	145.6	157.8
Number of reads (x1000)		1254.3	1107.8	1169.2	1434.7	1237.7	1340.7	1507.9	1410.0	1459.4	1256.5	1343.2
Number of trimmed reads (x1000)		1189.2	1055.8	1103.9	1374.6	1173.4	1257.8	1436.1	1339.9	1395.2	1179.1	1277.1
Maximum read length		301	301	301	301	301	301	301	301	301	301	301
Mean read length		209.8	201.1	210.3	200.5	209.4	214.7	210.5	208.7	200.4	200.1	194.5
Read insert size		290.5	278.1	291.7	274.0	288.2	301.6	288.8	284.5	274.1	275.0	268.7
Insert size StdDev		127.4	122.7	131.7	120.5	128.9	135.9	129.2	125.9	118.9	118.4	115.2
N50 (kbp)		13.4	12.1	520.2	12.1	12.7	479.6	479.8	15.9	11.5	9.2	11.5

Quality assessment made by the EQA-provider in-house quality control pipeline.
**: indicative QC ranges*
NA: not available
#: QC values unreliable due to assembly issues for Ion Torrent data
Lm: L. monocytogenes

Annex 14. EQA-5 laboratory questionnaire

This is a preview of all the fields and questions available.

Please keep in mind that, depending on your answers in the questionnaire, you will not necessarily have to answer all the questions.

1. Listeria EQA-6 2018

Dear participant,

Welcome to the sixth External Quality Assessment (EQA-6) scheme for typing of *Listeria* in 2018–2019. Please note that most of the fields are required to be filled in before the submission can be completed. Any comments can be written at the end of the form. You are always welcome to contact us at mailto:list.eqa@ssi.dk.

Please start by filling in your country, your laboratory name and your lab ID.

Available options in this submission form include:

- Click 'Options' and 'Pause' to save your results and finish at a later time (using the same link)
- Click 'Options' and 'Print' to print your answers. This can be done at any time, but before pressing 'Submit results'
- Click 'Previous' to go back to the questions you have already answered
- Click 'Options' and 'Go to..' to go back to a specific page number

Note: After pressing 'Submit results' you will not be able to review your results.

2. Country

☐ Austria
☐ Belgium
☐ Denmark
☐ Finland
☐ France
☐ Germany
☐ Greece
☐ Hungary
☐ Italy
☐ Latvia
☐ Luxembourg
☐ Norway
☐ Slovenia
☐ Spain
☐ Sweden
☐ Netherlands
☐ Turkey
☐ UK

3. Institute name

4. Laboratory name

5. Laboratory ID

Consisting of country code (two letters) lab ID on the vial, e.g. DK_SSI

6. E-mail

7. Serotyping of *Listeria*

8. Submitting results

(State one answer only)

☐ Online here (please fill in the isolate ID's in the following section) – Go to 9
☐ Did not participate in the serotyping part – Go to 14

9. Serotyping isolate ID's

Please enter the isolate ID (4 digits)

Listeria

Strain 1 ___
Strain 2 ___
Strain 3 ___
Strain 4 ___
Strain 5 ___
Strain 6 ___
Strain 7 ___
Strain 8 ___
Strain 9 ___
Strain 10 ___
Strain 11 ___

10. Submitting results – Serotyping of *Listeria*

(State one answer only)

☐ Both molecular and conventional serotyping – Go to 11
☐ Molecular serotyping – Go to 11
☐ Conventional serotyping – Go to 13

11. Method used for molecular serotyping of *Listeria*

(State one answer only)

☐ PCR-based
☐ WGS-based

12. Results for serotyping *Listeria* – molecular serotyping

Please select the serotype

Isolate	Molecular serotype					
	IIa	IIb	IIc	IVb	L	Untypeable
Isolate 1						
Isolate 2						
Isolate 3						
Isolate 4						
Isolate 5						
Isolate 6						
Isolate 7						
Isolate 8						
Isolate 9						
Isolate 10						
Isolate 11						

13. Results for serotyping *Listeria* – conventional serotyping

Please select the serotype

Isolate		Conventional serotype												
Isolate 1	1/2a	1/2b	1/2c	3a	3b	3c	4a	4ab	4b	4c	4d	4e	7	Autoagglutinable	Untypeable
Isolate 2															
Isolate 3															
Isolate 4															
Isolate 5															
Isolate 6															
Isolate 7															
Isolate 8															
Isolate 9															
Isolate 10															
Isolate 11															

14. Submitting cluster results

(State one answer only)

☐ Cluster analyses based on PFGE and/or WGS – Go to 15
☐ Did not participate in the cluster part – Go to 115

15. Cluster isolates ID's

Please enter the cluster isolate ID (4 digits)

Isolate	Cluster isolate ID
Isolate 1	
Isolate 2	
Isolate 3	
Isolate 4	
Isolate 5	
Isolate 6	
Isolate 7	
Isolate 8	
Isolate 9	
Isolate 10	
Isolate 11	

16. Submitting cluster results

(State one answer only)

☐ Cluster analysis based on PFGE – Go to 17
☐ Do not wish to submit any cluster results based on PFGE analysis – Go to 23

17. Cluster analysis based on PFGE data

18. Please list the ID for the isolates included in the cluster of closely related isolates detected by PFGE combining *ApaI*- and *AscI*- results:

Please use semicolon (;) to separate the ID's

19. Select a representative isolate with the cluster profile detected by PFGE:

Indicate the isolate ID

20. *ApaI* – Total number of bands (>33kb) in the selected representative cluster isolate

21. *AscI* – Total number of bands (>33kb) in the selected representative cluster isolate

22. Results for cluster analysis – PFGE (*ApaI* and *AscI*)

Please use 9999 for not analysed

	ApaI – Total number of bands (>33kb)	*ApaI* – Number of bands with same/shared position as the profile of the selected cluster isolate (>33kb)	*AscI* – Total number of bands (>33kb)	*AscI* – Number of bands with same/shared position as the profile of the selected cluster isolate (>33kb)
Isolate 1				
Isolate 2				
Isolate 3				
Isolate 4				
Isolate 5				
Isolate 6				
Isolate 7				
Isolate 8				
Isolate 9				
Isolate 10				
Isolate 11				

23. Submitting cluster results

(State one answer only)

☐ Cluster analysis based on WGS data – Go to 24
☐ Do not wish to submit any cluster results based on WGS data – Go to 115

24. Cluster analysis based on WGS data

25. Please select the analysis used to detect the cluster on data derived from WGS

As the basis for the cluster detection, only one data analysis can be reported. If more than one analysis is performed, please report later in this submission.

☐ SNP-based – Go to 27
☐ Allele-based – Go to 34
☐ Other – Go to 26

26. If another analysis is used, please describe your approach: – Go to 41

27. Please report the used SNP-pipeline (reference if publicly available or in-house pipeline)

28. Please select the approach used for the SNP analysis

(State one answer only)

☐ Reference-based – Go to 29
☐ Assembly-based – Go to 32

29. Reference genome used

Please indicate multilocus sequence type (e.g. ST8) and isolate ID (e.g. one of the isolates from the current cluster, ID of a public reference strain or an in-house isolate)

30. Please indicate the read mapper used (e.g. BWA, Bowtie2)

31. Please indicate the variant caller used (e.g. SAMtools, GATK)

32. Please indicate the assembler used (e.g. SPAdes, Velvet)

33. Please specify the variant caller used (e.g. NUCMER)

34. Please select tools used for the allele analysis

(State one answer only)

☐ BioNumerics – Go to 36
☐ SeqPhere – Go to 36
☐ BIGSdb-*Lm* – Go to 36
☐ Other – Go to 35

35. If another tool is used, please enter here:

36. Please indicate allele calling method

(State one answer only)

☐ Assembly-based and mapping-based – Go to 37
☐ Only assembly-based – Go to 37
☐ Only mapping-based – Go to 38

37. Please indicate the assembler used (e.g. SPAdes, Velvet)

38. Please select scheme used for the allele analysis

(State one answer only)

☐ Applied Math (wgMLST) – Go to 40
☐ Applied Math (cgMLST/Pasteur) – Go to 40
☐ Pasteur (cgMLST) – Go to 40
☐ Ruppitsch (cgMLST) – Go to 40
☐ Other – Go to 39

39. If another scheme (e.g. in-house) is used, please give a short description

40. Please report the number of loci in the used allelic scheme

41. Cluster detected by analysis on data derived from WGS

On this page, you have to report the results for the cluster detected by the selected analysis (e.g. SNP-based). If another additional analysis (e.g. allele-based or another SNP-based analysis) is performed, please report results later, but you will not be asked to submit the ID's for isolates in the cluster detected with the additional analysis.

42. Please list the ID's for the isolates included in the cluster

Please use semicolon (;) to separate the ID's

43. Select a representative isolate in the cluster

Indicate the isolate ID

44. Results for cluster analysis (e.g. SNP- or allele-based)

Please use 9999 for not analysed.

Isolate	ST	Distance (e.g. SNP) to the selected cluster isolate
Isolate 1		
Isolate 2		
Isolate 3		
Isolate 4		
Isolate 5		
Isolate 6		
Isolate 7		
Isolate 8		
Isolate 9		
Isolate 10		
Isolate 11		

45. Would you like to add results performed with another additional analysis on the data derived from the WGS?

e.g. If SNP based results are submitted, you can also report allele-based results or results from a second SNP analysis.

(State one answer only)

☐ Yes – Go to 46
☐ No – Go to 85

46. Please select the additional analysis used on data derived from WGS

(State one answer only)

☐ SNP-based – Go to 48
☐ Allele-based – Go to 55
☐ Other – Go to 47

47. If another analysis is used please describe your approach: - Go to 62

48. Please report the used SNP pipeline (reference if publicly available or in-house pipeline)

49. Please select the approach used for the SNP analysis reference genome used

(State one answer only)

☐ Reference-based – Go to 50
☐ Assembly-based – Go to 53

50. Reference genome used

Please indicate multilocus sequence type (e.g. ST8) and isolate ID (e.g. one of the isolates from the current cluster, ID of a public reference strain or an in-house isolate)

51. Please indicate the read mapper used (e.g. BWA, Bowtie 2)

52. Please indicate the variant caller used (e.g. SAMtools, GATK)

53. Please indicate the assembler used (e.g. SPAdes, Velvet)

54. Please specify the variant caller used (e.g. NUCMER)

55. Please select tool used for the allele analysis

☐ BioNumerics – Go to 57
☐ SeqPhere – Go to 57
☐ BIGSdb-*Lm* – Go to 57
☐ Other – Go to 56

56. If another tool is used, please list here:

57. Please indicate allele calling method

(State one answer only)

☐ Assembly based and mapping-based – Go to 58
☐ Only assembly-based – Go to 58
☐ Only mapping-based – Go to 59

58. Please indicate the assembler used (e.g. SPAdes, Velvet)

59. Please select scheme used for the allele analysis

(State one answer only)

☐ Applied Math (wgMLST) – Go to 61
☐ Applied Math (cgMLST/Pasteur) – Go to 61
☐ Pasteur (cgMLST) – Go to 61
☐ Ruppitsch (cgMLST) – Go to 61
☐ Other – Go to 60

60. If another scheme (e.g. in-house) is used, please give a short description

61. Please report the number of loci in the used allelic scheme

62. Additional analysis on data derived from WGS

63. Select a representative isolate in the cluster detected by the additional analysis

64. Results for the additional cluster analysis (e.g. SNP- or allele-based)

Please use 9999 for not analysed

Isolate	ST	Distance (e.g. SNP) to the selected cluster isolate
Isolate 1		
Isolate 2		
Isolate 3		
Isolate 4		
Isolate 5		
Isolate 6		
Isolate 7		
Isolate 8		
Isolate 9		
Isolate 10		
Isolate 11		

65. Would you like to add results performed with a third analysis on the data derived from the WGS?

e.g. If SNP based results are submitted you can also report allele based results or results from a second SNP analysis

(State one answer only)

☐ Yes – Go to 66
☐ No – Go to 85

66. Please select the third analysis used on data derived from WGS

☐ SNP-based – Go to 68
☐ Allele-based – Go to 75
☐ Other – Go to 67

67. If another analysis is used please describe your approach: - Go to 82

68. Please report the used SNP-pipeline (reference if publicly available or in-house pipeline)

69. Please select the approach used for the SNP analysis

(State one answer only)

☐ Reference-based – Go to 70
☐ Assembly-based – Go to 73

70. Reference genome used

Please indicate multilocus sequence type (e.g. ST8) and isolate ID (e.g. one of the isolates from the current cluster, ID of a public reference strain or an in-house isolate).

71. Please indicate the read mapper used (e.g. BWA, Bowtie2)

72. Please indicate the variant caller used (e.g. SAMtools, GATK)

73. Please indicate the assembler used (e.g. SPAdes, Velvet)

74. Please specify the variant caller used (e.g. NUCMER)

75. Please select tool used for the allele analysis

(State one answer only)

☐ BioNumerics – Go to 77
☐ SeqPhere – Go to 77
☐ BIGSdb-*Lm* – Go to 77
☐ Other – Go to 76

76. If another tool is used, please enter here:

77. Please indicate allele calling method

☐ Assembly based and mapping-based – Go to 78
☐ Only assembly-based – Go to 78
☐ Only mapping-based – Go to 79

78. Please indicate the assembler used (e.g. SPAdes, Velvet)

79. Please select scheme used for the allele analysis

(State one answer only)

☐ Applied Math (wgMLST) – Go to 81
☐ Applied Math (cgMLST/Pasteur) – Go to 81
☐ Pasteur (cgMLST) – Go to 81
☐ Ruppitsch (cgMLST) – Go to 81
☐ Other – Go to 80

80. If another scheme (e.g. in-house) is used, please give a short description

81. Please report the number of loci in the used allelic scheme

82. Third analysis on data derived from WGS

83. Select a representative isolate in the cluster detected by the third analysis

Indicate the isolate ID

84. Results for the third cluster analysis (e.g. SNP or Allele based)

Please use 9999 for not analysed

Isolate	ST	Distance (e.g. SNP) to the selected cluster isolate
Isolate 1		
Isolate 2		
Isolate 3		
Isolate 4		
Isolate 5		
Isolate 6		
Isolate 7		
Isolate 8		
Isolate 9		
Isolate 10		
Isolate 11		

85. Additional questions to the WGS part

86. Where was the sequencing performed?

(State one answer only)

☐ In own laboratory
☐ Externally

87. Protocol used to prepare the library for sequencing:

(State one answer only)

☐ Commercial kits – Go to 88
☐ Non-commercial kits – Go to 90

88. Please indicate name of commercial kit:

89. If relevant please list deviation from commercial kit shortly in few bullets: - Go to 91

90. For non-commercial kit, please indicate a short summary of the protocol:

91. The sequencing platform used

(State one answer only)

☐ Ion Torrent PGM – Go to 93
☐ Ion Torrent Proton – Go to 93
☐ Genome Sequencer Junior System (454) – Go to 93
☐ Genome Sequencer FLX System (454) – Go to 93
☐ Genome Sequencer FLX+ System (454) – Go to 93
☐ PacBio RS – Go to 93
☐ PacBio RS II – Go to 93
☐ HiScanSQ – Go to 93
☐ HiSeq 1000 – Go to 93
☐ HiSeq 1500 – Go to 93
☐ HiSeq 2000 – Go to 93
☐ HiSeq 2500 – Go to 93
☐ HiSeq 4000 – Go to 93
☐ Genome Analyzer lix – Go to 93
☐ MiSeq – Go to 93
☐ MiSeq Dx – Go to 93
☐ MiSeq FGx – Go to 93
☐ ABI SOLiD – Go to 93
☐ NextSeq – Go to 93
☐ MinION (ONT) – Go to 93
☐ Other – Go to 92

92. If another platform is used, please list here:

93. Criteria used to evaluate the quality of sequence data

In this section, you can report criteria used to evaluate the quality of sequence data.

Please first reply on the use of 5 selected criteria that were the most frequently reported by the participants in the Listeria EQA-5 scheme, 2017.

Next you will be asked to report 5 additional criteria of your own choice.

For each criteria, please also report the threshold or procedure used to evaluated the current criteria.

94. Did you use confirmation of genus to evaluate the quality of sequence data?

☐ Yes
☐ No – Go to 96

95. Procedure used to evaluate confirmation of genus:

96. Did you use coverage to evaluate the quality of sequence data?

(State one answer only)

☐ Yes
☐ No – Go to 98

97. Procedure or threshold used for coverage:

98. Did you use Q score (Phred) to evaluate quality of sequence data?

(State one answer only)

☐ Yes
☐ No – Go to 100

99. Threshold or procedure used to evaluate Q score (Phred):

100. Did you use genome size to evaluate the quality of sequence data?

(State one answer only)

☐ Yes
☐ No – Go to 102

101. Procedure or threshold used for genome size:

102. Did you evaluate the number of good cgMLST loci?

(State one answer only)

☐ Yes
☐ No – Go to 104

103. Threshold or procedure used to evaluate the number of good cgMLST loci:

104. Other criteria used to evaluate the quality of sequence data.

Please list up to 5 additional criteria (e.g. N50, read length, contamination)

105. Other criteria used to evaluate the quality of sequence data – additional criteria 1:

106. Threshold or procedure used to evaluate the additional criteria 1:

107. Other criteria used to evaluate the quality of sequence data - additional criteria 2:

108. Threshold or procedure used to evaluate the additional criteria 2:

109. Other criteria used to evaluate the quality of sequence data – additional criteria 3:

110. Threshold or procedure used to evaluate the additional criteria 3:

111. Other criteria used to evaluate the quality of sequence data – additional criteria 4:

112. Threshold or procedure used to evaluate the additional criteria 4:

113. Other criteria used to evaluate the quality of sequence data – additional criteria 5:

114. Threshold or procedure used to evaluate the additional criteria 5:

115. Comment(s):

e.g. Remarks to the submission, data analyses or laboratory methods

116.

Thank you for filling out the submission form for the *Listeria* EQA-6.

For questions, please contact list.eqa@ssi.dk or phone +45 3268 8341 +45 3268 8372.

We highly recommend to document this submission form by printing it. You will find the print option after pressing the "Options" button.

Important: After pressing "Submit results", you will no longer be able to edit or print your information.

For final submission, remember to press "Submit results" after printing.

European Centre for Disease Prevention and Control (ECDC)

Gustav III:s Boulevard 40, 16973 Solna, Sweden

Tel. +46 858601000
Fax +46 858601001
www.ecdc.europa.eu

An agency of the European Union
www.europa.eu

Subscribe to our publications
www.ecdc.europa.eu/en/publications

Contact us
publications@ecdc.europa.eu

Follow us on Twitter
@ECDC_EU

Like our Facebook page
www.facebook.com/ECDC.EU

ECDC is committed to ensuring the transparency and independence of its work

In accordance with the *Staff Regulations for Officials and Conditions of Employment of Other Servants of the European Union* and the *ECDC Independence Policy*, ECDC staff members shall not, in the performance of their duties, deal with matters in which they may, directly or indirectly, have a personal interest that could impair their independence. Declarations of interest must be received from any prospective contractor before a contract can be awarded.
www.ecdc.europa.eu/en/aboutus/transparency

Publications Office
of the European Union