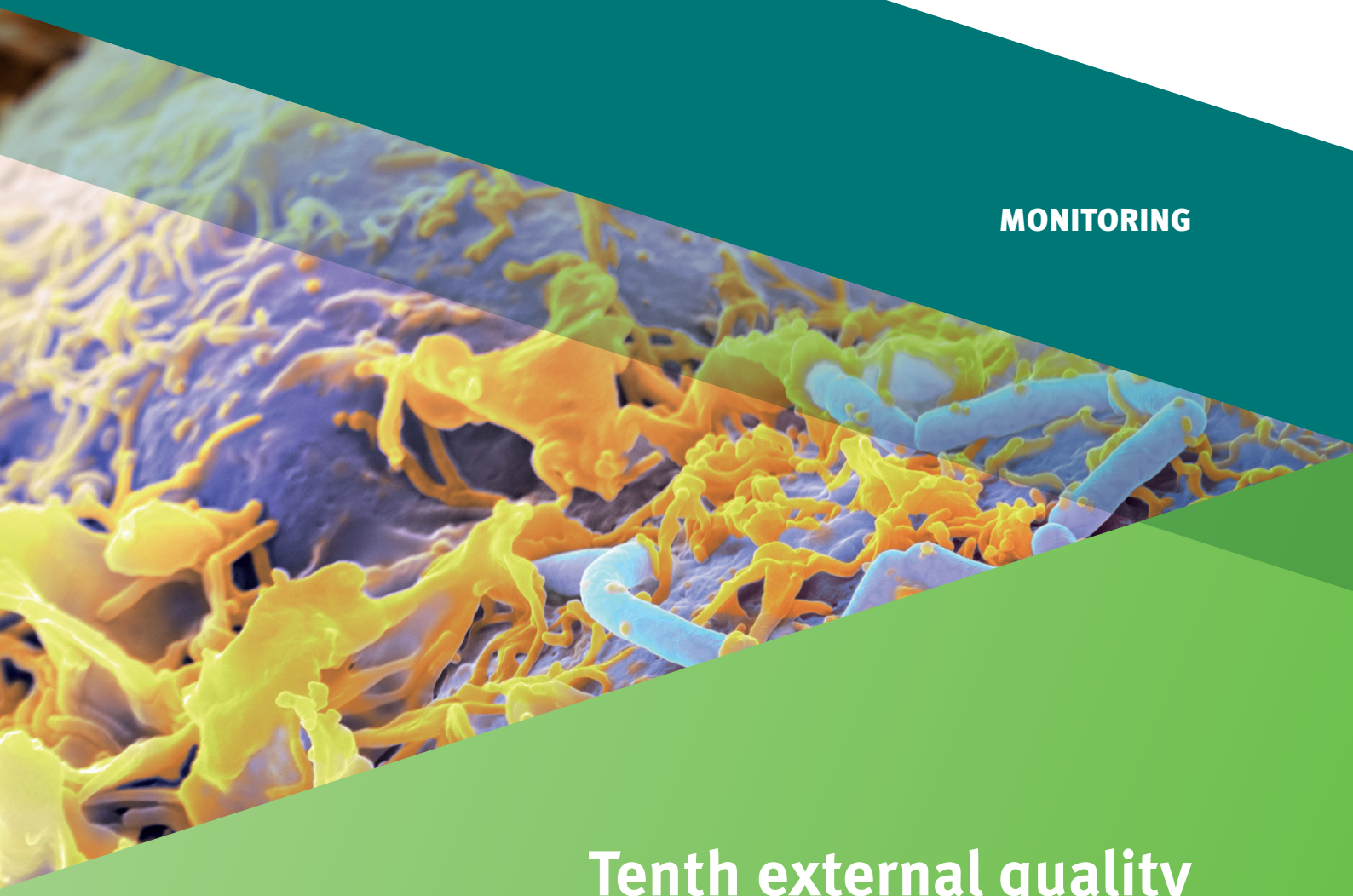
ecdc
EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL



MONITORING

Tenth external quality assessment scheme for *Listeria monocytogenes* typing in EU/EEA and EU enlargement countries, 2023

ECDC MONITORING

Tenth external quality assessment scheme for *Listeria monocytogenes* typing in EU/EEA and EU enlargement countries, 2023

This report was commissioned by the European Centre for Disease Prevention and Control (ECDC), coordinated by Taina Niskanen (ECDC Emerging, Food- and Vector-borne Diseases Programme) and produced by Susanne Schjørring, Anne Sophie M Uldall, Kristoffer Kiil 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. Ninth external quality assessment scheme for *Listeria monocytogenes* typing in EU/EEA and EU enlargement countries, 2023. Stockholm: ECDC; 2025.

Stockholm, February 2025

ISBN 978-92-9498-777-8
doi: 10.2900/6471065
Catalogue number TQ-01-25-008-EN-N

Contents

Figures

Tables

Abbreviations

AD	Allelic differences
BN	BioNumerics
cgMLST	Core 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
NPHRL	National Public Health Reference Laboratory
QC	Quality control
R1/R2	Read1 (forward)/Read2 (reverse) from a paired-end run (FASTQ file)
SNP	Single-nucleotide polymorphism
SNV	Single-nucleotide variant (based on cgMLST)
SSI	Statens Serum Institut
ST	Sequence type
TESSy	The European Surveillance System
wgMLST	Whole genome multilocus sequence type
WGS	Whole-genome sequencing

Executive summary

This report presents the results of the tenth round of the external quality assessment (EQA-10) scheme for *Listeria monocytogenes (L. monocytogenes)* typing in EU/EEA and EU enlargement countries, organised for national public health reference laboratories (NPHRLs) 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 series of framework contracts with ECDC.

Human listeriosis is a relatively rare but serious food-borne disease with a European Union (EU) notification rate of 0.62 cases per 100 000 population in 2022 [3]. With an increase of 16% from 2021 to 2022 (2 365 and 2 738 cases, respectively) the number of cases in 2022 was even higher than before the COVID-19 pandemic (2019, 2 621 cases) [3].

Since 2007, ECDC has been responsible for EU-wide surveillance of listeriosis, including facilitating the detection and investigation of food-borne outbreaks. Surveillance data, including certain basic typing parameters, are reported by European Union/European Economic Area (EU/EEA) countries to The European Surveillance System (TESSy). Since 2012, the EQA scheme has covered molecular typing methods used for EU-wide surveillance.

EQA-10, conducted between May and November 2023, involved serotyping/grouping and molecular typing-based cluster analysis. The objective of this EQA was to assess the quality and comparability of typing data reported by NPHRLs participating in FWD-Net. Test strains for the EQA were selected to cover strains that are currently pertinent for public health in Europe and to represent a broad range of clinically relevant types of invasive listeriosis. Seven test strains were selected for serotyping/grouping and molecular typing-based cluster analyses. An additional ten sequences were included for the molecular typing-based cluster analysis. Of the 23 laboratories that signed up, 21 completed the exercise. This represented an increase of one laboratory compared with EQA-9; however, the composition was different, as two laboratories from EQA-9 did not participate in EQA-10, two new laboratories joined EQA-10, and one laboratory from EQA-8 rejoined for EQA-10. Most participating laboratories (15/21; 71%) completed the full EQA scheme.

In total, 18 laboratories (86%) participated in the serotyping part of the exercise; all of them conducted PCR-based/WGS molecular serogrouping, and two laboratories (11%, 2/18) also conducted conventional antigen-based serotyping. On average, molecular serogrouping was performed well, with 95% correct results. For the conventional method, 100% of the participants correctly serotyped all seven test strains. One laboratory, participating for the first time, mistyped four of the seven strains in the molecular serogrouping, swapping two of the isolates. Since the first EQA in 2012, the trend has been towards replacing conventional serotyping with molecular serogrouping, showing strong performance.

Of the 21 laboratories participating in the EQA-10, 18 (86%) performed molecular typing-based cluster analysis using a method of their choice. The intent of the cluster analysis component of the EQA was to evaluate the NPHRLs' capacity to identify a genetically closely related cluster. In other words, the goal was to accurately categorise the cluster test strains – regardless of the method used – rather than strictly be able to follow a specific procedure.

The cluster of seven closely related strains (three test strains and four strain sequences) was predefined by the EQA provider using data derived from whole genome sequencing (WGS). Therefore, as expected, the correct cluster delineation was not possible to obtain using less discriminatory methods (e.g. pulsed-field gel electrophoresis (PFGE)). Nevertheless, the three cluster strains were correctly identified by the participant that used PFGE, though they could not include the sequences provided in their analysis. Seventeen laboratories performed cluster analysis using WGS-derived data; only one used single-nucleotide polymorphism (SNP) as the main analysis. The submitted allelic differences (AD) clearly showed coherence despite the different approaches and schemes that the participants used. The most widely used core genome multilocus sequence type (cgMLST) scheme was Ruppitsch (9/16 laboratories), while the Pasteur scheme was less common (5/16). Most laboratories (11/16) reported 0–8 AD for the strains in the predefined cluster; however, not all included the strains with 7 or 8 AD in the reported cluster and others excluded the modified strain with low read R2 quality (strain13). When analysing the predefined cluster of the seven closely related strains, 47% (8/17) of the participants reported the same list of strains as the EQA provider. Two laboratories reported the cluster without strain13 and two laboratories employed a stringent cut-off, excluding borderline cluster strains. Furthermore, three of the laboratories that failed to identify the predefined cluster were suspected of submitting errors. Only two laboratories exhibited concerning errors: one laboratory interchanged strains and another included all sequence type 5 (ST5) in the cluster.

In general, most of the participants were able to identify the different characteristics and modifications of the EQA-provided sequences. All laboratories detected issues with the sequence that had a mix of sequence types and 94% of participants identified the quality control (QC) issues of low coverage. Different observations were made for both the strain with slightly reduced coverage and the strain with low read quality of R2.

A feedback survey was sent to the participants to assess their experience of EQA-10; 48% (10/21) of the participants responded. Of note, the QC evaluation of participant-sequenced data and the inclusion of low-quality data were considered useful by all respondents. Additionally, two respondents suggested uploading a TSV file instead of entering the data in a survey format to reduce typing errors and one suggested using an even bigger dataset to assess the overall cluster congruence between pipelines.

1 Introduction

1.1 Background

ECDC is an EU agency with a mission to identify, assess and communicate current and emerging threats to human health from communicable diseases. ECDC's founding regulation outlines its mandate as fostering the development of sufficient capacity within EU/EEA-dedicated surveillance networks for diagnosis, detection, identification and characterisation of infectious agents that may threaten public health. Under this mandate, ECDC supports the implementation of quality assurance schemes [1].

External quality assessments (EQAs) are an essential part of quality management. An external organiser is used to assess the performance of laboratories on test samples supplied specifically for quality assessment purposes.

ECDC has outsourced the organisation of EQA schemes for EU/EEA countries to the disease networks. The aim of EQAs is to identify areas of improvement in the laboratory diagnostic capacities relevant for epidemiological surveillance of communicable diseases, as in Decision No 1082/2013/EU [2], and to ensure the reliability and comparability of results generated by laboratories across all EU/EEA countries.

The main objectives of the EQA schemes are:

- to assess the general standard of performance ('state of the art');
- to assess the effects of analytical procedures (method principle, instruments, reagents, calibration);
- to support method development;
- to evaluate individual laboratory performance;
- to identify problem areas;
- to provide continuing education;
- to 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 typing of *Salmonella enterica* ssp. *enterica*, Shiga toxin/verocytotoxin-producing *Escherichia coli* (STEC/VTEC) and *L. monocytogenes*. In 2021, SSI won the new round of tenders (2022–2025) for *Listeria* and STEC. The *Listeria* EQA covers serotyping and molecular typing-based cluster analysis. This report presents the results of the *Listeria* EQA-10.

1.2 Surveillance of listeriosis

Human listeriosis is a relatively rare but serious food-borne disease, with high rates of morbidity, hospitalisation and mortality in vulnerable populations. From 2017 to 2019, the number of human listeriosis cases increased slightly in the EU (from 2 474 to 2 621 cases). However, after a decrease in the number of confirmed human listeriosis cases in 2020 (1 887), the number of cases rose in 2021 (2 365) and increased to higher than before the COVID-19 pandemic in 2022 (2 738). This increase can also be seen in the 2022 notification rate (0.62 cases per 100 000 population), which is higher than in 2017–2021, when the notification rate was stable (0.47–0.49 cases per 100 000 population) [3].

One of ECDC's key objectives is to improve and harmonise surveillance systems in the EU to increase scientific knowledge of aetiology, risk factors and burden of food- and waterborne diseases and zoonoses. Countries report surveillance data, including basic typing parameters for pathogens isolated from human infections, to TESSy. There is also public health value in using more discriminatory typing techniques for pathogen characterisation in the surveillance of foodborne infections. Since 2012, ECDC has enhanced EU surveillance by incorporating molecular typing data through isolate-based reporting for selected food-borne pathogens. Since March 2019, ECDC has been coordinating WGS-enhanced, real-time surveillance of invasive listeriosis within the EU/EEA. The overall aims of integrating molecular typing into EU-level surveillance are:

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

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

The current EQA scheme for *L. monocytogenes* typing is the tenth EQA organised for NPHRLs in the FWD-Net. The molecular typing-enhanced surveillance system, implemented as part of TESSy, relies on the capacity of FWD-Net laboratories to produce sequences of good quality and comparable typing results for cross-border cluster detections. Member States are asked to submit *L. monocytogenes* WGS data in real-time to TESSy to be used for EU-wide WGS-enhanced listeriosis surveillance. The submitting of WGS data is requested to be performed whenever new data are available in the laboratories or in relation to on-going multi-country outbreak investigations. Since the end of 2022, all human isolate clusters detected by ECDC are used to query the EFSA One Health WGS system on a weekly basis.

1.3 Objectives of the EQA-10 on *Listeria*

The EQA-10 on *Listeria* offered quality support for NPHRLs that perform molecular typing-enhanced surveillance at the national level.

1.3.1 Serotyping/serogrouping

EQA-10 assessed serotype determination by either conventional antigen-based typing of somatic 'O' antigens and flagellar 'H' antigens, or PCR-based/WGS molecular serogrouping.

1.3.2 Molecular typing-based cluster analysis

EQA-10 sought to assess a participating laboratory's ability to detect a cluster of closely related strains. Laboratories could perform their analysis using PFGE and/or WGS-derived data. The cluster analysis was to be conducted on seven test strains and an additional set of 10 test strains with provided genomic sequences. Some of the provided sequences were modified to include QC issues.

2 Study design

2.1 Organisation

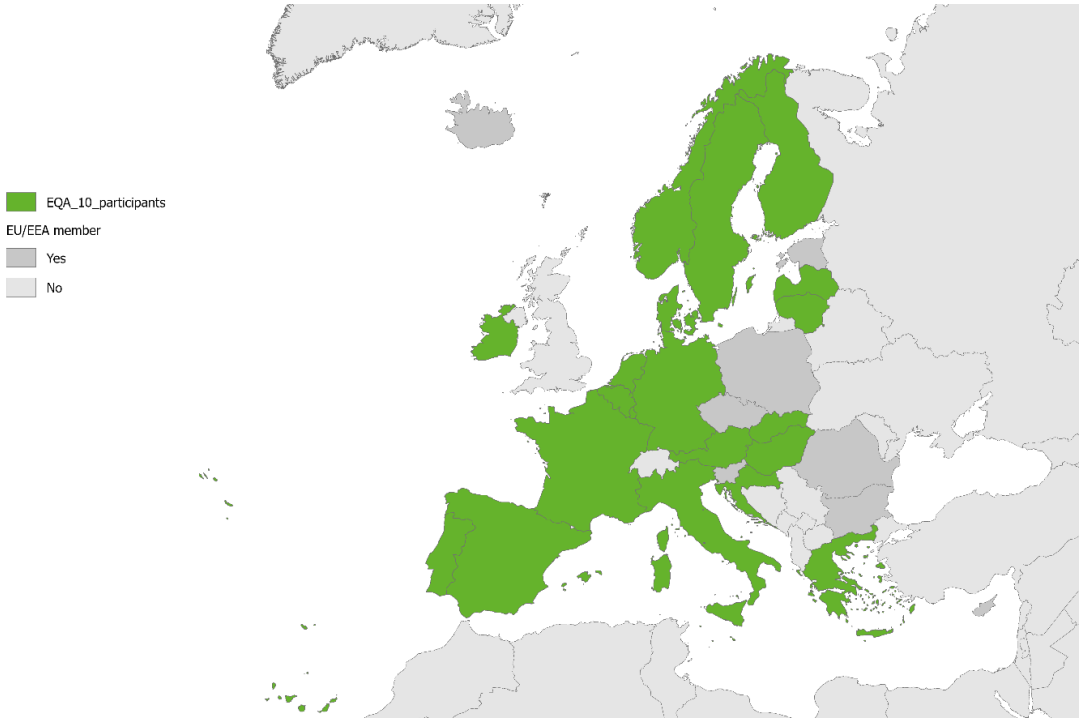
The EQA-10 on *Listeria* was funded by ECDC and arranged by SSI following ISO/IEC 17043:2010 [4]. EQA-10 was carried out between May and November 2023.

The EQA provider emailed invitations to ECDC's contact points in FWD-Net (30 countries) on 2 May 2023, with a deadline to respond by 8 May 2023. Invitations were also sent to EU candidate countries.

Twenty-three NPHRLs in the EU/EEA and EU candidate countries accepted the invitation to participate and 21 submitted results (Figure 1, Annex 1, Table 1A); details of participation in EQA-9 and EQA-10 are listed in Annex 1, Table 2A to provide an overview of the number of participants.

The EQA test strains were sent to participants on 31 May 2023. Participants were asked to submit their results by 30 August 2023, using the online form (Annex 5). If WGS was performed, submission of the raw reads (FASTQ files) was requested. The EQA submission protocol was distributed by email and made available online.

Figure 1. Countries participating in EQA-10, EU/EEA and candidate countries, 2023



Administrative boundaries: © EuroGeographics, UN-FAO, Turkstat. The boundaries and names shown on this map do not imply official endorsement or acceptance by the European Union. EQA-10 participating countries are shown in green.

2.2 Selection of test strains/sequences

Ten candidate strains were analysed using the methods set out in the EQA (serotyping and WGS) before and after re-culturing. All candidate strains remained stable using these methods and seven final test strains were selected: five test strains and a set of technical duplicates (twice from the same culture). In addition, ten sequences (representing the genomes of nine additional test strains, as one sequence was the same as the technical duplicate) were provided for the participants to include in the cluster analysis, and four of the sequences were modified by the EQA provider to have various QC issues.

Seventeen *L. monocytogenes* test strains/sequences were selected to fulfil the following criteria:

- cover a broad range of the commonly reported, clinically relevant strains of invasive listeriosis in Europe;
- include genetically closely related strains;
- remain stable during the preliminary test period at the organising laboratory;
- include three 'repeat strains' from EQA-1 to EQA-10; and
- include a set of technical duplicates in the serotyping/grouping.

The seven test strains for serotyping were selected to cover different serotypes/-groups (1/2a/IIa, 1/2b/IIb, 1/2c/IIc, and 4b/IVb). Again, this year all of the provided test strains and sequences had to be assessed in the cluster analysis. The provided test strains and the sequences had different 7-gene multilocus sequence types (ST2, ST5, ST7 and ST9) and varied in relatedness.

To follow the development of each laboratory's performance (the reproducibility), three strains of different serotypes/groups were included in EQA-1 to 10: strain1 (4b-IVb), strain5 (1/2c-IIc) and strain6 (1/2a-IIa). Based on the WGS-derived data, the selected cluster of closely related strains consisted of at least five *L. monocytogenes* ST5 strains (including the technical duplicate set strain2/strain3 and provided sequence strain16). The maximum cluster also included strain7, which was borderline, with 8 AD/13 SNPs and strain13, which was a cluster from strain2 manipulated to have poor read quality of R2. A single-end assembly of R1 could be used in the analysis 0 AD/0 SNP. Characteristics of all the *L. monocytogenes* test strains are listed in Table 1 and Annex 2, Tables 3A–5A, 7A–14A The EQA provider found at most 7 AD or 10 SNPs between any two strains in the minimum cluster or 8 AD/13 SNPs in the maximum cluster (Annex 2, Table 8A og 9A). The EQA provider's cluster analysis of WGS-derived data was based on an allele-based (cgMLST [7]) and SNP analysis (NASP [8]). The participant using PFGE as a cluster method could only evaluate the seven test strains from the package, and only three belonged to the cluster of closely related strains based on WGS. The provided sequences represented four ST5 strains, one ST2 strain and five modified sequences: one ST217 mixed with approximately 6% *L. welshimeri,* one ST5 genome with massively reduced coverage, one ST5 cluster genome with poor read quality of R2, one sequence with slightly reduced coverage (core percentage of approximately 94%), and one sequence was a mix of ST7 and ST224. (Table 1).

Table 1. Characteristics of test strains and sequences

Strain no.	Serotyping analysis	Cluster analysis		
	Strains	ST	QC status	Part of the predefined cluster?
Strain1[a]	4b/IVb	2	NA	No
Strain2[b]	1/2b/IIb	5	NA	Yes
Strain3[b]	1/2b/IIb	5	NA	Yes
Strain4	1/2b/IIb	5	NA	No
Strain5[a]	1/2c/IIc	9	NA	No
Strain6[a]	1/2a/IIa	7	NA	No
Strain7	1/2b/IIb	5	NA	Yes
Strain8-sequence	IVb	2	A	No
Strain9-sequence	IIb	5	A	Yes
Strain10-sequence[c]	NA	NA	C	NA
Strain11-sequence[c]	NA	NA	C	NA
Strain12-sequence	IIb	5	A	Yes
Strain13-sequence[c]	IIb	5	B	Yes
Strain14-sequence[c]	IIb	5	B	No
Strain15-sequence[c]	NA	NA	C	NA
Strain16-sequence	IIb	5	A	Yes
Strain17-sequence	IIb	5	A	No

A: acceptable quality; B: quality only acceptable for outbreak situations (less good quality); C: unacceptable quality (strain not analysed); NA: not applicable (quality too poor to include in the cluster analysis); ST: sequence type; QC: quality control.
[a] `Repeat strains included in EQA-1 to 10 (strain1, strain5 and strain6). Strain6 was different from the strain used in previous years, although it was the same serotype/group.
[b] Technical triplicates were strain2, strain3 and strain16-sequence (Annex 2, Table 3A-4A).
[c] Modified sequences included: strain10, a non-cluster sequence (ST217) contaminated with approximately 6% L. welshimeri*; strain11, a non-cluster sequence (ST5) with low coverage; strain13, a cluster sequence (conducted from strain2/3) with low read quality of R2 (The provider's QC pipeline only accept paired-end reads. However, only R2´s read quality is too low for proper analysis and the R1 read quality is acceptable. Therefore, a single-end assembly was conducted for use in outbreak investigation.); strain14, a non-cluster with core percentage of approximately 94%; strain15, a non-cluster sequence of ST7 and ST224 combined.*

2.3 Distribution of strains and sequences

The seven test strains were blinded and shipped on 31 May 2023. The protocol for the EQA exercise and a letter stating the unique strain IDs were included in the packages and emailed to participants on the same day. The packages were shipped from SSI, labelled 'UN3373 Biological Substance'. Eighteen participants received their dispatched strains within two days and five within five days after shipment. No participants reported damage to the shipment or errors in the unique strain IDs.

On 14 June 2023, instructions regarding the procedure for submitting results were emailed to the participants. This included the link to the website where participants could upload sequences and download the additional test strains 8–17 (FASTQ genomic sequences) and the empty submission form.

2.4 Testing

In the serotyping part of the EQA, seven *L. monocytogenes* strains 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 serogrouping (multiplex PCR, according to the protocol suggested by Doumith et al. [5, 6]) or in silico PCR by using WGS data. The results of serotyping/grouping were submitted in the online form.

In the cluster analysis part, participants could choose to perform PFGE (*Apa*I and *Asc*I profiles) and/or use WGS-derived data. The participants were instructed to report the IDs of the strains included in the cluster of closely related strains, by method. Laboratories performing WGS could use their own analysis pipeline for cluster analysis, e.g. single-nucleotide polymorphism analysis (SNP-based) or whole/core genome multilocus sequence typing (wgMLST)/cgMLST (allele-based). The participants were asked to report the number of loci in the allelic scheme used for cluster analysis and/or the name of the SNP pipeline used.

The participants were asked to report the strains identified as a cluster of closely related strains based on the analysis used. Laboratories could submit results from up to three analyses (one main and two additional analyses), but the detected cluster had to be based on results from the main analysis. The results were reported as SNP distance or AD between each test strain and a strain (strain16) selected by the EQA provider. The 7-gene multilocus sequence types and the serotype of strains in the cluster analysis could also be submitted.

Each participant also needed to undertake QC by assessing the quality of the provided sequences. The three possible QC results were:

A: acceptable quality;
B: quality only acceptable for outbreak situations (less good quality);
C: unacceptable quality (strain not analysed).

The participants were instructed to describe the observations and considerations that lead to their QC status decisions. The EQA provider modified five sequences (strain10, strain11, strain13, strain14 and strain15) (Table 6, Annex 2, Tables 10A–14A).

The laboratories uploaded the raw reads (FASTQ files) for further analysis by the EQA provider.

2.5 Data analysis

The submitted serotyping and cluster analysis results, as well as the raw reads, were imported to a dedicated *Listeria* EQA-10 BioNumerics (BN) database. Five laboratories did not submit the raw reads by the deadline. One additional laboratory uploaded incomplete data due to upload errors. The EQA provider contacted these laboratories to obtain these data and the sequences were analysed in BioNumerics.

Serotyping results were evaluated according to the percentage of correct results, generating a score from 0−100%.

Molecular typing-based cluster analysis was evaluated as correct or incorrect identification of the expected cluster of closely related strains. This was determined according to predefined categorisation that the EQA provider based on WGS-derived data. This categorisation was obtained by allele-based analysis (cgMLST [7] and SNP analysis (NASP, [8])). The correct number of closely related *L. monocytogenes* strains/sequences by WGS was seven ST5 strains: strain2, strain3, strain9, strain12, strain16 (strain7 and strain13), with 0–8 AD/0–13 SNPs. The provider included strain7 in the cluster despite being borderline, with 8 AD/13 SNPs. Strain13 was also included in the cluster despite having been manipulated to have poor read quality of R2; therefore, a single-end assembly of R1 needed to be used if the laboratory's system demanded pair-end reads.

The EQA provider accepted submitted clusters without strain13, if the participant excluded the sequence due to assigning QC level C. Strain2 and strain3 were from the same culture and were sent to participants as two strains and as the provided sequence strain16. (Annex 2; Table 8A–9A. The participants using PFGE only evaluated strain1 to strain7 and only three belonged to the cluster of closely related strains based on WGS.

The participant's descriptions and the QC status of the EQA provider's modified sequences are listed in Annex 2, Table 10A–14A.

Individual evaluation reports and certificates of attendance were distributed by email to participants in November 2023. If WGS data was 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 contained comments on the QC status of the submitted sequences.

3 Results

3.1 Participation

Laboratories could participate in the full EQA scheme or in one part only (serotyping/serogrouping or molecular typing-based cluster analysis). Of the 23 laboratories that signed up, 21 completed and submitted their results. The majority of the participants (15/21; 71%) completed the full EQA scheme. In total, 18 (86%) participants participated in serotyping and 18 (86%) in cluster analysis. Molecular serogrouping results were also provided by 86% of the participants, two of which also performed conventional serotyping.

Most participants (17/18; 94%) reported cluster analysis using only WGS-derived data, while one (6%) reported using only PFGE data (Table 2).

Table 2. Number and percentage of laboratories submitting results for each method

	Serotyping			Cluster analysis		
	Molecular only	Conventional and molecular	Total	PFGE only	WGS only	Total
Number of participants	16	2	18	1	17	18
Percentage of participants	89	11	86[a]	6	94	86[a]

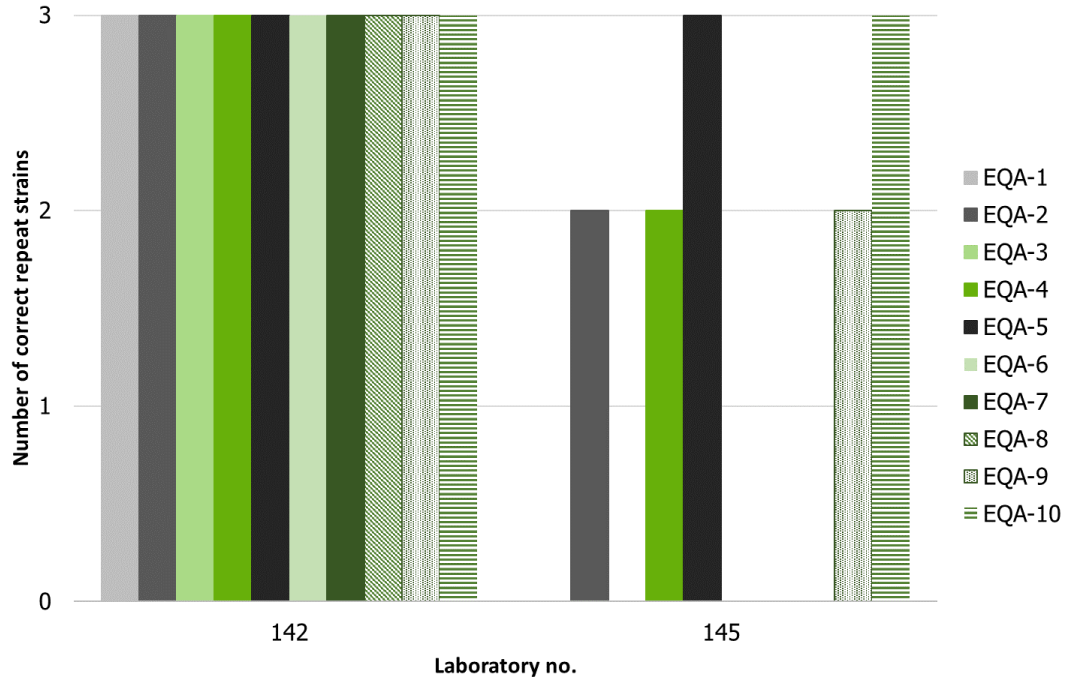
PFGE: pulsed-field gel electrophoresis; WGS: whole genome sequencing.
Of the 21 participating laboratories, 15 (71%) completed both parts (serotyping and cluster analysis) of EQA-10.
[a] Percentage of the total number of laboratories (21) participating in the EQA.

3.2 Serotyping

3.2.1 Conventional serotyping

Two participants performed conventional serotyping of *L. monocytogenes* with a strong performance, achieving 100% correctness (Annex 2, Table 3A). Figure 2 and Table 3 show the reproducibility of the individual participants' performances in conventional serotyping of the three repeat strains from EQA-1 to EQA-10 for those two laboratories. The reproducibility of conventional serotyping results for the repeat strains shows stability and strong performance for the one laboratory participating every year (laboratory 142). Laboratory 145 demonstrated strong performance in EQA-10 despite encountering some issues in EQA-2, EQA-4 and EQA-9, and not participating every year.

Figure 2. Correctly assigned conventional serotypes for three repeat strains from EQA-1 to EQA-10 for two laboratories participating in EQA-10



Arbitrary numbers are used to represent the participating laboratories. Bars represent the number of correctly assigned serotypes for the three repeat strains (strain1, strain5 and strain6).

Table 3. Correctly assigned conventional serotypes for three repeat strains from EQA-1 to EQA-10 for two laboratories participating in EQA-10

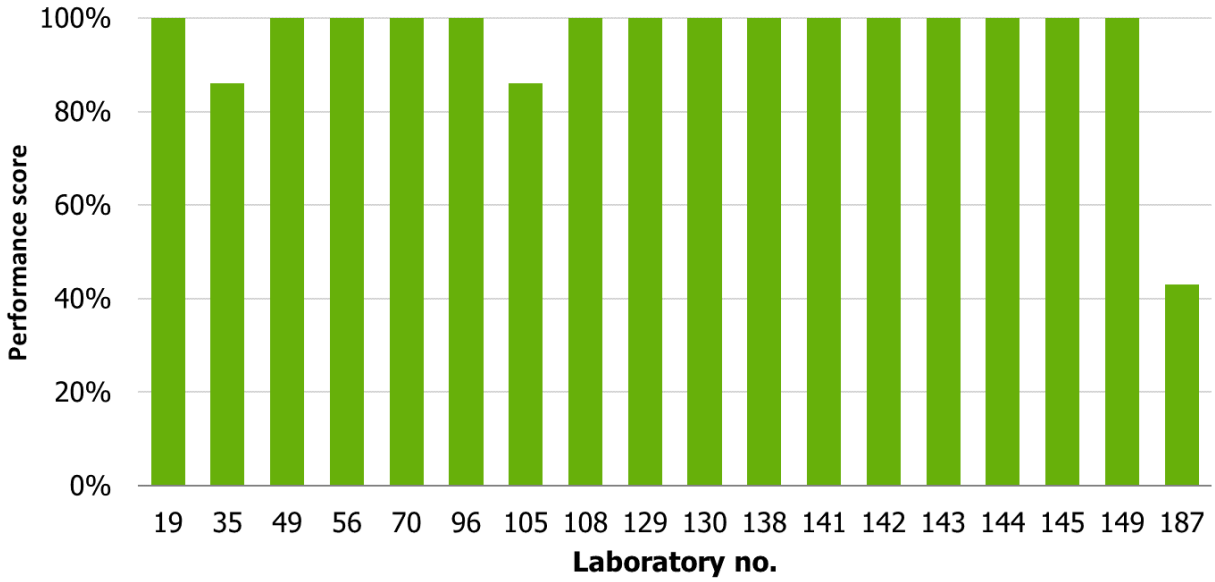
EQA round	Number of correctly assigned serotypes[a]	
	Laboratory no.	
	142	145
EQA-1	3	NA
EQA-2	3	2
EQA-3	3	NA
EQA-4	3	2
EQA-5	3	3
EQA-6	3	NA
EQA-7	3	NA
EQA-8	3	NA
EQA-9	3	2
EQA-10	3	3

NA: not applicable, as the laboratory did not participate in that EQA round.
[a] The number of correctly assigned serotypes for the three repeat strains (strain1, strain5 and strain6).
Arbitrary numbers were used to represent the participating laboratories.

3.2.2 Molecular serogrouping

Eighteen participants performed molecular serogrouping of *L. monocytogenes* in EQA-10 (Figure 3). Molecular serogrouping was carried out in accordance with guidelines [5] and nomenclature [6] by Doumith et al. Fifteen (83%) participants were able to correctly serogroup all seven EQA test strains, with an average performance of 95%. Twelve of the 18 participants reported using WGS-based analysis (in silico PCR) for molecular serogrouping. Most (4/6) errors were reported by laboratory 187.

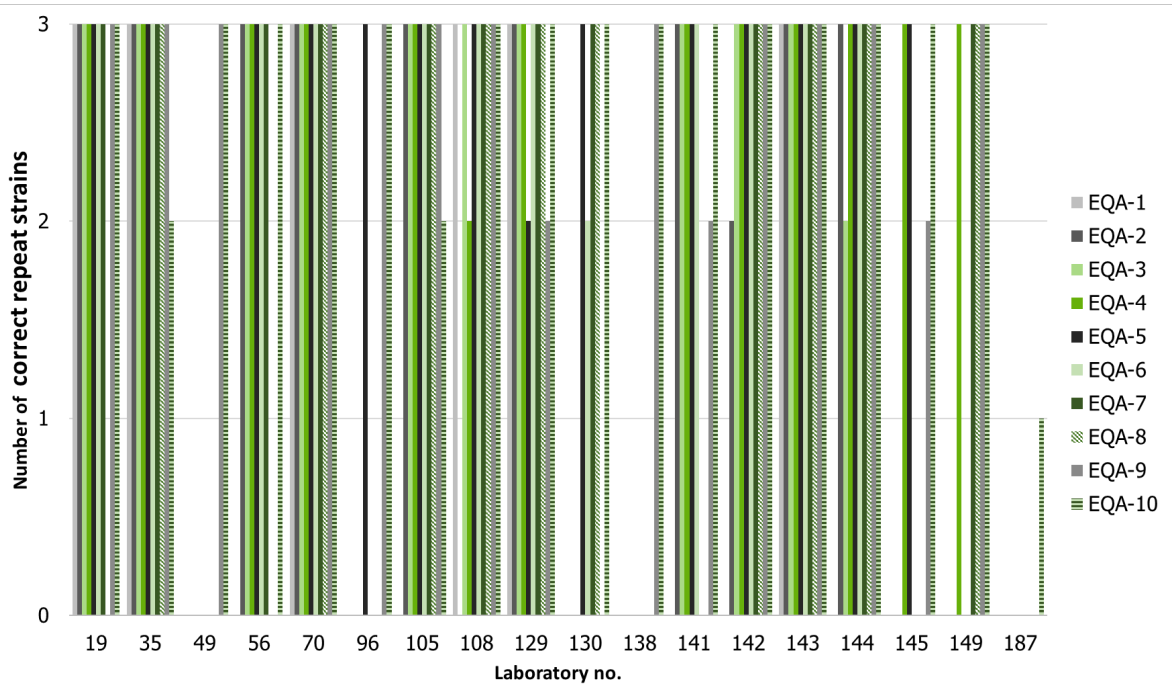
Figure 3. Participant scores for molecular serogrouping of seven *L. monocytogenes* test strains



Arbitrary numbers were used to represent the participating laboratories. Bars represent the percentage of correctly assigned serogroups for the seven test strains (strain1 to strain7).

Figure 4 and Table 4 show the individual reproducibility of participants' performances in molecular serogrouping when assessing the three repeat strains during the 10 EQAs. Out of the 18 laboratories that participated in EQA-10, 7/18 (39%) correctly serogrouped all three repeat strains in all the EQA rounds they participated in.

Figure 4. Number of correctly assigned molecular serogroups for three repeat strains from EQA-1 to EQA-10 for laboratories participating in EQA-10



Arbitrary numbers were used to represent the participating laboratories. Bars represent the number of correctly assigned serotypes for the three repeat strains (strain1, strain5 and strain6).

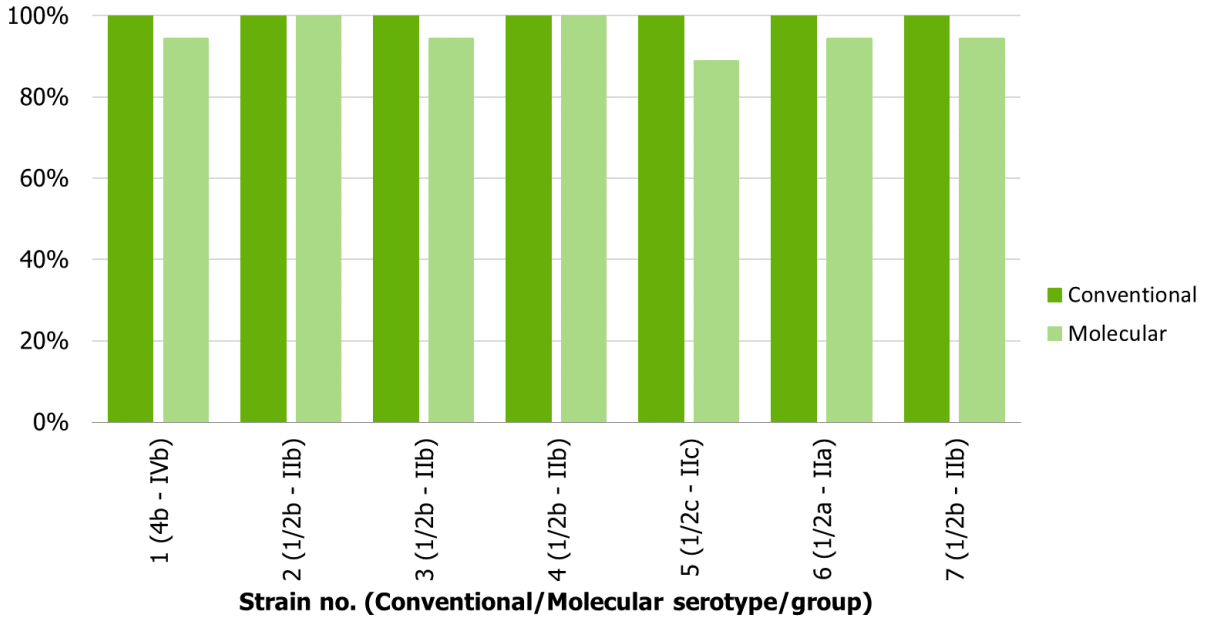
Table 4. Number of correctly assigned molecular serogroups for three repeat strains from EQA-1 to EQA-10 for laboratories participating in EQA-10

EQA round	Number of correctly assigned serogroups[a] Laboratory no.																
	19	35	49	50	70	96	105	108	129	138	141	142	143	144	145	149	187
EQA-1	3	3	NA	NA	NA	NA	NA	NA	3	3	NA	NA	NA	NA	3	NA	NA
EQA-2	3	3	NA	NA	3	NA	3	3	NA	3	NA	NA	3	2	3	3	NA
EQA-3	3	3	NA	NA	3	NA	3	3	3	3	NA	NA	3	3	3	2	NA
EQA-4	3	3	NA	NA	3	0	3	3	2	3	NA	NA	3	3	3	3	NA
EQA-5	3	3	NA	NA	3	3	3	3	3	2	3	NA	3	3	3	3	NA
EQA-6	3	3	NA	NA	3	NA	3	3	3	3	2	NA	3	3	3	3	NA
EQA-7	3	3	NA	NA	3	NA	3	3	3	3	3	NA	NA	3	3	3	NA
EQA-8	NA	3	NA	NA	0	NA	3	3	3	3	3	NA	NA	3	3	3	NA
EQA-9	3	3	3	NA	NA	2	3	3	3	2	NA	3	2	3	3	3	NA
EQA-10	3	2	3	3	3	3	2	3	3	3	3	3	3	3	3	3	1

NA: not applicable, as the laboratory did not participate in that EQA round.
[a] The number of correctly assigned serogroups for the three repeat strains (strain1, strain5 and strain6).
Arbitrary numbers were used to represent the participating laboratories.

Figure 5 shows the reported error distributed per strains. The six errors were reported in different strains. Only two of the errors was in strain5, one reporting IIa or IIb instead of IIc.

Figure 5. Average scores for seven test strains in EQA-10



Bars represent the percentage of correctly assigned serotypes/groups. *Two laboratories performed conventional serotyping and 18 performed molecular serogrouping.*

3.3 Molecular typing-based cluster analysis

Participants were tested on their ability to correctly identify the cluster of closely related strains among the seven test strains and 10 provided sequences.

3.3.1 PFGE-derived data

One participant (1/18; 6%) performed cluster analysis using PFGE-derived data. The cluster categorisation was based on WGS data including the provided sequences; as expected, it was not possible to obtain the correct cluster delineation using this less discriminatory method. The participant correctly identified the three strains as belonging to a cluster, and did not report any additional strains as belonging to the cluster.

Table 5 provides an overview of the strains that the participant included or excluded in their cluster identification.

Table 5. Results of Laboratory 130's cluster analysis based on PFGE-derived data in EQA-10

Strain no.	ST	Part of the cluster
Strain1	2	No
Strain2[a,b]	5	Yes
Strain3[a,b]	5	Yes
Strain4	5	No
Strain5	9	No
Strain6	7	No
Strain7	5	Yes
Included the three strains from the WGS cluster[c]		**Yes**
Included any additional strains in the cluster		**No**

ST: sequence type; WGS: whole genome sequencing.
[a] *Closely related strains predefined by WGS (in grey).*
[b] *Technical duplicate strains (in bold).*
[c] *Predefined categorisation by WGS-derived data.*

3.3.2 WGS-derived data

3.3.2.1 Details reported on equipment and method

Seventeen participants (17/18; 94%) performed cluster analysis using WGS-derived data. One laboratory reported using an external laboratory for sequencing, while 94% used their own laboratory. Different sequencing platforms were listed by the participants: one MiniSeq, seven MiSeq, eight NextSeq, and one Ion GeneStudio S5 System. All reported using commercial kits for library preparation. Of the 17 participants, 15 (88%) used Illumina's Nextera kit. Two participants had modified the manufacturer's protocol by changing volumes (Annex 2, Table 6A).

3.3.2.2. Assessment of the QC status of the sequences provided

Almost all the provided sequences that were not modified were reported as acceptable quality (QC status A) by the participants, except strain17, by laboratory 177 (as they reported N50 as below 30 000bp).

For strain10 – a non-cluster sequence ST217 contaminated with approximately 6% *L. welshimeri* – 100% of the participants correctly observed the contamination of the sequence and 94% (16/17) reported QC status C.

For strain11 – a non-cluster sequence ST5 with massively reduced coverage and removal of genes – all but one participant (16/17) correctly identified the poor quality of the sequence and 81% (13/16) correctly excluded the sequence from the cluster analysis (QC status C). One of the four participants that included the sequence in the analysis also incorrectly listed strain11 as part of the cluster.

For strain13 – a cluster sequence with low read quality of R2 – 41% (7/17) of the participants reported the sequence as acceptable (QC status A), 18% (3/17) reported the sequence as only acceptable for outbreak investigation (QC status B), and 41% (7/17) of the participants reported the sequence to be of unacceptable for analysis (QC status C). When observing quality issues on only one of the reads, a single-end assembly was an option to still utilise the data.

For strain14 – a non-cluster sequence with core percentage reduced to 94% – 35% (6/17) of the participants reported the sequence as acceptable (QC status A) and 47% (8/17) reported the sequence as only acceptable for an outbreak investigation (QC status B). The remaining three participants – laboratories 56, 88 and 141 – reported the sequence to be unacceptable for analysis (QC status C). The participants reported that the percentage of good cgMLST targets was below their threshold, and some even specified below 95%.

For strain15 – a mix of two non-cluster sequences ST7 and ST224 combined into one FASTQ file – 100% of the participants correctly observed the contamination of the sequence and 94% (16/17) reported QC status C. One participant (laboratory 177) reported that the quality was not optimal (QC status B). They reported a warning, as the number of contigs was over 200 and N50 was below 30 000bp.

Table 6. Results of participants' quality control assessment of the modified sequences provided for EQA-10

Strain no.	Strain characteristics	QC status results			
		EQA provider	A	B	C
Strain10	A non-cluster sequence contaminated with approximately 6% *Listeria welshimeri*	C	0	1	16
Strain11	A non-cluster sequence with low coverage	C	1	3	13
Strain13	A cluster sequence with low read quality of R2	B	7	3	7
Strain14	A non-cluster sequence with core percentage of approximately 94%	B	6	8	3
Strain15	A non-cluster sequence of ST7 and ST224 combined	C	0	1	16

QC: quality control.
Raw data are available in Annex 2, Tabel 10A–13A1. QC status is ascertained as: acceptable quality (A), quality only acceptable for outbreak situations (less good quality) (B) or unacceptable quality – strain not analysed (C).

3.3.2.3. Cluster analysis results

Each participant was required to use their own produced sequences and the sequences provided (after assessment of QC status) in the cluster analysis. They were instructed to report which strains/sequences were part of the cluster of closely related strains, thereby mimicking an urgent outbreak situation where it is impossible to re-run the sequencing and less than optimal quality sequences may have to be assessed.

In general, the performance was strong in the cluster analysis with WGS-derived data; however, this year the EQA provider complicated the analysis by including borderline strains with 7 or 8 AD (data from the provider analysis), and a cluster sequence with modified R2 read quality (originally strain2). Therefore, the results were divided into different groups (Tables 7–9). Eight participants (47%) correctly identified the cluster of seven closely related predefined by the EQA-provider. Additionally, two participants (12%) identified the cluster without including the modified strain13, as they excluded the sequence from the analysis. Additionally, three laboratories used a strict cut-off and

consequently excluding borderline cluster sequences, two included a sequence/strain above the range of expected cut-off and two laboratories had concerning errors, one laboratory swapped strains and one included all ST5 in the cluster.

Laboratories were instructed to report the data analysis used for cluster identification, using a provided sequence (strain16) as reference for reporting SNP distance or AD. Laboratories could report results from one main analysis and one or two additional analyses, but the detected cluster had to be based on results from the main analysis. Only two participants (Laboratories 19 and 177) reported additional analyses (Table 7).

Table 7. Results of cluster analysis based on WGS-derived data

Strain no.	ST	Laboratory no.																
		19	35	49	56	70	88	105	108	129	135	138	141	142	144	149	177	187
Strain1	2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+
Strain2[a,b]	5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Strain3[a,b]	5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	–
Strain4	5	–	+	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–
Strain5	9	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Strain6	7	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Strain7[b]	5	+	+	-	-	+	+	+	+	+	+	+	+	-	+	+	+	–
Strain8-sequence	2	–	–	–	–	–	–	–	–	ND	–	–	–	–	–	–	–	–
Strain9-sequence[b]	5	+	+	–	–	+	+	+	+	+	(+)	+	+	-	+	+	+	+
Strain10-sequence[c]	ND	ND	ND	ND	ND	–	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND
Strain11-sequence[c]	ND	ND	ND	ND	ND	–	ND	ND	ND	–	ND	ND	–	ND	ND	ND	ND	+	ND
Strain12-sequence[b]	5	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+
Strain13-sequence[b,c]	5	+	+	ND	ND	+	+	ND	+	+	ND	ND	+	ND	+	+	+	ND
Strain14-sequence[c]	5	–	–	–	ND	–	ND	–	–	–	–	+	ND	–	–	–	+	–
Strain15-sequence[c]	–	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	ND	–	ND
Strain16-sequence[a,b]	5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Strain17-sequence	5	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	+	–
Main analysis		A	A	A	A	A	A	A	S	A	A	A	A	A	A	A	A	A
Additional analysis		S	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	S	NA
Cluster identified (without strains)[a]		Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	(Yes)	No	Yes	No	Yes	Yes	No	No
Cluster identified		Yes	No	No	No	Yes	Yes	No	Yes	Yes	No	No	Yes	No	Yes	Yes	No	No

+: the participant reported the strain as part of the cluster; (): partial accepted, as the participant communicated additional information in a comment field; –: the participant reported the strain as not being part of the cluster; A: allele-based analysis (cgMLST) (Annex 2, Table 7A–9A); NA: not applicable; ND: not analysed; S: single-nucleotide polymorphism (SNP) (Annex 8A); ST: 7 multilocus sequence type.

[a] *Technical triplicates (strains or sequence) (in bold)*

[b] *Closely related strains (in grey)*

[c] *Modified sequences*[a] *Cluster (without the strain7 (borderline difference) and modified strain13, QC status B)*

Of the 17 participants, one used SNP as their main analysis and two used SNP as an additional analysis (Table 8). The three participants that did an SNP analysis used a reference-based approach, two with strain16 (provided) as the reference. One used CLC Genomics Workbench for both read mapper and variant caller, and the other used Burrows-Wheeler Aligner (BWA) as the read mapper and Genome Analysis Toolkit (GATK) as the variant caller. The third participant used MINtyper (CGE tool).

Tables 8 and 9 show the overview of the submitted data. Each laboratory reported SNP distances/AD by strain (Annex 2, Table 8A).

Table 8. Results of SNP-based cluster analysis

Laboratory no.	SNP-based analysis							
	SNP pipeline	Approach	Reference	Read mapper	Variant caller	Identified predefined cluster	Difference reported within cluster (SNP)	Included strain7 and strain13 (difference to strain7/13 in SNP)
Provider	NASP [8]	Rb	Strain16	BWA	GATK	Yes	0–13	Yes (13/0)
19[a]	NASP	Rb	EQA strain 0016	BWA	GATK	Yes	0–13	Yes (13/0)
108	In-house	Rb	EQA strain 0016	CLC Assembly Cell	CLC Assembly Cell	Yes	0–14	Yes (14/0)
177[a]	MINTyper 1.0	Rb	ST5	MINtyper	MINtyper	No	0–30	Yes (0/0)

BWA: Burrows-Wheeler Aligner; CLC: CLC Genomics Workbench; GATK: Genome Analysis Toolkit; Rb: reference based.
[a]Additional analysis
For more details see Annex 2, Table 8A.

Almost all (16/17) participants used allele-based analysis as the main analysis for cluster detection (Table 9). Two (13%) used SNP as an additional method. The majority (75%) used an only assembly-based (OAB) allele calling method (Table 9). All 16 reported using cgMLST: nine used cgMLST Ruppitsch [9] (eight with 1701 loci, one with 2866 loci), six used cgMLST Pasteur (five with 1748 loci and one with 95% 1660 loci) and one used an in-house cgMLST scheme with only 1503 loci.

Table 9. Results of allele-based cluster analysis

Laboratory no.	Allele-based analysis							
	Approach	Allelic calling method	Assembler	Scheme	No. of loci	Identified predefined cluster/without strain 13?	AD reported within cluster	Included strain7 and strain13 (AD to strain7/13)
Provider	BioNumerics	A&M	SPAdes	Applied Maths (cgMLST/Pasteur)	1748	Yes	0–8	Yes (8/0)
19	BioNumerics	A&M	SPAdes	Applied Maths (cgMLST/Pasteur)	1748	Yes	0–8	Yes (8/0)
35	SeqSphere	A&M	Velvet, SPAdes	Ruppitsch (cgMLST)	1701	No/No	0–20	Yes (6/0)
49	BioNumerics	OAB	SPAdes	Applied Maths (cgMLST/Pasteur)	1748	No/No	0–3	No (8/NA)
56	BIGSdb-Lm	A&M	INNUca v4.2.2	Pasteur (cgMLST)	1748	No/No	0–4	No (8/NA)
70	SeqSphere	A&M	SPAdes v3.15.4	Ruppitsch (cgMLST)	1701	Yes	0–8	Yes (6/0)
88	INNUca_v4.2.2, chewBBACA_v.3.1[a]	OAB	SPAdes INNUca_v4.2.2	Pasteur (cgMLST)	1748	Yes	0–8	Yes (8/0)
105	SeqSphere	OAB	SPAdes	Ruppitsch (cgMLST)	1701	No/Yes	0–8	Yes (6/NA)
129	SeqSphere	OAB	Velvet	The Ridom SeqSphere+ software's Target Definer	1503	Yes	0–5	Yes (3/0)
135	SeqSphere	OAB	SPAdes	Ruppitsch (cgMLST)	1701	No/Yes	0–6	Yes (8/NA)
138	chewbbaca,3.1.2[a]	OAB	shovill,1.1.0	Ruppitsch (cgMLST)	1701	No/No	0–65	Yes (10/NA)
141	SeqSphere	OAB	SPAdes 3.15.5	Ruppitsch (cgMLST)	1701	Yes	0–8	Yes (6/0)
142	BIGSdb-Lm	OAB	SPAdes 3.13.0	Pasteur (cgMLST)	1748	No/No	0–4	No (13/NA)
144	SeqSphere	OAB	SKESA	Ruppitsch (cgMLST)	1701	Yes	0–8	Yes (6/0)
149	SeqSphere	OAB	SKESA 2.4.0	Ruppitsch (cgMLST)	1701	Yes	0–8	Yes (6/0)
177	BIGSdb-Lm	OAB	Shovill	Pasteur (cgMLST)	1660	No/No	0–30	Yes (0/0)
187	SeqSphere	OAB	SKESA	Ruppitsch (cgMLST)	2866	No/No	0–10	No (2 362/NA)

A&M: assembly and mapping based; AD: allelic difference; NA: not analysed; OAB: only assembly based.
[a] Modified from submitted information.
For more details see Annex 2, Table 8A.

Laboratory 108, the sole participant that conducted SNP analysis as the primary method, successfully identified the correct cluster of closely related strains, including strain13. They reported 0–14 SNPs within the cluster, yielding results comparable to the EQA-provider. The participant was unaware of the low read quality of R2, as their SNP analysis remained unaffected.

Laboratories 19 and 177 performed SNP analysis as an additional analysis. Laboratory 19 correctly identified the cluster of closely related strains using cgMLST as the main analysis and reported SNP distances ranging from 0–13 within the cluster, including strain13 for the additional analysis. Both Laboratory 19 and 108 exhibited a clear separation (Figure 6). Laboratory 177 utilised an SNP-based MINTyper for the additional analysis and reported 0–2 SNPs.

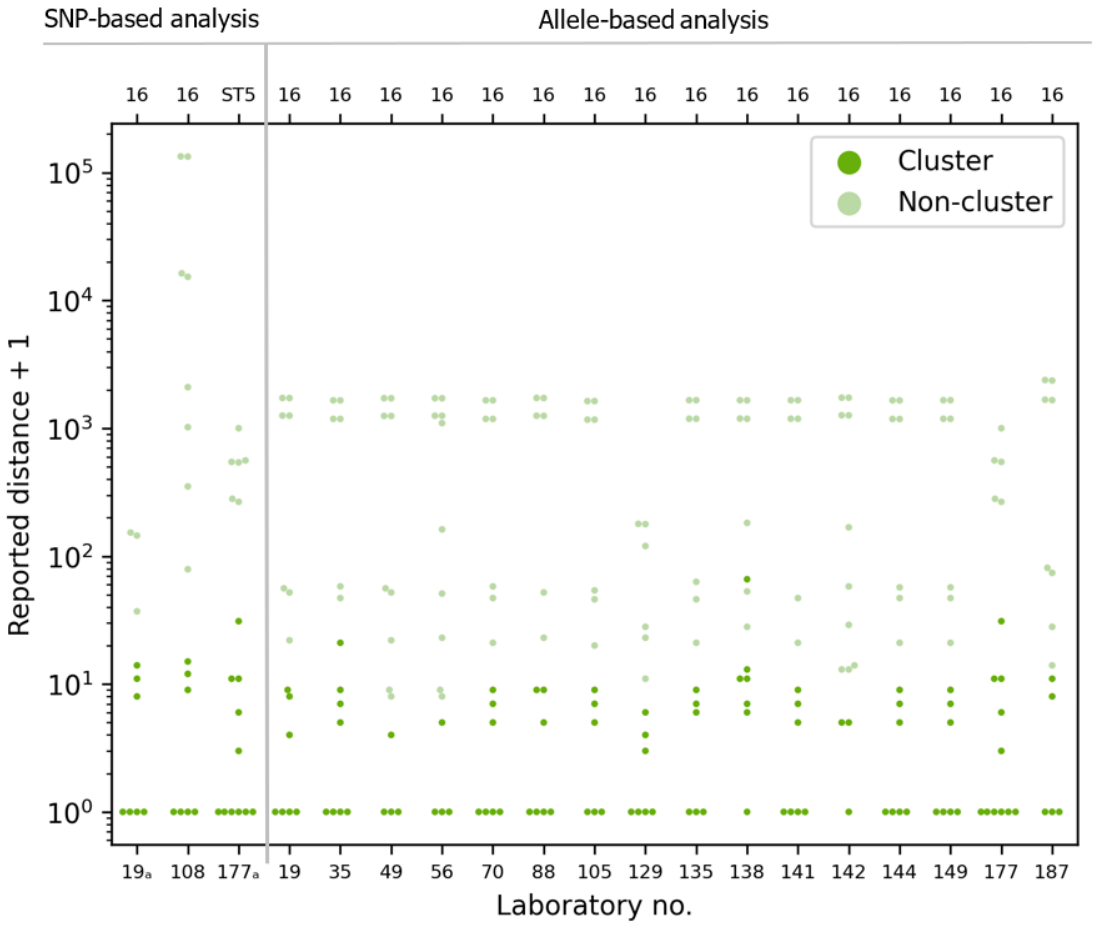
The laboratories using an allele-based method can be divided into four groups (Table 7 and 9; Figure 6):

- Group 1 (7/16) correctly identified the predefined cluster of seven strains, with a maximum of 0–8 AD within the cluster, including strain13.
- Group 2 (2/16) excluded strain13 based on their QC analysis of the provided sequences and, consequently, did not report strain13 as part of the cluster (0–8 AD).
- Group 3 (3/16) neither included strain13 in their cluster analysis nor reported strains/sequences with borderline AD (7–13 AD) between the reference strain (strain16) and strains 7 and 9. This implied the exclusion of strains from the reported cluster.
- Group 4 (4/16) consisted of Laboratory 35, which mistakenly included strain4 with 20 AD in the reported cluster. After receiving the evaluation report, it was confirmed to be a submitting error. Laboratory 138 incorrectly included strain14 with 65 AD in the reported cluster but correctly excluded strain17 with 52 AD, suggesting a submitting error as well. Laboratory 187 reported four mistakes in the serogroup section, and the results of the reported cluster suggest that strain1 was mixed up with strain3 and strain5 with strain7. However, the sequences were correctly named when submitted to the sFTP.

Excluding laboratory 187, most laboratories obtained similar results (0 SNP/AD) for the technical triplicates (strain2, strain3 and strain16). However, Laboratory 142 reported 4 AD and Laboratory 138 reported 5/6 AD, once again surpassing the majority.

Three of the test strains/provided sequences (strain4, strain14 and strain17) were also ST5, but were not predefined by the EQA provider as part of the cluster (cgMLST (AD 20–55)/SNP (36–718)). Based on cgMLST, the thirteen laboratories (Group 1, Group 2 and Group 3) reported AD ranging from 10–62 for this group of strains, excluding laboratory 142, which – in general – reported a higher number of AD. Figure 6 presents all the laboratories' reported SNP distances and AD, illustrating the similarities within allele-based analysis.

Figure 6. Reported single-nucleotide polymorphism distances or allelic differences for each test srain to selected cluster representative strain16

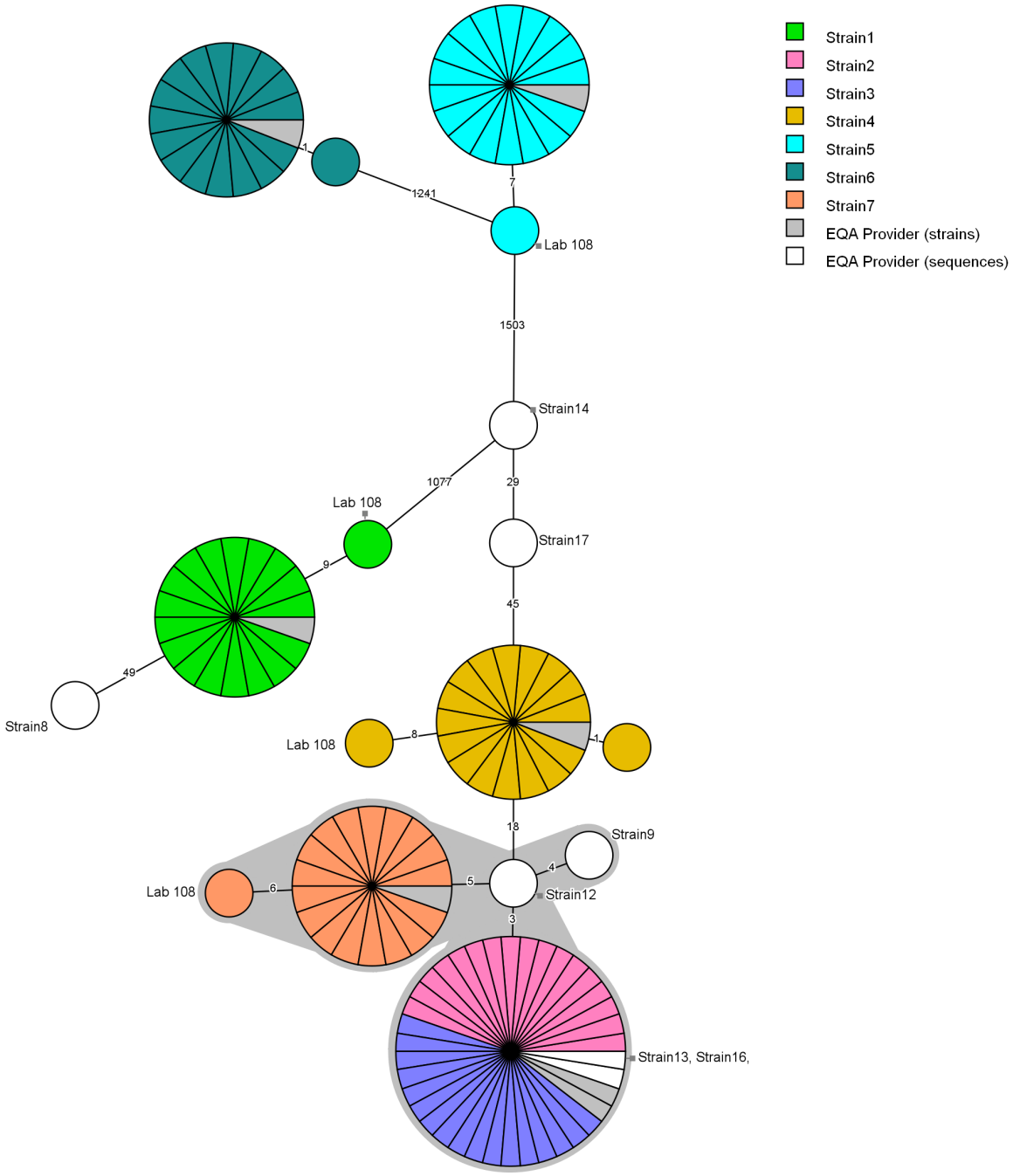


SNP: single nucleotide polymorphism
[a] Additional analysis
Dark green: reported cluster of closely related strains; Light green: not reported as part of cluster. 3.3.2.4. Analysis of raw reads uploaded by participants.

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) [7] and evaluated by the EQA provider's in-house quality control pipeline [10].

The overall cgMLST analysis, as shown in the minimum spanning tree (MST), is based on raw reads (FASTQ files) submitted by 17 laboratories and the sequences provided by the EQA provider. This analysis excludes three sequences due to poor quality (strain10, strain11 and strain15). Figure 7 reveals a clear clustering of the results for each test strain; only data notes from Laboratory 108 are separated with distance from the other results.

Figure 7. Minimum spanning tree of core genome multilocus sequence typing of participants' FASTQ files



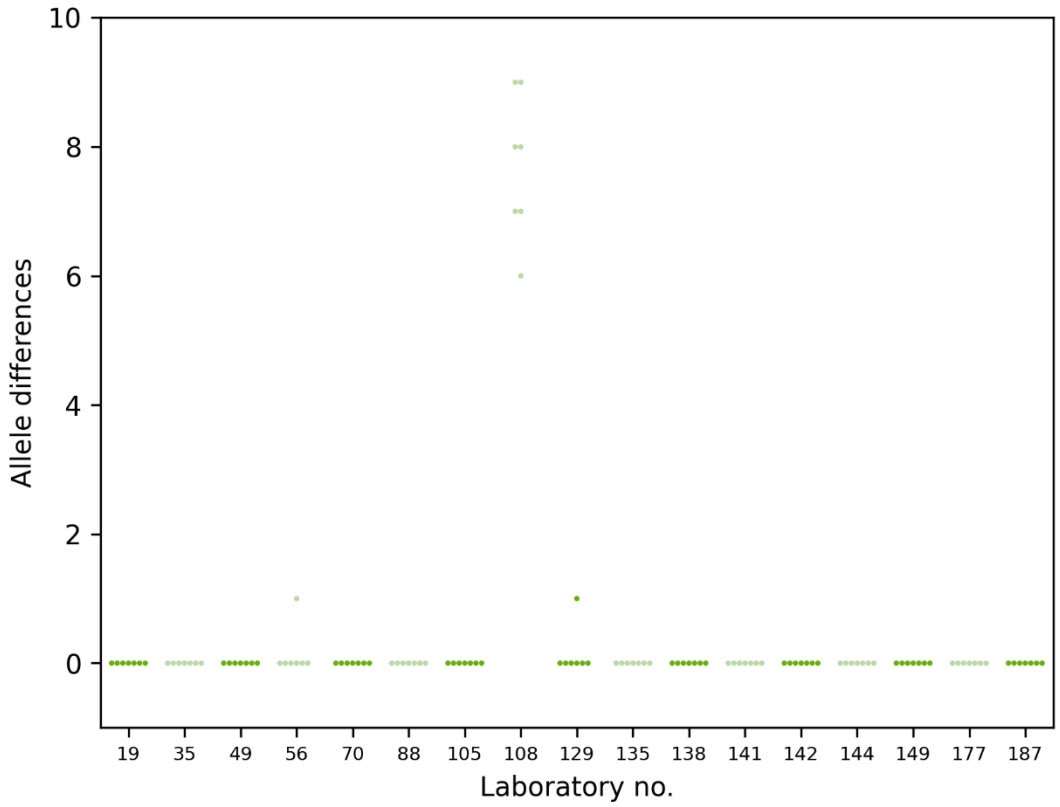
Minimum spanning tree (MST) in log scale of core genome multilocus sequence typing (cgMLST) [7] based on submitted raw reads (FASTQ files). Each of the strain1–7 test strains have a different colour. Sequences (strain1–7) by the EQA provider are in grey, the sequences (strain9–17) are in white. The modified sequences of poor quality (strain10, strain11 and strain15) and three sequences from Laboratory 108 were not included in the analysis.

The AD in Figures 7 and 8 do not precisely match those illustrated in the individual reports. This discrepancy is caused by loci being dropped if they do not pass QC for all strains in the analysis. Therefore, the joint analysis contains fewer loci.

For each laboratory, cgMLST was conducted on the raw reads submitted (FASTQ files), applying Applied Maths allele calling using the Pasteur scheme [7]. Hierarchical single linkage clustering was then carried out for each laboratory, using the submitted data alongside the EQA provider's reference strains. As seen in Figures 7 and 8, all laboratories exhibited minor differences from the reference strains. Laboratory 108 had the most sequences different from those of the EQA provider.

Figure 8 shows the AD between each submitted sequence and the corresponding reference.

Figure 8. Allelic differences of participant strains from corresponding EQA provider results (strain1–7), based on submitted raw reads (FASTQ files)



For 110 of 119 (92%) results, no differences were identified. As seen in Figure 8, in two (2%) instances, a difference of one allele from the corresponding EQA-provided strain was calculated. Results from Laboratory 108 exhibited AD for all seven strains, ranging from six to nine; these results were attained using Ion Torrent data analysed in BioNumerics. Notably, Laboratory 56 also used Ion Torrent data but had no differences.

The laboratories reported the QC parameters used to evaluate their data separately. Coverage and confirmation of the genus were the most common QC parameters, used by 100% and 94% of the laboratories, respectively (Table 10). Participants reported different thresholds of coverage, ranging from 20–60x. The laboratories reported using different programmes for the contamination check of the genus, with six reporting having used Kraken [11]. The number of good cgMLST loci was used as a QC parameter by 82% of the participants, with thresholds ranging from 89–98%. Genomic size was used by 88%, ranging from 2.5–3.5 Mb, and q score was only reported by 47% of participants. Additional QC parameters were provided by some of the participants (listed in Annex 2, Table 15A) and several participants listed N50 and GC% content as a parameter used, with a threshold of 30 000 bp.

Table 10. Summary of selected quality control parameters reported by participants

Laboratory no.	Confirmation of genus	Coverage	Q score (Phred)	Genome size	No. of good cgMLST loci
19	Kraken analysis and <5% contamination with other genus	Min. 50x	NR	2.8–3.1 Mbp	Min. 95% core percentage and max. 30 loci with multiple consensus
35	Multiplex-PCR, Kraken, SeqSphere	Min. 20x	NR	2 8–3.2 Mbp	90%
49	PubMLST Species ID Ribosomal MLST	>30x	>30	2.6–3.2Mb	>2 800
56	NR	30x	No	NR	1 660 genes mapped
70	Mash Screen in SeqSphere and ID species in PubMLST	≥50x	No	Length of contigs assembled < Ref genome + 10%	≥98%
88	Kraken (as implemented in INNUca_v4.2.2) and ConFindr (≤3 Num contam SNVs)	INNUca_v4.2.2 employs several coverage thresholds throughout the analysis (e.g. 15x for the first estimated coverage; 30x for the assembly coverage)	INNUca_v4.2.2, and Trimmomatic (INNUca default settings were applied for these steps)	INNUca_v4.2.2, using 3.0 Mbp as the expected genome size, and INNUca default criteria	Allele calling perfomed with chewBBACA (using default settings) and filtering (≥95% loci called) with ReporTree during clustering analysis
105	Assembled genomes aligned against a *Listeria monocytogenes* genome (Threshold: >90%nucleotide identity)	>45x	Trimming performed with Trimmomatic, Phred <10 or an average Phred <15 in a sliding window of four nucleotides. Sequences with a length <70 bases were removed too.	≤3.3Mb	≥95%
108	Species identified by BLAST against an in-house database with reference sequences	>20x	NR	2.8–3.3 Mbp	NR
129	Presence of prfA gene (LIPI)	>29x	NR	NR	>89%
135	Species identification tool (Juno) built into in-house assembly pipeline	>30x	>30	2.7–3.23 Mb	>90% of alleles
138	Kraken2 with database built from all RefSeq genomes; rMLST from PubMLST	30x	≥30; fastp: v0.22.0	2.918–3.156 Mb	NR
141	Contamination check (Mash Screen) in SeqSphere, CGE tools, JSpecies	Min. 30x	NR	2.9–3.1 Mb	Min. 98% good targets
142	Kraken and PRS/prfA gene	> 30x	>30	2.5–3.5 Mb	>95%
144	SeqSphere	>50x	>30	2.7–3.3 Mb	>95%
149	Kraken	>50x	No	Approx. 2.9 Mb (range: 2.8–3.0 Mb)	>90%
177	PubMLST	20–60x	NR	2.7–3.2 Mb	NR
187	rMLST, KmerFinder, MLST	Quast	33	2.88 Mb	95%
Percentage of laboratories using QC parameter	**94%**	**100%**	**47%**	**88%**	**82%**

Approx.: approximately; BLAST: Basic Local Alignment Search Tool; CGE: Center for Genomic Epidemiology; cgMLST: core genome multilocus sequence type; LIPI: Listeria Pathogenicity Island; Max: maximum; Min: minimum; MLST: multilocus sequence typing; PCR: polymerase chain reaction; prfA: positive regulatory factor A; PRS: polygenic risk score; QC: quality control; rMLST: ribosomal multilocus sequence typing; SNVs: single nucleotide variants.

For each laboratory, the submitted raw reads (FASTQ files) were evaluated by the EQA provider's in-house quality control pipeline [10]. Table 11 shows the QC parameters and range of QC values per laboratory. For the full QC evaluation of all strains, see Annex 3, Table 16A–32A.

According to the QC parameters, sequencing quality was uniformly good and overall coverage was high. Only five laboratories had issues; four out of five reported average coverage (below 50 for some strains) and one (Laboratory 108) reported 'Length at >25X coverage' below 2.8. Notably, Bifrost was developed for Illumina data and not Ion Torrent (Annex 3, Table 16A–32A.).

Table 11. Summary of results of raw reads submitted by participants and evaluated by the EQA provider's quality control pipeline

Laboratory no.	Detected species	% Species 1	% Species 2	Unclassified reads (%)	Length at >25 x min. coverage (Mbp)	Length (1-25) x min. coverage (kbp)	No. of contigs at 25 x min. coverage	Contigs at (1,25)X coverage	Average coverage	No. of reads (x 1000	Average read length	Average insert size	N50 (kbp)	QC status (Bifrost)
Ranges[a]	Lm			5%	2.8-3.1	<250	>0	<1000	>50					
19	Lm	92.0–93.4	0.0–0.1	6.5–7.7	2.9–3.0	0.0–0.0	21.0–45.0	0.0–0.0	135.0–173.0	2967.0–3909.0	135.0–140.0	214.0–238.0	173.0–435.0	OK
35	Lm	96.4–97.3	0.0–0.1	2.6–3.3	2.9–3.0	0.0–8.7	32.0–112.0	0.0–9.0	182.0–340.0	3269.0–6958.0	147.0–276.0	311.0–372.0	66.0–189.0	OK
49	Lm	98.1–98.8	0.0–0.2	1.1–1.7	2.9–3.0	0.0–0.0	11.0–18.0	0.0–0.0	88.0–157.0	951.0–1689.0	269.0–285.0	322.0–424.0	344.0–1500.0	OK
56	Lm	97.4–98.2	1.0–1.4	0.7–1.0	2.9–3.0	0.0–0.0	13.0–19.0	0.0–0.0	162.0–229.0	1919.0–2609.0	301.0–301.0	293.0–327.0	344.0–1500.0	OK
70	Lm	97.1–97.7	0.0–0.1	2.3–2.7	2.9–3.0	0.0–0.0	16.0–26.0	0.0–0.0	46.0–100.0	945.0–1956.0	147.0–149.0	323.0–332.0	344.0–541.0	Warning
88	Lm	94.8–95.6	0.0–0.1	4.3–5.0	2.9–3.0	0.0–0.0	17.0–34.0	0.0–0.0	141.0–157.0	2819.0–3310.0	144.0–147.0	255.0–290.0	191.0–1500.0	OK
105	Lm	98.0–98.6	0.0–0.1	1.4–1.9	2.9–3.0	0.0–48.3	14.0–24.0	0.0–2.0	45.0–162.0	921.0–3293.0	147.0–149.0	270.0–407.0	261.0–509.0	Warning
108	Lm	95.5–96.4	0.0–0.1	3.6–4.5	2.6–2.8	0.0–2.4	2436.0–2584.0	0.0–11.0	65.0–98.0	728.0–1046.0	275.0–293.0	0.0–0.0	2.0–2.0	Warning
129	Lm	91.2–96.8	0.0–0.1	3.1–8.4	2.9–3.0	0.0–14.5	24.0–112.0	0.0–12.0	59.0–167.0	1260.0–3615.0	134.0–145.0	263.0–390.0	48.0–396.0	OK
135	Lm	96.4–96.9	0.3–0.8	2.4–3.0	2.9–3.0	0.0–0.0	14.0–25.0	0.0–0.0	196.0–534.0	4030.0–10566.0	151.0–151.0	342.0–367.0	344.0–1500.0	OK
138	Lm	97.5–98.0	0.0–0.1	1.9–2.3	2.9–3.0	0.0–0.0	14.0–22.0	0.0–0.0	61.0–233.0	1253.0–4801.0	148.0–148.0	342.0–400.0	344.0–1500.0	OK
141	Lm	96.6–97.2	0.0–0.1	2.7–3.2	2.9–3.0	0.0–0.0	10.0–21.0	0.0–0.0	74.0–112.0	803.0–1239.0	273.0–283.0	381.0–422.0	344.0–1500.0	OK
142	Lm	97.8–98.3	0.1–0.2	1.4–2.0	2.9–3.0	0.0–0.8	12.0–22.0	0.0–1.0	50.0–61.0	623.0–760.0	242.0–246.0	454.0–473.0	284.0–1500.0	OK
144	Lm	97.1–97.8	0.0–0.1	2.1–2.8	2.9–3.0	0.0–10.3	14.0–30.0	0.0–6.0	136.0–152.0	2792.0–3158.0	149.0–149.0	253.0–297.0	244.0–541.0	OK
149	Lm	92.8–93.9	0.3–0.9	5.3–6.2	2.9–3.0	0.0–0.0	17.0–43.0	0.0–0.0	117.0–168.0	2416.0–3602.0	151.0–151.0	265.0–308.0	244.0–508.0	OK
177	Lm	88.3–91.5	0.0–0.2	8.4–11.4	1.5–3.0	2.7–1414.5	68.0–271.0	2.0–100.0	27.0–52.0	657.0–1341.0	113.0–121.0	150.0–189.0	18.0–75.0	Warning
187	Lm	97.0–97.6	0.0–0.1	2.2–2.9	2.6–3.0	0.0–351.6	16.0–27.0	0.0–3.0	26.0–90.0	545.0–1846.0	146.0–147.0	328.0–398.0	310.0–1500.0	Warning

Lm: L. monocytogenes.
[a] *Quality control range*
'Warning' indicates that Bifrost produced a warning result to indicate issues with the submitted sequences (see Annex 3).

3.4 Feedback survey – evaluation of the EQA scheme

After the results of the individual laboratory reports were sent to the participants, the EQA provider circulated a feedback survey to assess the EQA-10 scheme for *L. monocytogenes*. The questionnaire contained questions on aspects of accreditation, information on the individual report, the actions taken if errors were detected, the usefulness of the QC evaluation of the participants' sequenced data, the usefulness of including low-quality data and any suggestions for improvements. The survey response rate was 48% and the results are summarised in Table 12.

Table 12. Results of the feedback survey to evaluate EQA-10

Questions	Positive response (Yes)	Comments/actions
1) Used for accreditation/licensing purposes?	5/10 (50%)	One reported that their method is not yet subjected to accreditation and/or licensing.
		One reported not yet, but starting from 2024 they want accredited NGS.
		One reported that their method is not accredited yet, but hopefully in the near future.
2) SatisfieBoifrost with the format/comments?	9/10 (90%)	One reported that they were satisfied with the format of the report, but not with the comments provided.
		One participant reported that finding a quick summary or conclusion proved to be challenging. Upon closer examination of the report, all the necessary information was present; however, the participant suggested that the presentation could be more clear and concise.
3) Did any of your analytical test results differ?	2/10 (20%)	One reported, regarding the serogroup misidentifications (IIa vs IIc for Strain5), that they would like to draw attention to this issue in SeqSphere.
		One reported that no corrective actions were needed, but there were some suggestions that justify follow-up for educational purposes.
		One reported that strain11 and strain13 were incorrectly reported as good quality (QC status A), despite the low coverage of strain11 and the less optimal quality of R2 for strain13. They indicated that not all QC parameters were relevant for them, as they use Ion Torrent data.
		One participant reported that the only difference they had in the results pertained to the decision on whether or not one strain should be classified as QC status B: Quality only acceptable for outbreak situations (lower quality). The artificial discrepancy from the expected results could have been avoided if the criteria for classification were indicated in the instructions or if comments were taken into consideration.
4) Usefulness of the manipulated sequences?	8/9 (89%)	One reported that they have their own QC pipelines that accept single-end reads. Therefore, they found the inclusion of a data set with a corrupt R2 file not useful.
		One answered yes, but not always and not all QC values. They are using another sequencing platform and do not look at all the mentioned QC parameters in routine analysis.
5) Usefulness of the QC status of your submitted sequences?	9/9 (100%)	The same participant responded with the same comment as made previously. One answered yes, but not always and not all QC values. They are using another sequencing platform and do not look at all the mentioned QC parameters in routine analysis.
		One reported that they discuss the discrepancies.
6) Improvements/remarks	One reported the following: • Include criteria for strain classification in the instructions; • Reply to our comments in the report; • Allow a 'save and submit later' option in the submission portal; • If possible, allow submission of results in batches (e.g. table in TSV format) instead of requiring participants to write each response one by one (more prone to errors); • Consider including an exercise to assess the overall cluster congruence between pipelines (using, e.g., larger sequence datasets), instead of or as a complement to the fine analysis of a single outbreak.	
	No, they work very well.	
	Some clarity of the report could be improved.	
	To prevent accidental mistakes in the transfer of the analysis results, it was suggested to provide the possibility of a batch upload of tabulated data (instead of clicking through multiple choices for each individual strain). Another source of mistakes is the assignment of new strain identifiers during results submission (e. g. strain5 ID). This latter point was particularly confusing.	

NGS: next-generation sequencing; QC: quality control.
Ten laboratories responded to the main questions (questions 1–3 and 6). Nine laboratories responded to the whole genome sequencing-related questions (questions 4 and 5).

4 Discussion

Based on the feedback survey results, we believe that most of the participants were satisfied with the format of the EQA scheme, the individual results reports and additional feedback from the EQA provider, as only two participants suggested a different upload system like TSV files. As the evaluation responses are anonymous, it is not possible to follow up on the questions one by one; however, the comment that SeqSphere lists strain5 to be IIa instead of IIc should be investigated before the next round. The inclusion of the modified sequences in the cluster analysis and the QC feedback of the uploaded sequences was well received by the participants.

4.1 Serotyping/serogrouping

Eighteen (86%) laboratories participated in the serotyping part of EQA-10, all 18 (100%) provided molecular serogrouping results and two (11%) conducted conventional serotyping.

4.1.1 Conventional serotyping

The number of participants that conducted conventional serotyping decreased from 10 laboratories in EQA-1 to two in EQA-10, highlighting the transition towards the use of molecular serogrouping. None of the laboratories only used conventional serotyping. The performance was above 85% (range: 87–100%) in EQA-1 to EQA-7, decreased below 85% in EQA-8 and EQA-9 (82% and 80%, respectively) and was 100% in EQA-10.

4.1.2 Molecular serogrouping

Since EQA-2, the number of participants performing molecular serogrouping has ranged from 13 to 18 participants. From EQA-6 to EQA-8, three laboratories reported the use of in silico PCR (WGS) serogrouping, and this increased to 10 in EQA-9 and 12 in EQA-10. In recent years, the lower number of strains to sequence likely made it possible or more attractive for a higher number of laboratories to participate. Regarding molecular serogrouping, performance was consistently high in both EQA-9 and EQA-10, with 95–96% accuracy. The transition from conventional serotyping to molecular serogrouping has reached a point where molecular serogrouping can be considered best practice at NPHRLs in the EU/EEA.

4.2 Molecular typing-based cluster analysis

The adjustment of the EQA to focus on WGS rather than PFGE appears to have been well accepted by the countries, as 18 of the 21 laboratories (86%) participated in the cluster analysis, which is three more than in EQA-8. Only one laboratory participated in cluster identification using PFGE as the sole method, while the remaining 17 laboratories participated using WGS-derived data.

4.2.1 PFGE-derived data

Of the 18 laboratories participating in the cluster analysis, only one (6%) performed cluster analysis using PFGE-derived data. As the criteria of the predefined cluster was based on WGS-derived data, the correct cluster delineation can be difficult to obtain using a less discriminatory method. The participant correctly identified the three cluster strains (among the seven strains in the package).

The number of participants only submitting cluster analysis based on PFGE-derived data has decreased with each of the EQAs and this time 94% (17/18) submitted analyses based on WGS-derived data.

4.2.3 WGS-derived data

Of the 18 participating laboratories, 17 (94%) performed cluster analysis using WGS-derived data. Overall, there has been increased participation since the cluster analysis part of the EQA was introduced. In EQA-5, 12 laboratories participated in WGS-based cluster analysis. Since then, the number of participants has varied but increased overall. In addition, in EQA-10 two laboratories participated for the first time using WGS-derived data. Almost all laboratories (94%) reported that the sequencing was done at their own premises. Most participants (16/17) also reported using an Illumina platform and all reported using commercial kits for library preparation.

The EQA provider's QC evaluation of the raw reads submitted by the participants showed good quality data. Only five participants received warnings from the Bifrost QC pipeline. Two participants received warnings as the 'average coverage' was below the threshold of 50 for multiple sequences that were within the range of 26–43; the cut-off of 50 in the Bifrost pipeline is very strict. Two additional laboratories had the same warning, but only on one sequence each. The fifth laboratory received a warning due to the 'Length at >25x' coverage being below 2.8; however, as the participant submitted Ion Torrent data, the EQA provider acknowledges that some of the QC values provided by Bifrost are unreliable due to assembly issues for Ion Torrent data.

As in previous years, the main QC parameters reported in EQA-10 were coverage threshold and the control of genus/species confirmation. The percentage of participants using assessment of the genome size has increased from 71% in EQA-8 to 88% in EQA-10, and only two of the participants using allele-based analysis did not use the number of cgMLST alleles as a QC parameter.

Sixteen laboratories (94%) reported using an allele-based method as the main analysis and one laboratory (6%) reported using SNP analysis. Compared with EQA-6, there was a 75% increase in the use of allele-based analyses as the main analysis. Over the course of these EQAs, both Laboratory 56 and 105 changed their main analyses from SNP- to allele-based, and two new laboratories also selected allele-based analysis as their main approach.

As in previous EQAs, many participants (57%) used the Ruppitsch cgMLST scheme for the main analysis.

For the second time in this EQA series, participants had to include the EQA-provided sequences in the main cluster analysis. All (100%) of the laboratories identified the sequence (strain10) contaminated with *L. welshimeri,* and only one used the sequence in their analysis (QC status B). This is a much higher identification rate compared with EQA-9, where 69% identified the sequence contaminated with 9% *L. innocua*. Almost all laboratories (94%) correctly identified QC issues, such as low coverage of the sequence strain11, and 81% consequently disregarded this sequence from the analysis. Comparable to the levels observed in EQA-9, the slightly modified strain14, with a 94% core percentage, divided the participants: six (35%) reported QC status A, eight (47%) reported QC status B, and three (8%) reported QC status C. A similar rate was observed in EQA-9, with seven (44%) reporting the sequence with a 94% core percentage as QC status A.

The new modification in EQA-10 was a cluster sequence, strain13, with low read quality of R2. Of the 17 participants that conducted this analysis, seven (41%) reported the sequence to be unacceptable quality (QC status C), indicating that the sequence was not analysed. Since it was a cluster sequence, disregarding it had an impact on the reported cluster. However, a single-end assembly of the good read quality R1 could still be used. Twelve participants exclusively used assembly-based allele calling, and they might benefit from undertaking this challenge in the future.

Sequence strain14 was a non-cluster strain with a slightly reduced coverage. This presented a challenge for participants, who assigned different QC statuses. Fourteen laboratories (82%) accepted the sequence either as QC status A or B, while three laboratories discarded it, citing that the cgMLST core percentage fell below their specified threshold. In EQA-9, only one laboratory discarded the modified sequence with a 94% core percentage (Annex 2, Table 10A–14A).

All 17 laboratories (100%) reported QC issues for sequence strain15, identifying a mix of two sequence types. As a result, 94% of participants excluded the sequence from their analysis, an increase from 75% in EQA-9.

This year, Laboratory 138 (a new participant as of EQA-9) reported results using an allele-based approach instead of SNP based. The analysis clearly showed a separation between the correct cluster (0−12 AD) and the closest non-cluster ST5 strain4 with 27 SNPs; however, mistakenly, the non-cluster strain14 (65 ADs) was listed as a cluster strain, which the EQA provider concluded was a submitting error. Comparing the reported AD, they were similar to those of Laboratory 142 and slightly higher than most other participants' reported AD. Laboratory 142 used the only assembly-based BIGSdb-Lm (1748 loci) with the SPAdes assembler and reported 0–13 AD within the predefined cluster, compared with the most commonly reported 0–8 AD. The highest difference was observed in the non-cluster sequence strain14 (94% core percentage): 165 AD. This higher AD was also identified in EQA-9 when using BIGSdb-Lm. In EQA-9, a participant proposed an explanation suggesting that the calling of new alleles (likely partial alleles) and the inclusion of missing loci were counted as AD. Additionally, Laboratory 187 (a new laboratory as of EQA-10) exhibited slightly higher AD. However, given their use of Ruppitsch with 2866 loci, this outcome was anticipated.

In addition, Laboratory 177 – participating for the first time – only reported 0–2 AD within the predefined cluster, compared with the majority of 0–8 AD. They used only assembly-based BIGSdb-Lm with the SKESA assembler, but with a lower number of loci (1660) than normally reported by BIGSdb. SKESA is known to be a fast but not accurate assembler that usually requires above 40x to reach an error rate of <2.0% [12]. According to the QC evaluation of the participant's raw reads (Annex 3), they received a warning response for six of the seven strains, as the average coverage was below 50 (range: 27–42). The laboratory reported all ST5 to be a part of the cluster, with only 0–10 AD, compared with the 0–57 AD reported by most other participants. Laboratory 129 also used a lower number of loci (1503); however, they reported 0–3 AD within the cluster and 0–22 AD for all ST5 strains/sequences.

From the submitted raw sequences, the EQA provider concluded that Laboratory 187 correctly analysed the sequence; however, it made errors when reporting, mixing up the results of four strains. Additionally, Laboratories 35 and 177 presumably made submitting errors.

In this challenging EQA, most (11/16) participants reported 6–8 AD for strain7 and 7–8 AD for strain9, compared with the selected reference strain16. Those with borderline AD influenced the reported cluster. Most participants considered these strains to be part of the cluster, with one laboratory (Laboratory 135) even describing strain9 as a

potential case with 8 AD. However, Laboratories 49 and 56 reported 8 AD and 7 AD for strain7 and strain9, respectively, and consequently excluded them from the reported cluster. Microbiology is often highly diverse, and adhering to a fixed cut-off can be convenient. However, as *Listeria* outbreaks sometimes evolve gradually, it may be necessary to include probable cases at the beginning of an outbreak investigation. In this dataset, the AD between strain16 and strain12 was 3, while the AD between strain12 and strain7 was 5 AD and between strain12 and strain9 was 4 AD (according to the EQA provider's analysis). From the literature, it is known that Palma and colleagues concluded in 2022 that all examined workflows (BIGSdb, INNUENDO, GENPAT, SeqSphere, MentaLIST) require a depth of coverage >40 and high loci detection >99.54% (BioNumerics only 97.78%) to maintain consistent cluster definitions when using the reference cut-off of 7 AD [13]. However, data is not always perfect, and sometimes a case-by-case assessment must be conducted.

When the data from Laboratory 108 was analysed by the EQA provider (utilising the standardised cgMLST/Pasteur analysis), it exhibited AD for most of the test strains (Figure 8). This laboratory provided Ion Torrent data, which the EQA provider's analysis is not optimised for, making correct assembly challenging (as observed in previous EQAs for Laboratory 108, but not for Laboratory 56, which also provided Ion Torrent data). Therefore, the observed AD may be method artifacts. However, the use of Ion Torrent data can complicate the communication and investigation of multi-country outbreaks when solely relying on the allelic method.

5 Conclusions

Twenty-one laboratories participated in the EQA-10 scheme, with 18 (86%) performing serogrouping and 18 (86%) performing cluster identification. It was very encouraging to see an increase in participation for both serogrouping and cluster analyses compared with previous EQAs, as EQA-10 had two new participants.

Most laboratories (16/18; 89%) performed only molecular serogrouping, while 11% (2/18) performed molecular serogrouping in combination with conventional serotyping. In general, there has been a trend towards replacing conventional serotyping with molecular serogrouping throughout the 10 EQAs. The average quality of conventional serotyping was the lowest last year (EQA-9), at 80%; however, in EQA-10, the performance was high (100%). The performance of the molecular serogrouping was also good, at 95%. The transition from conventional serotyping to molecular serogrouping has reached a level where molecular serogrouping can be considered the best practice in NPHRLs across the EU/EEA.

One laboratory used PFGE for cluster analysis. Since the cluster pre-categorisation was based on WGS data, it was expected that it might be difficult to obtain the correct cluster delineation using less discriminatory methods, such as PFGE. However, the participant successfully identified the three cluster strains among the shipped strains but could not include the provided sequences in their cluster analyses.

In general, most of the participants were able to identify the different characteristics and modifications of the EQA-provided sequences. For the sequence with very low coverage, 94% of the participants identified the QC issues and, consequently, 76% did not proceed with the cluster analysis. However, there were disagreements among the participants, who did not reach a consensus on the QC status of the modified cluster sequence with slightly reduced coverage, given the subjective nature of this threshold and its dependence on experience. Additionally, the challenging modified read quality of the cluster sequence strain13 led to varying reported QC statuses and strains being identified as part of the cluster. All laboratories identified issues with the sequence, noting a combination of ST7 and ST224.

Seventeen laboratories performed cluster analysis using WGS-derived data, one using SNP and 16 (94%) using allele-based analyses. The results are encouraging when evaluating the reported number of AD per strain. Most (11/16) reported 0–8 AD in the predefined cluster; however, not all included a strain with 7 or 8 AD in the reported cluster. Additionally, three laboratories reported a slightly increased number of AD.

When analysing the predefined cluster of the seven closely related strains, 47% (8/17) reported the same list of strains; however, two laboratories excluded the modified strain13 from the analysis, thus omitting it from the reported cluster, and two laboratories excluded the borderline strain7 and strain9, as they exceeded their cut-off. Therefore, 71% (12/17) of the participants reported results that were in line with the expected outcome. Three of the laboratories that did not identify the predefined cluster were suspected to have made submitting errors.

In general, most of the reported cgMLST results for participants' own sequence data and the non-modified sequences were at a comparable level of AD (0–8) within the cluster strains, despite being analysed using different schemes. The reported SNP result showed the laboratory was able to identify the correct cluster. Therefore, both methods seem to work for cluster detection, even though it was less obvious than in previous EQAs. Furthermore, standardised cgMLST analyses leave little room for error when only including good quality data, which results in good inter-laboratory comparability.

6 Recommendations

6.1 Laboratories

It is recommended that laboratories use EQA-provided data and strains to validate their analysis methods if incorrect results are obtained (e.g. participate in EQA) or if implementing new methods and procedures.

When laboratories re-name/change the strains from the EQA-provided ID to an ID that fits into their pipelines, it might be useful to introduce a control procedure.

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 such as EQA schemes, expert exchange visits and workshops.

6.3 EQA provider

The EQA has evolved over the years, as the EQA provider has included additional sequences and modified some to introduce QC issues. Positive feedback from the evaluation indicates that this approach should be continued in future EQAs. The QC assessment of the participants' submitted sequences was also appreciated. In future, the EQA provider will assess whether the online submission form could be made more user-friendly. The EQA provider will also investigate why SeqSphere indicated that strain5 was serogroup IIa.

In EQA-9, one laboratory requested that EQA schemes for different pathogens be more separated across the year; however, this year another participant indicated that it would prefer to have all EQA strains shipped at the same time, so they could fill up a WGS run with EQA samples instead of having more WGS runs.

References

1. European Commission (EC). 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. Brussels: EC; 2004. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32004R0851
2. European Commission (EC). Regulation (EU) 2022/2371 of the European Parliament and of the Council of 23 November 2022 on serious cross-border threats to health and repealing Decision No 1082/2013/EU (Text with EEA relevance). Brussels: EC; 2022. Available from: http://data.europa.eu/eli/reg/2022/2371/oj
3. European Food Safety Authority (EFSA)/European Centre for Disease Prevention and Control (ECDC). The European Union One Health 2022 Zoonoses Report. EFSA J. 2023;21:e8442 Available from: https://doi.org/10.2903/j.efsa.2023.8442
4. International Organization for Standardization (ISO). 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. 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.
7. 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
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 (SSI). Serum QC "Bifrost" [Internet, software package]. Copenhagen: SSI; 2020 [cited 24 January 2020]. Available from: https://github.com/ssi-dk/bifrost
11. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biology. 2014;15:R46.
12. Liu YY, Chen BH, Chen CC, Chiou CS. Assessment of metrics in next-generation sequencing experiments for use in core-genome multilocus sequence type. PeerJ. 2021 Aug 19;9:e11842.
13. Palma F, Mangone I, Janowicz A, Moura A, Chiaverini A, et al. In vitro and in silico parameters for precise cgMLST typing of *Listeria monocytogenes*. BMC Genomics. 2022; Mar 26;23(1):235. Available from: https://doi.org/10.1186/s12864-022-08437-4

Annex 1. Participation details

Table 1A. List of participants in EQA-10

Country	Laboratory	National institute
Austria	NRL Listeria Austria	Austrian Agency for Health and Food Safety (AGES)
Belgium	NRC Listeria	Sciensano
Germany	Consultant Laboratory for Listeria	Robert Koch Institute
Denmark	Laboratory of Gastrointestinal Bacteria	Statens Serum Institut
Spain	Neisseria, Listeria and Bordetella Unit	National Centre for Microbiology, Instituto de Salud Carlos III
Finland	Expert microbiology	Finnish Institute for Health and Welfare (THL)
France	National Reference Centre and WHO Collaborating Centre Listeria	Institut Pasteur
Greece	Reference Centre for Salmonella, Shigella, Listeria	University of West Attica
Croatia	Virology Department	Croatian Institute of Public Health
Hungary	FWD – Laboratory	National Center for Public Health and Pharmacy
Ireland	National Salmonella, Shigella & Listeria Reference Laboratory (NSSLRL)	Galway University Hospital (HSE)
Italy	Department of Food Safety, Nutrition and Veterinary Public Health	Istituto Superiore di Sanità
Lithuania	Rare and dangerous infectious agents investigation division	National Public Health Surveillance Laboratory
Luxembourg	Epidemiology and Microbial Genomics (EPIGEM)	Laboratoire National de Santé
Latvia	Laboratory Service, National Microbiology Reference Laboratory	Riga East University Hospital
Malta	Molecular Diagnostics	Mater Dei Hospital
Netherlands	IDS	RIVM
Norway	National Reference Laboratory for Enteropathogenic Bacteria	Norwegian Institute of Public Health
Portugal	URGI	National Institute of Health Dr. Ricardo Jorge (INSA)
Sweden	Unit for laboratory surveillance of bacterial pathogens	Public Health Agency of Sweden
Slovakia	Department of Environmental Microbiology	Regional Public Health Authority Kosice

Table 2A. Overview of participation in EQA-9 and EQA-10

Laboratory no.	EQA-9 (2020–2021)					EQA-10 (2022–2023)				
	All[a]	Serotyping/group		Cluster		All[a]	Serotyping/group		Cluster	
		Conventional	Molecular	PFGE	WGS		Conventional	Molecular	PFGE	WGS
19	X		X		X	X		X		X
35	X		X		X	X		X		X
49	X		X		X	X		X		X
56	X				X	X		X		X
70	X		X		X	X		X		X
88	X		X		X	X				X
96	X		X			X		X		
100	X	X	X		X					
105	X		X		X	X		X		X
108	X		X		X	X		X		X
114	X	X		X						
129	X		X		X	X		X		X
130						X		X	X	
135[b]	X				X	X				X
138	X	X	X		X	X		X		X
141	X		X		X	X		X		X
142	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	X	X		
149	X		X		X	X		X		X
177						X				X
187						X		X		X
Total no. of participants	**20**	5	17	2	16	**21**	2	18	1	17

[a] *Participation in at least one element of the EQA*
[b] *Previously Laboratory 77*
X indicates participation.

Annex 2. Reported results

Table 3A. Result scores for conventional serotyping

Strain no.	Provider	Laboratory no.	
		142	145
Strain1	4b	4b	4b
Strain2[a]	1/2b	1/2b	1/2b
Strain3[a]	1/2b	1/2b	1/2b
Strain4	1/2b	1/2b	1/2b
Strain5	1/2c	1/2c	1/2c
Strain6[a]	1/2a	1/2a	1/2a
Strain7	1/2b	1/2b	1/2b

[a] Technical duplicates (triplicates with sequence strain16)
Purple shading: repeat strains that were in EQA-1 to EQA-10 (strain1, strain5 and strain6). Strain6 was a different strain from the one used in previous years, but the same serotype/group.

Table 4A. Result scores for molecular serogrouping

Strain no.	Provider	Laboratory no.																	
		19	35	49	56	70	96	105	108	129	130	138	141	142	143	144	145	149	187
Strain1	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IVb	IIb
Strain2[a]	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb
Strain3[a]	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IVb
Strain4	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb
Strain5	IIc	IIc	IIa	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIc	IIb
Strain6[a]	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IVb	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa	IIa
Strain7	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIb	IIc
Method	WGS	WGS	WGS	WGS	WGS	WGS	PCR	WGS	WGS	WGS	PCR	WGS	WGS	PCR	PCR	PCR	PCR	WGS	WGS

[a] Technical duplicates (triplicates with sequence strain16)
Purple shading: repeat strains that were in EQA-1 to EQA-10 (strain1, strain5 and strain6). Strain6 was a different strain from the one used in previous years, but the same serotype/group.
Pink shading: incorrect results.

Table 5A. Reported cluster of closely related strains based on PFGE-derived data

Laboratory no.	Reported cluster	Corresponding strains	Included the three strains from the WGS cluster[a]	Included additional strains in the cluster
EQA Provider	NA	strain2, strain3, strain7	Yes	
130	1002 1071 1535	strain2, strain3, strain7	Yes	No

NA: not applicable.
PFGE: Pulsed-field gel electrophoresis; WGS: whole genome sequencing.
[a] Predefined categorisation using WGS-derived data.

Table 6A. Reported sequencing details

Laboratory no.	Sequencing performed	Protocol (library prep)	Commercial kit	Sequencing platform
19	In own laboratory	Commercial kits	Nextera XT Kit (Illumina)	NextSeq
35	In own laboratory	Commercial kits	Nextera	NextSeq
49	In own laboratory	Commercial kits	Illumina DNA Prep and V3 600 Cycle flowcell and cartridge	MiSeq
56	Externally	Commercial kits	Illumina DNA prep (M) Tagmentation 96 samples	MiSeq
70	In own laboratory	Commercial kits	DNA Prep Illumina[a]	MiniSeq Illumina
88	In own laboratory	Commercial kits	Nextera XT DNA Library Prep, Illumina#15031942 and Denature and Dilute LibrariesIllumina#15048776)[b]	NextSeq
105	In own laboratory	Commercial kits	Illumina DNA Prep, (M) Tagmentation (24 samples, IPB)	MiSeq
108	In own laboratory	Commercial kits	Ion XpressTM Plus Fragment Library Kit for AB Library BuilderTM System (Thermo Fisher Scientific)	Ion S5 XL system (Thermo Fisher Scientific)
129	In own laboratory	Commercial kits	Illumina Nextera XT	MiSeq
135	In own laboratory	Commercial kits	Illumina DNA Prep	NextSeq
138	In own laboratory	Commercial kits	Illumina DNA Prep	NextSeq
141	In own laboratory	Commercial kits	Illumina DNA Prep	MiSeq
142	In own laboratory	Commercial kits	Miseq Reagent Kit v3 - 500 cycles	MiSeq
144	In own laboratory	Commercial kits	DNA Prep	NextSeq
149	In own laboratory	Commercial kits	Kapa HyperPlus (Kapa Biosystems)	NextSeq
177	In own laboratory	Commercial kits	Nextera XT	MiSeq
187	In own laboratory	Commercial kits	Illumina DNA Prep	NextSeq

[a] All the volumes of reagents were divided by 2.5.
[b] The commercial protocol described in Nextera XT DNA Library Prep Reference Guide (Illumina, Document # 15031942, v07, May 2023) was followed using half volume of all reagents until the library pooling, but keeping the original amount of DNA (1 ng).

Table 7A. EQA provider cluster analysis, based on WGS-derived data

Laboratory no.	Reported cluster	Corresponding to EQA provider strains	Correct cluster without borderline (strain7) or modified sequence (strain13)	Predefined correct cluster
Provider	NA	Strain2, Strain3, Strain9, Strain12, Strain16, (Strain7, Strain13) (2/3 duplicates)	NA	NA
19	1423, 1984, 1301, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
35	1766, 1555, 1377, 1848, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain4, Strain7, Strain9, Strain12, Strain13, Strain16	No	No
49	1858, 1313, 0012, 0016	Strain2, Strain3, Strain12, Strain16	No	No
56	1094, 1096, 0012, 0016	Strain2, Strain3, Strain12, Strain16	No	No
70	1461, 1186, 1685, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
88	1785, 1177, 1522, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
105	1877, 1967, 1571, 0009, 0012, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain16	Yes	No
108	1566, 1011, 1106, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
129	1187, 1964, 1324, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
135	1184, 1498, 1314, 0012, 0016	Strain2, Strain3, Strain7, (Strain9), Strain12, Strain16	Yes	No
138	1193, 1154, 1832, 0009, 0012, 0014, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain14, Strain16	No	No
141	1514, 1204, 1837, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
142	1538, 1395, 0016	Strain2, Strain3, Strain16	No	No
144	1911, 1896, 1437, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
149	1418, 1904, 1135, 0009, 0012, 0013, 0016	Strain2, Strain3, Strain7, Strain9, Strain12, Strain13, Strain16	Yes	Yes
177	1833, 1429, 1890, 1826, 0009, 0011, 0012, 0013, 0014, 0016, 0017	Strain2, Strain3, Strain4, Strain7, Strain9, Strain11, Strain12, Strain13, Strain14, Strain16, Strain17	No	No
187	1607, 1983, 0009, 0012, 0016	Strain1, Strain2, Strain9, Strain12, Strain16	No	No

NA: not applicable.
Strains8–17 were provided sequences. Strains 10, 11, 13, 14 and 15 were modified by the EQA provider.

Table 8A. Reported SNP distances

Strain no.	ST	Provider	Laboratory no.		
			19[a]	108	177[a]
Strain1	2	NA	NA	15297	560
Strain2[b,c]	5	0	0	0	0
Strain3[b,c]	5	0	0	0	0
Strain4	5	36	36	78	5
Strain5	9	NA	NA	133872	1000
Strain6	7	NA	NA	133366	280
Strain7[b]	5	13	13	14	0
Strain8	2	NA	NA	16283	545
Strain9[b]	5	10	10	11	2
Strain10[d]	NA	NA	NA	NA	540
Strain11[d]	NA	NA	NA	2097	30
Strain12[b]	5	7	7	8	0
Strain13[b,d]	5	0	0	0	0
Strain14[d]	5	276	152	350	10
Strain15[d]	NA	NA	NA	NA	265
Strain16[b,c]	5	0	0	0	0
St'ain17	5	718	144	1020	10

NA: Not analysed; ST: sequence type.
[a] Additional analysis
[b] Closely related strains
[c] Technical triplicates (strain/sequence) in bold
[d] Modified sequences
Strain16 used as the representative to report the AD/SNP distance.

Table 9A. Reported allelic differences

Strain no.	ST	Provider	Laboratory no.															
			19	35	49	56	70	88	105	129	135	138	141	142	144	149	177	187
Strain1	2	1248	1256	1183	1249	1253	1184	1253	1168	119	1187	1191	1187	1264	1182	1186	560	0
Strain2[a,b]	5	0	0	0	0	0	0	0	0	0	0	6	0	4	0	0	0	0
Strain3[a,b]	5	0	0	0	0	0	0	0	0	0	0	5	0	4	0	0	0	1661
Strain4	5	20	21	20	21	22	20	22	19	10	20	27	20	28	20	20	5	27
Strain5	9	1724	1724	1654	1719	1717	1654	1726	1631	178	1659	1654	1657	1734	1652	1659	1000	13
Strain6	7	1720	1723	1653	1717	1718	1654	1724	1631	177	1659	1658	1657	1735	1651	1658	280	2362
Strain7[a]	5	8	8	6	8	8	6	8	6	3	6	10	6	13	6	6	0	2385
Strain8	2	1255	1255	1184	1248	1252	1183	1251	1168	NA	1186	1186	1186	1262	1181	1185	545	1673
Strain9[a]	5	7	7	8	7	7	8	8	8	5	8	12	8	12	8	8	2	10
Strain10[c]	NA	NA	NA	NA	NA	1097	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Strain11[c]	NA	NA	NA	NA	NA	161	NA	NA	NA	NA	NA	181	NA	NA	NA	NA	30	NA
Strain12[a]	5	3	3	4	3	4	4	4	4	2	5	10	4	12	4	4	0	7
Strain13[a,c]	5	0	0	0	NA	NA	0	0	NA	0	NA	NA	0	NA	0	0	0	NA
Strain14[c]	5	55	55	57	55	NA	57	NA	53	27	62	65	NA	167	56	56	10	80
Strain15[c]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	265	NA
Strain16[a,b]	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Strain17	5	51	51	46	51	50	46	51	45	22	45	52	46	57	46	46	10	73

NA: Not analysed; ST: sequence type.
[a] Closely related strains
[b] Technical triplicates (strain/sequence) in bold
[c] Modified sequences
Strain16 used as the representative to report the AD/SNP distance.

Table 10A. Reported quality control statuses of strain10

Laboratory no.	Sero	ST	Cluster	QC status	Description strain10
EQA provider				**C**	**A non-cluster sequence contaminated with approximately 6% *L. welshment***
19				C	Strain 0010 has not acceptable sequence quality and cannot be used in cluster analysis. Strain 0010 is contaminated with *Listeria welshimeri*, resulting in many multiple consensus calls and in a poorly assembled genome (large genome, many contigs, low N50 value and many unidentified bases (Ns)). Strain 0010 has to be restreaked for pure culture and resequenced.
35				C	91x coverage, 72.8% good cgMLST targest, 463 missing alleles, 5.8 Mbp approximated genome size -> contamination possible? *Listeria welshimeri* contamination detected
49				C	Genome Length 5.75 MB, numerous small contigs, rMLST 75% *Listeria monocytogenes*, 25% *Listeria welshimeri* query contamination. Would repeat sequencing if this was sequenced in house.
56	IVb	217	No	B	good coverage 197X but only 1514 genes mapped in cgMLST
70				C	Contaminated above 10% with *L. welshimiri*- Mixed culture
88				C	OVERALL QC: FAILED (contamination / strain admixture)
105				C	Strain 0010 FAILS several QCs: Genome size was 5.5 Mbases, cgMLST percentage of good targets found was only 75.7%.
108				C	The sample is contaminated with *Listeria welshimeri*. The genome size is too large.
129				C	Sequence quality was poor. Percentage of good targets was 70.1 and Ridom did not give any MLST type or serogroup. In Rematch the serogroup was IVb. String MLST 217
135				C	Too many contigs, total length too high, GC% too low, N50 too low, contamination too high
138				C	QC - Fail: contigs (>=0bp) - 443 ($x \leq 212$), average coverage - 96.33 ($x > 30$), GC% - 37.08 ($37.8 < x \leq 38$), N50 - 27701 ($x > 60010$), assembly length - 5763901 ($2918000 < x \leq 3156000$. Contamination with *Listeria welshimeri*
141				C	QC not OK: contaminated 75.1% good targets - too low, av. Coverage (assembled): 76 OK, contig count: 726, appr. genome size: 5.8 too big possible contamination with *Listeria welshimeri*
142				C	Contamination with *L welshimeri*. Double allele calling. Low % cgMLST calling
144				C	Not acceptable quality, parameters considered: N50: 10 975, approx. genom size: 5.6 Mb (> 3.3 Mb), contig count: 1062 (>200), aver. coverage: 93x cgMLST good target (%): 76.9% (< 95%), Probably contaminated sample with *Listeria welshimeri* (ID 98%)
149				C	<90% good targets: 76.9%, Wrong genome size: 5.6 M, N50 too low (10.975), no of contigs too high (1062), KRAKEN and SeqSphere: contamination with *Listeria welshimeri* (mix of *L. monocytogenes* and *L. welshimeri*)
177				C	QC thresholds used: Genome size 2.7 - 3.2 Mb, N50: 30,000bp, Number of Contigs: 200, Strain 10: Genome Size: > 3.2Mb, N50: < 30,000bp Number of Contigs: >200
187				C	It's a fail, too large genome size, 2 different *Listeria*

QC: quality control; ST: sequence type; Sero: serogroup.

Table 11A. Reported quality control statuses of strain11

Laboratory no.	Sero	ST	Cluster	QC status	Description strain11
EQA provider				**C**	**A nonCluster sequence with low coverage**
19				C	Strain 0011 has not acceptable sequence quality and cannot be used in cluster analysis. Strain 0011 has low read coverage (< 50) resulting in poor coverage of the cgMLST scheme (low cgMLST core%, < 95%) and in a poorly assembled genome (many contigs, low N50 value and many unidentified bases (N's)). Strain 0011 has to be resequenced.
35				C	30x coverage, 60% good cgMLST targets, 684 missing alleles
49				C	Core percent = 67, Numerous small contigs, Coverage of 31, much lower than we normally get. Would repeat sequencing if this was sequenced in house.
56	IIb	5	No	B	low coverage 32X and only 1534 genes mapped
70				C	Bad quality, number of contigs >200, N50<30 000, Coverage <50, Number of targets found cgMLST <90%
88				C	OVERALL QC: FAILED (low depth of coverage yielding a low % cgMLST loci called - 78.7%, i.e., below 95%)
105				C	Strain 0011 FAILS the QC: - cgMLST percentage of good targets found was only 61.8%.
108	IIb	5	No	A	No deviations observed.
129				C	Sequence quality was poor. Percentage of good targets was 66.1 and Avg. Coverage was 27. Ridom did not give any serogroup. In Rematch the serogroup was IIb.
135				C	Number of contigs too high, GC% too high, N50 too low, coverage too low
138	1/2b, 3b, 7 - IIb	5	No	B	QC - Fail: contigs >=0bp - 682 ($x \leq 212$) average coverage - 29.91 ($x > 30$) GC% - 38.16 ($37.8 < x \leq 38$) N50 - 6682 ($x > 60010$) assembly length - 2924268 ($2918000 < x \leq 3156000$), Undersequenced. Possibly due to low input DNA arising from problems in extraction or storage as there are also signs of a highly fragmented genome. Still typable.
141				C	QC not OK: 89.4% good targets - too low, av. Coverage (assembled): 28 – low, contig count: 767 appr. genome size: 3.0 OK
142				C	Low coverage. Low % cgMLST calling
144				C	Not acceptable quality, parameters considered: N50: 3600, approx. genom size: 2.8 Mb, contig count: 1415 (>200), aver. coverage: 32x (<50x), cgMLST good target (%): 70.7% (< 95%)
149				C	<90% good targets: 70.7%, Correct genome size: 2.8 Mb, Low coverage (31). N50 too low (3.600), no of contigs too high (1415), KRAKEN: low coverage, *L. monocytogenes*
177		5	Yes	B	QC thresholds used: Genome size 2.7 - 3.2 Mb, N50: 30,000bp, Number of Contigs: 200, Strain 11: Number of Contigs: >200 N50: < 30,000bp
187				C	The assembly is not good to many contigs and bad N50 and i50

QC: quality control; ST: sequence type; Sero: serogroup.

Table 12A. Reported quality control statuses of strain13

Laboratory no.	Sero	ST	Cluster	QC status	Description strain13
EQA provider				**B**	**A cluster sequence (strain2/3) with low read quality of R2**
19	IIb	5	Yes	B	Strain 0013 is only accepted for outbreak situations. The forward and reverse read file is not of same size. This is likely a result of error in the file transfer. An assembled genome was made from one read file only and cgMLST analysis was made on assembly-based calls only. For SNP analysis both read files can be used as input. However, Strain 0013 has to be either resequenced or examined for correct file transfer.
35	IIb	5	Yes	A	129x coverage, 98.5% good cgMLST targets 25 missing alleles
49				C	Quality score of 18
56				C	Error in mapping
70	IIb	5	Yes	B	No Reverse raw reads in the assembly but with forward raw reads, number og contigs are less than 200, N50 is above 30 000, size of the genome and coverage (130) are OK. Number of targets found = 98.6%
88		5	Yes	B	OVERALL QC: FAILED (R2 without quality data); assembly: OK; contamination checking: OK; % cgMLST loci called: OK (>=95%)
105				C	Strain 0013 FAILS the QC "Q Score (Phred)".
108	IIb	5	Yes	A	No deviations observed.
129	IIb	5	Yes	A	Percentage of good targets and average coverage were acceptable. QC passed.
135				C	QC could not be performed, pipelines reported "empty files" when attempting to process downloaded fastq files
138				C	QC - Fail: Very low quality read2. Either corrupted file during upload or sequencing failed on read2.
141	IIb	5	Yes	A	QC OK: 98.7% good targets OK, av. Coverage (assembled): 77 OK, contig count: 209, appr. genome size: 3.0 OK
142				C	Poor quality Reverse sequences (only short reads)
144	IIb	5	Yes	A	A, parameters considered: N50: 62 852, approx. genom size: 3.0 Mb, contig count: 129, aver. coverage: 130x, cgMLST good target (%): 98.5%
149	IIb	5	Yes	A	>90% good targets: 98.5%, Good coverage: 134 (raw reads), Correct genome size: 3 Mb, N50 (62.852) and number of contigs (129): ok, KRAKEN: *L. monocytogenes*. We accept the quality but noticed the following: SeqSphere: fastQC Per Base Sequence Quality (Reverse read) failed. We also had problems making fasta file in Spades due to low quality of the reverse read
177		5	Yes	A	QC thresholds used: Genome size 2.7 - 3.2 Mb, N50: 30,000bp, Number of Contigs: 200, Strain 13: All parameters within the threshold limits
187				C	Bad fastq r2 file all reads are below 20

QC: quality control; ST: sequence type; Sero: serogroup.

Table 13A. Reported quality control statuses of strain14

Laboratory no.	Sero	ST	Cluster	QC status	Description strain14
EQA provider				**B**	**A nonCluster with core percentage approximately 94%**
19	IIb	5	No	B	Strain 0014 is only accepted for outbreak situations. The cgMLST core% is below 95% (93%) and the strain has to be resequenced for better core%.
35	IIb	5	No	B	111x coverage, 94% good cgMLST targets, 95 missing alleles, assembly base count 2.85 Mbp (somewhat below average)
49		5	No	B	Quality= 34, Coverage >30, Length between 2.6-3.2MB, Core Percent 93%. Would repeat this if this was sequenced in house
56				C	Good coverage 110X but only 1598 genes mapped in cgMLST
70	IIb	5	No	B	Low QC, number of contigs, N50, size of the genome and coverage OK but the % of targets found <95%
88				C	OVERALL QC: FAILED (potential contamination / strain admixture and low % cgMLST loci called - 93.4%, i.e., below 95%)
105	IIb	5	No	B	In Strain 0014 the QC "cgMLST percentage of good targets" was 92.9% (WARN).
108	IIb	5	No	B	The analysis shows very small traces of other *Listeria* species but it is most likely not relevant. The genome size is within threshold and there are no other deviations. To be sure we would rerun the sample in routine analysis.
129	IIb	5	No	A	Percentage of good targets and average coverage were acceptable. QC passed.
135		5	No	B	completeness (CheckM) too low (94.85%, threshold >96%), all other criteria were within range
138	1/2b, 3b, 7 - IIb	5	Yes	A	QC - Warn: contigs >=0bp - 137 (x ≤212), average coverage - 110.2 (20 < x ≤ 30), GC% - 37.85 (37.8 < x ≤ 38), N50 - 40154 (x > 60010), assembly length - 2832365 (2918000 < x ≤ 3156000), Overall good quality but the N50 is relatively low, possible issues with DNA quality, still usable for typing.
141				C	QC not OK: 94.4% good targets - too low, av. Coverage (assembled): 106 OK contig count: 224 appr. genome size: 2.9 OK
142	IIb	5	No	B	Low % cgMLST calling
144	IIb	5	No	A	A, parameters considered: N50: 40 596, approx. genom size: 2.8 Mb, contig count: 165, aver. coverage: 113x, cgMLST good target (%): 93.4% (< 95%) It is a bit less then the expected, but we accept it because the difference is minimal.
149	IIb	5	No	A	>90% good targets: 93.4%, Good coverage: 108 (raw reads), Correct genome size: 2.8 Mb, N50 (40.596) and number of contigs (165): ok, KRAKEN: *L. monocytogenes*
177		5	Yes	A	QC thresholds used: Genome size 2.7 - 3.2 Mb, N50: 30,000bp, Number of Contigs: 200, Strain 14, All parameters within the threshold limits
187	1/2b-3b-7	5	No	A	Assemblies were good, n50 and i 50 are good, genome size is ok average cov is good

QC: quality control; ST: sequence type; Sero: serogroup.

Table 14A. Reported quality control statuses of strain15

Laboratory no.	Sero	ST	Cluster	QC status	Description strain15
EQA provider				**C**	**A nonCluster sequence of ST7 and ST224 combined**
19				C	Strain 0015 has not acceptable sequence quality and cannot be used in cluster analysis. Strain 0015 is likely contaminated with other *Listeria monocytogenes*, resulting in many multiple consensus calls and in a poorly assembled genome (large genome, many contigs, low N50 value and many unidentified bases (N's)). Strain 0015 has to be restreaked for pure culture and resequenced.
35				C	180x coverage, 52% good cgMLST targets, 821 missing alleles, 5.7 Mbp appr. genome size -> possible contamination
49				C	Core percent = 36, Numerous small contigs, Genome Length 5.2MB, NrBAFMultiple 298, Would repeat sequencing if this was sequenced in house
56				C	Error in mapping cgMLST
70				C	QC is bad and suggests a mix culture. Number of contigs far exceeds 200, N50 is far below 30 000, size of the genome is around twice the normal size expected (5.7 Mb). Number of targets found is around 40%.
88				C	OVERALL QC: FAILED (contamination / strain admixture)
105				C	Strain 0015 FAILS several QCs: Genome size was 4.2 Mbases, cgMLST percentage of good targets found was only 60.1%
108				C	Too large genome size. Possibly contaminated with another strain of *Listeria monocytogenes*. Not able to define sequence type.
129				C	Sequence quality was poor. Percentage of good targets was 24.8 and Ridom did not give MLST type or serogroup. In Kraken analysis, the strain was *Listeria monocytogenes*, but Rematch gave contradictory serogroup. StrinMLST 0 abcZ
135				C	number of contigs too high, total length too high, GC% too high, N50 too low, contamination too high
138				C	QC - Fail: contigs >=0bp - 1629 (x ≤ 212) average coverage - 132.08 (20 < x ≤ 30) GC% - 38.2 (37.8 < x ≤ 38) N50 - 3565 (x > 60010) assembly length - 4147836 (2918000 < x ≤ 3156000) Probably heavily degraded DNA.
141				C	QC not OK: contaminated, 42.6% good targets - too low, av. Coverage (assembled): 76 OK, contig count: 6157 - too many, appr. genome size: 5.7 too big possible contamination with another *Listeria monocytogenes*; multiple MLST hits found
142				C	Double calling cgMLST. Mix *L. monocytogenes*
144				C	Not acceptable quality, parameters considered: N50: 1214, approx. genom size: 4.3 Mb (> 3.3 Mb), contig count: 4949 (>200), aver. coverage: 100x, cgMLST good target (%): 15.6% (< 95%)
149				C	<90% good targets: 15.9%, Too large genome size: 4.3 Mb, N50 too low (1.214) and number of contigs too high (4949), KRAKEN: *L. monocytogenes* (two hits on each gene in 7 MLST scheme, probably two *L. monocytogenes* genomes)
177		7	No	B	QC thresholds used: Genome size 2.7 - 3.2 Mb, N50: 30,000bp, Number of Contigs: 200, Strain 15: Number of Contigs > 200, N50: < 30,000bp
187				C	bad quality, 2 different mlst types of *Listeria*

QC: quality control; ST: sequence type; Sero: serogroup.

Table 15A. Additional reported quality control parameters

Laboratory no.	1		2		3		4		5	
	Parameter	Threshold	Parameter	Threshold	Parameter	Threshold	Parameter	Threshold	Parameter	Threshold
19	N50	Available from QC analysis but no threshold	Number of contigs	Available from QC analysis but no threshold	Number of unidentified bases (N) or ambigiues sites	Available from QC analysis but no threshold				
35	number of missing alleles (<100)	N50>20000								
49	Core Percent	>97% pass, 90-96% Warn, <90% fail	NrBAF Multiple	<20 pass, >20 further investigation required	N50		NrContigs	If high investigate further	NrBAFPerfect	
70	number of contigs	<200	N50	>30 000						
88	Inter- and intra-species contamination	INNUca (using default Kraken PASS criteria) and ConFindr (<=3 NumContamSNVs)	Number of contigs	INNUca_v4.2.2 (default setting)						
135	number of contigs	<= 300	GC%	between 37.6 and 38.2%	N50	>=30000	completeness (CheckM)	>96%	contamination (CheckM)	<4%
138	N50	x > 60010	GC%	37.8 < x ≤ 38	number of contigs >=0bp	x ≤ 212	rMLST_Support_%	>90% of alleles		
141	contig size	200 - contigs shorter than 200 bases were ignored								
144	N50	>100 kb	contig count	<200						
149	N50 >30.000	No. of contigs < 500								
177	N50	> 30,000 bp	Number of Contigs	< 200	Completeness of the genome	98%				
187	N50	>20kb	contamination	kmerfinder	Number of unidentified bases (N) or ambigiues sites	Available from QC analysis but no threshold				

Annex 3. Calculated qualitative and quantitative parameters

Table 16A. Calculated qualitative/quantitative parameters for Laboratory 19

Qualitative/quantitative	QC ranges	Laboratory 19						
		1017	1301	1423	1472	1658	1835	1984
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		92.0	92.8	92.7	93.3	92.9	93.4	92.4
% Species 2	{<5%}	0.1	0.0	0.0	0.1	0.1	0.0	0.0
Unclassified reads (%)		7.7	7.1	7.2	6.6	6.9	6.5	7.5
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	3.0	2.9	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	45	36	29	23	36	21	32
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	159	173	164	135	147	147	145
No. of reads (x 1000)		3641	3909	3672	2967	3254	3073	3246
Average read length		135	136	136	139	139	140	136
Average insert size		218	214	223	236	227	238	222
N50 (kbp)		173	175	316	435	177	430	191
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 17A. Calculated qualitative/quantitative parameters for Laboratory 35

Qualitative/quantitative	QC ranges	Laboratory 35						
		1034	1374	1377	1555	1766	1848	1865
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		96.5	97.0	96.4	97.2	97.1	96.9	97.3
% Species 2	{<5%}	0.1	0.0	0.1	0.0	0.0	0.0	0.0
Unclassified reads (%)		3.3	3.0	3.3	2.8	2.8	3.0	2.6
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	2.9	3.0	3.0	3.0	3.0	2.9
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	8.7	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	46	39	112	67	44	35	32
No. of contigs [1-25] x min. coverage	{<1000}	0	0	9	0	0	0	0
Average coverage	{>50}	294	319	182	334	331	340	306
No. of reads (x 1000)		6028	6421	3761	6764	6736	6958	3269
Average read length		148	147	148	149	148	148	276
Average insert size		327	311	324	349	337	327	372
N50 (kbp)		132	189	66	98	162	162	172
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 18A. Calculated qualitative/quantitative parameters for Laboratory 49

Qualitative/quantitative	QC ranges	Laboratory 49						
		1085	1313	1570	1638	1858	1879	1953
Detected species	{Lm}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		98.3	98.4	98.7	98.1	98.5	98.5	98.8
% Species 2	{<5%}	0.0	0.0	0.0	0.2	0.0	0.0	0.0
Unclassified reads (%)		1.6	1.5	1.1	1.7	1.5	1.4	1.1
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	3.0	2.9	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	14	15	17	18	15	11	16
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	157	88	109	112	123	144	149
No. of reads (x 1000)		1689	951	1187	1203	1322	1476	1684
Average read length		281	282	281	283	282	285	269
Average insert size		390	382	385	390	389	424	322
N50 (kbp)		514	541	537	344	541	1500	541
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 19A. Calculated qualitative/quantitative parameters for Laboratory 56

Qualitative/quantitative	QC ranges	Laboratory 56						
		1094	1096	1230	1623	1655	1761	1821
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		97.8	97.8	97.9	97.4	97.5	98.2	98.1
% Species 2	{<5%}	1.1	1.4	1.2	1.3	1.1	1.0	1.1
Unclassified reads (%)		0.9	0.7	0.8	1.0	0.7	0.8	0.7
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	3.0	2.9	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	14	15	17	18	19	13	14
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	162	214	210	193	217	229	220
No. of reads (x 1000)		1919	2601	2494	2258	2609	2564	2581
Average read length		301	301	301	301	301	301	301
Average insert size		318	293	310	327	307	326	320
N50 (kbp)		541	541	477	344	555	1500	541
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 20A. Calculated qualitative/quantitative parameters for Laboratory 70

Qualitative/quantitative	QC ranges	Laboratory 70						
		1186	1388	1428	1461	1653	1685	1722
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		97.5	97.4	97.1	97.7	97.4	97.4	97.4
% Species 2	{<5%}	0.0	0.1	0.1	0.0	0.1	0.0	0.0
Unclassified reads (%)		2.4	2.4	2.7	2.3	2.4	2.5	2.5
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	2.9	3.0	3.0	3.0	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	16	19	26	19	25	19	21
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	80	100	46	85	78	72	77
No. of reads (x 1000)		1628	1956	945	1731	1601	1477	1556
Average read length		148	148	148	148	148	147	149
Average insert size		332	332	329	323	323	323	328
N50 (kbp)		541	360	344	476	348	476	477
QC-status (Bifrost)		OK	OK	Warning	OK	OK	OK	OK

QC: quality control; Warning: Average coverage is below 50 for one strain.

Table 21A. Calculated qualitative/quantitative parameters for Laboratory 88

Qualitative/quantitative	QC ranges	Laboratory 88						
		1038	1162	1177	1334	1522	1785	1875
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		95.3	95.6	95.6	95.2	94.8	95.6	95.2
% Species 2	{<5%}	0.1	0.0	0.0	0.1	0.0	0.0	0.1
Unclassified reads (%)		4.6	4.3	4.3	4.5	5.0	4.3	4.6
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	2.9	3.0	3.0	3.0	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	34	17	27	33	21	17	29
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	156	142	141	153	157	147	142
No. of reads (x 1000)		3254	2819	2918	3234	3310	3043	2981
Average read length		145	147	147	145	144	147	145
Average insert size		258	279	288	263	255	290	264
N50 (kbp)		278	1500	244	348	348	349	191
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 22A. Calculated qualitative/quantitative parameters for Laboratory 105

Qualitative/quantitative	QC ranges	Laboratory 105						
		1170	1320	1508	1569	1571	1877	1967
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		98.0	98.1	98.3	98.5	98.6	98.4	98.6
% Species 2	{<5%}	0.1	0.1	0.0	0.0	0.0	0.0	0.0
Unclassified reads (%)		1.9	1.7	1.6	1.4	1.4	1.5	1.4
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
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	48.3	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	24	24	14	14	20	16	20
No. of contigs [1-25] x min. coverage	{<1000}	0	0	2	0	0	0	0
Average coverage	{>50}	71	89	45	79	147	162	75
No. of reads (x 1000)		1443	1813	921	1555	2985	3293	1522
Average read length		149	149	149	147	148	148	149
Average insert size		407	394	365	270	282	287	341
N50 (kbp)		261	353	509	430	476	476	316
QC-status (Bifrost)		OK	OK	Warning	OK	OK	OK	OK

QC: quality control; Warning: Average coverage is below 50 for one strain.

Table 23A. Calculated qualitative/quantitative parameters for Laboratory 108

Qualitative/quantitative	QC ranges	Laboratory 108						
		1059	1192	1295	1423	1747	1848	1985
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		96.7	97.0	97.3	97.3	96.5	97.0	97.2
% Species 2	{<5%}	0.0	0.1	0.0	0.0	0.0	0.0	0.0
Unclassified reads (%)		3.2	2.8	2.7	2.7	3.5	3.0	2.8
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	2.9	3.0	2.9	3.0	3.0	2.9	2.9
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	13	21	14	19	16	14	17
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	47	47	36	55	47	48	57
No. of reads (x 1000)		950	983	738	1151	993	970	1164
Average read length		144	146	145	145	143	144	144
Average insert size		335	334	341	341	369	358	283
N50 (kbp)		1500	344	1490	509	478	1491	478
QC-status (Bifrost)		Warning	Warning	Warning	Warning	Warning	Warning	Warning

QC: quality control; Warning: Length at >25X coverage: below 2.8.
Some QC values are unreliable due to assembly issues for Ion Torrent data (contigs, average insert size, N50).

Table 24A. Calculated qualitative/quantitative parameters for Laboratory 129

Qualitative/quantitative	QC ranges	Laboratory 129						
		1027	1187	1294	1324	1335	1352	1964
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		95.8	95.9	91.2	95.8	96.7	94.4	96.8
% Species 2	{<5%}	0.1	0.0	0.1	0.0	0.0	0.0	0.0
Unclassified reads (%)		3.9	4.0	8.4	4.1	3.2	5.5	3.1
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	3.0	2.9	3.0
Length [1-25] x min. coverage (kbp)	{<250}	5.6	0.6	14.5	0.0	0.4	0.7	0.9
No. of contigs at 25 x min. coverage	{>0}	56	43	112	26	24	103	41
No. of contigs [1-25] x min. coverage	{<1000}	6	1	12	0	1	1	1
Average coverage	{>50}	72	167	59	153	155	157	149
No. of reads (x 1000)		1533	3615	1260	3333	3284	3378	3144
Average read length		144	139	145	139	143	134	143
Average insert size		390	289	322	307	344	263	331
N50 (kbp)		96	161	48	241	396	55	156
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 25A. Calculated qualitative/quantitative parameters for Laboratory 135

Qualitative/quantitative	QC ranges	Laboratory 135						
		1184	1249	1288	1290	1314	1498	1637
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		96.8	96.9	96.5	96.4	96.7	96.7	96.9
% Species 2	{<5%}	0.4	0.3	0.3	0.4	0.3	0.8	0.4
Unclassified reads (%)		2.6	2.7	2.8	3.0	2.9	2.4	2.5
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	2.9	3.0	3.0	3.0	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	17	14	25	16	17	17	24
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	336	534	300	276	422	491	196
No. of reads (x 1000)		6926	10566	6347	5709	8650	10071	4030
Average read length		151	151	151	151	151	151	151
Average insert size		353	347	342	367	359	347	351
N50 (kbp)		541	1500	349	514	476	541	344
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 26A. Calculated qualitative/quantitative parameters for Laboratory 138

Qualitative/quantitative	QC ranges	Laboratory 138						
		1066	1123	1154	1193	1366	1710	1832
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		97.5	97.5	97.9	97.7	97.8	98.0	97.9
% Species 2	{<5%}	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Unclassified reads (%)		2.3	2.3	2.0	2.2	2.1	1.9	2.0
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	2.9	2.9	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	22	19	16	16	14	18	22
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	65	233	224	61	76	164	130
No. of reads (x 1000)		1342	4801	4578	1253	1492	3301	2662
Average read length		148	148	148	148	148	148	148
Average insert size		354	400	354	343	342	346	355
N50 (kbp)		344	476	541	477	1500	477	350
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 27A. Calculated qualitative/quantitative parameters for Laboratory 141

Qualitative/quantitative	QC ranges	Laboratory 141						
		1204	1273	1348	1431	1514	1581	1837
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		96.9	96.6	96.7	97.0	96.8	97.2	97.0
% Species 2	{<5%}	0.0	0.1	0.0	0.1	0.0	0.0	0.0
Unclassified reads (%)		3.0	3.1	3.2	2.8	3.1	2.7	2.9
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	2.9	3.0	3.0	2.9	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	14	21	10	17	13	13	17
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	110	83	90	98	74	107	112
No. of reads (x 1000)		1211	913	964	1105	803	1154	1239
Average read length		279	283	277	273	283	278	276
Average insert size		393	401	409	381	422	390	396
N50 (kbp)		508	508	1500	344	541	514	541
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 28A. Calculated qualitative/quantitative parameters for Laboratory 142

Qualitative/quantitative	QC ranges	Laboratory 142						
		1019	1325	1395	1538	1568	1626	1773
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		98.1	97.9	97.9	98.0	97.8	97.9	98.3
% Species 2	{<5%}	0.2	0.1	0.2	0.1	0.1	0.1	0.1
Unclassified reads (%)		1.6	1.6	1.8	1.8	2.0	1.9	1.4
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	3.0	2.9	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.8
No. of contigs at 25 x min. coverage	{>0}	21	18	16	17	12	12	22
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	1
Average coverage	{>50}	60	58	50	60	58	61	57
No. of reads (x 1000)		742	724	623	749	698	760	697
Average read length		245	245	246	245	242	246	245
Average insert size		459	461	473	461	454	472	458
N50 (kbp)		344	499	542	444	1500	508	284
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 29A. Calculated qualitative/quantitative parameters for Laboratory 144

Qualitative/quantitative	QC ranges	Laboratory 144						
		1014	1276	1437	1484	1665	1896	1911
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		97.7	97.8	97.2	97.4	97.5	97.3	97.1
% Species 2	{<5%}	0.1	0.1	0.0	0.1	0.1	0.1	0.1
Unclassified reads (%)		2.1	2.1	2.6	2.4	2.2	2.6	2.8
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	2.9	3.0	3.0	3.0	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	10.3	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	30	16	14	26	27	16	23
No. of contigs [1-25] x min. coverage	{<1000}	6	0	0	0	0	0	0
Average coverage	{>50}	150	142	139	152	147	136	138
No. of reads (x 1000)		3071	2800	2850	3158	3024	2792	2845
Average read length		149	149	149	149	149	149	149
Average insert size		258	259	297	253	273	292	259
N50 (kbp)		478	430	541	538	244	349	316
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

QC: quality control.
All strains passed the QC.

Table 30A. Calculated qualitative/quantitative parameters for Laboratory 149

Qualitative/quantitative	QC ranges	Laboratory 149						
		1135	1291	1401	1418	1780	1904	1930
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		93.9	93.3	92.8	93.6	93.0	92.8	93.9
% Species 2	{<5%}	0.5	0.9	0.7	0.6	0.8	0.7	0.3
Unclassified reads (%)		5.3	5.7	6.2	5.5	5.8	6.1	5.4
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	2.9	3.0	3.0	3.0	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	18	17	26	17	43	36	29
No. of contigs [1-25] x min. coverage	{<1000}	0	0	0	0	0	0	0
Average coverage	{>50}	155	158	136	147	168	131	117
No. of reads (x 1000)		3274	3214	2881	3094	3602	2742	2416
Average read length		151	151	151	151	151	151	151
Average insert size		270	288	273	277	265	307	308
N50 (kbp)		508	437	344	508	244	253	477
QC-status (Bifrost)		OK	OK	OK	OK	OK	OK	OK

All strains passed the QC.

Table 31A. Calculated qualitative/quantitative parameters for Laboratory 177

Qualitative/quantitative	QC ranges	Laboratory 177						
		1429	1527	1541	1603	1826	1833	1890
Detected species	{*Lm*}	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		91.4	91.5	88.3	90.2	89.4	90.2	89.1
% Species 2	{<5%}	0.0	0.1	0.2	0.0	0.0	0.0	0.1
Unclassified reads (%)		8.5	8.4	11.4	9.7	10.5	9.7	10.6
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	3.0	3.0	3.0	1.5	3.0	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	9.0	12.2	21.1	1414.5	8.9	2.7	36.1
No. of contigs at 25 x min. coverage	{>0}	89	74	264	68	159	127	271
No. of contigs [1-25] x min. coverage	{<1000}	11	10	29	100	8	2	37
Average coverage	{>50}	39	32	41	27	42	52	38
No. of reads (x 1000)		981	790	1102	657	1120	1341	1008
Average read length		120	121	114	118	113	117	115
Average insert size		182	189	158	180	150	169	163
N50 (kbp)		57	75	20	28	40	47	18
QC-status (Bifrost)		Warning	Warning	Warning	Warning	Warning	OK	Warning

Warning: Average coverage is below 50 for all strains except strain 1826, which passed the QC.

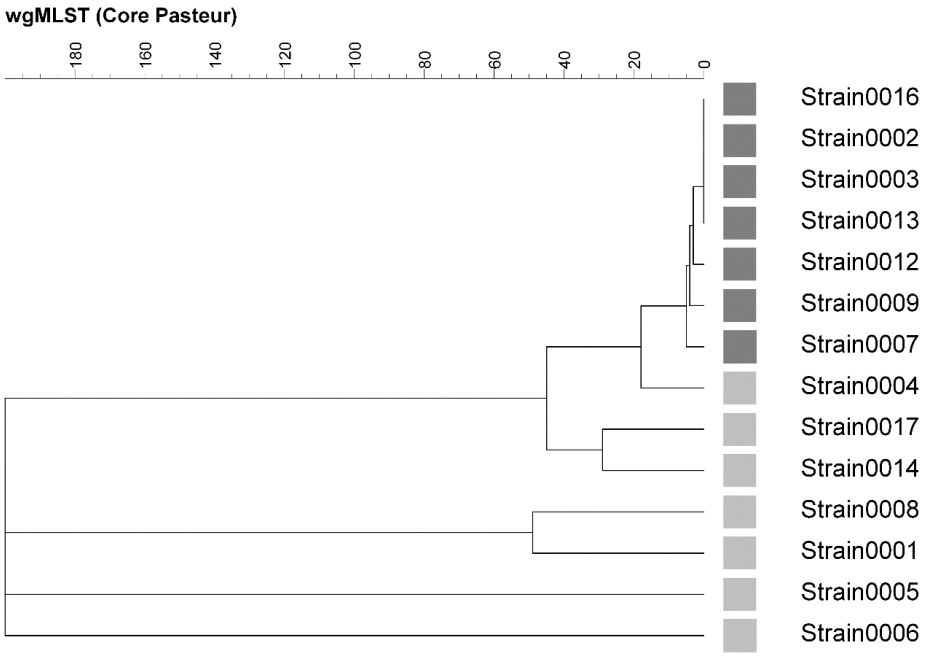
Table 32A. Calculated qualitative/quantitative parameters for Laboratory 187

Qualitative/quantitative	QC ranges	Laboratory 187						
		1190	1326	1362	1607	1730	1751	1983
Detected species	{*Lm*}	*Lm*	*Lm,*	*Lm*	*Lm*	*Lm*	*Lm*	*Lm*
% Species 1		97.5	97.0	97.2	97.6	97.1	97.6	97.5
% Species 2	{<5%}	0.0	0.1	0.0	0.1	0.0	0.0	0.0
Unclassified reads (%)		2.5	2.7	2.7	2.2	2.9	2.3	2.4
Length at >25 x min. coverage (Mbp)	{>2.8 ∧ <3.1}	2.9	3.0	2.6	3.0	2.9	3.0	3.0
Length [1-25] x min. coverage (kbp)	{<250}	0.0	0.0	351.6	0.0	47.6	0.0	0.0
No. of contigs at 25 x min. coverage	{>0}	16	27	23	23	22	17	17
No. of contigs [1-25] x min. coverage	{<1000}	0	0	3	0	3	0	0
Average coverage	{>50}	54	43	26	90	27	60	64
No. of reads (x 1000)		1059	891	545	1846	566	1234	1310
Average read length		147	147	147	147	146	147	147
Average insert size		377	388	396	328	398	371	336
N50 (kbp)		1500	310	316	344	478	518	476
QC-status (Bifrost)		OK	Warning	Warning	OK	Warning	OK	OK

Warning: Average coverage is below 50 for some strains.
Quality assessment made using the EQA provider's in-house quality control pipeline.

Annex 4. EQA provider cluster analysis

Figure 1A. EQA provider cluster analysis, based on WGS-derived data



Single linked dendrogram of core genome multilocus sequence typing (cgMLST) profiles of Listeria *EQA-10 strains (cgMLST, Pasteur, Moura et al., 2016).*
Analysed in BioNumerics: maximum distance of 200 exceeded; results clipped.
Dark grey: cluster strains.
Light grey: outside cluster strains.

Annex 5. EQA-10 laboratory questionnaire

Listeria EQA-10 2023–2024

Dear participant,

Welcome to the Tenth External Quality Assessment (EQA-10) scheme for typing of Listeria in 2023–2024.

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 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:

- Fill in your email to receive a link with your answers. The email with the link will be sent after pressing "Finish" in the last slide in the survey.
- If the survey is shut down before you are finished, the answers are saved and it is possible to return to the survey through the same link.

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

1. Country

(State one answer only)

☐ Australia
☐ Austria
☐ Belgium
☐ Croatia
☐ Denmark
☐ Finland
☐ France
☐ Germany
☐ Greece
☐ Hungary
☐ Iceland
☐ Ireland
☐ Italy
☐ Israel
☐ Latvia
☐ Lithuania
☐ Luxembourg
☐ Malta
☐ New Zealand
☐ Norway
☐ Portugal
☐ Scotland
☐ Slovakia
☐ Slovenija
☐ Spain
☐ Sweden
☐ The Netherlands
☐ Turkey
☐ United Kingdom
☐ United States

2. Institute name

3. Laboratory name

4. Laboratory ID

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

5. Email

6. Listeria EQA-10 Strain IDs

Please enter the isolate ID (4 digits)

Strain number

Strain 1 ___
Strain 2 ___
Strain 3 ___
Strain 4 ___
Strain 5 ___
Strain 6 ___
Strain 7 ___

Serotyping/grouping of *Listeria*

7. Would you like to submit serotyping/grouping results?

(State one answer only)

☐ Yes
☐ Did not participate in the serotyping/grouping part – Go to 12

8. Submitting results - serotyping/grouping of *Listeria*

(State one answer only)

☐ Both molecular and conventional serogrouping/serotyping – Go to 9
☐ Molecular serogrouping – Go to 9
☐ Conventional serotyping – Go to 11

9. Method used for molecular serogrouping of *Listeria*

(State one answer only)

☐ PCR-based
☐ WGS-based

10. Results for serotyping/grouping *Listeria* – molecular serogrouping

Please select the serogroup

(State only one answer per question)

Strain	Molecular serotype					
	IIa	IIb	IIc	IVb	L	Un-typeable
Strain1						
Strain2						
Strain3						
Strain4						
Strain5						
Strain6						
Strain7						

11. Results for serotyping *Listeria* – Conventional serotyping

Please select the serotype

(State only one answer per question)

Strain	Conventional serotype														
Strain1	1/2a	1/2b	1/2c	3a	3b	3c	4a	4ab	4b	4c	4d	4e	7	Autoagglutinable	Un-typeable
Strain2															
Strain3															
Strain4															
Strain5															
Strain6															
Strain7															

12. Submitting cluster results

(State one answer only)

☐ Cluster analyses based on PFGE and/or WGS
☐ Did not participate in the Cluster part – Go to 213.

13. Submitting cluster results

(State one answer only)

☐ Cluster analysis based on PFGE – Go to 14
☐ Do not wish to submit any cluster results based on PFGE analysis – Go to 17

Cluster analysis based on PFGE data

14. 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 a semicolon (;) to separate the IDs.

15. *ApaI* – Total number of bands (>33kb) in a cluster strain

(Use 9999 if not analysed)

16. *AscI* – Total number of bands (>33kb) in a cluster strain

(Use 9999 if not analysed)

17. Submitting cluster results

(State one answer only)

☐ Cluster analysis based on WGS data - Go to 18
☐ Do not wish to submit any cluster results based on WGS data - Go to 213.

Cluster analysis based on WGS data

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

The results of the cluster detection can only be reported once (main analysis). If more than one analysis is performed, please report later in this submission.

(State one answer only)

☐ SNP-based – Go to 20
☐ Allele-based – Go to 27
☐ Other – Go to 19

19. If another analysis is used, please describe in detail your approach (including: assembler, number of loci, variant caller, read mapper or reference ID etc.)

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

21. Please select the approach used for the SNP analysis

(State one answer only)

☐ Reference based – Go to 22
☐ Assembly based – Go to 25

22. Reference genome used

Preferable use EQA strain 0016 (downloaded sequences) as reference. Otherwise indicate multilocus sequence type (e.g. ST8) and identification of the reference used.

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

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

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

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

27. Please select tools used for the allele analysis

(State one answer only)

☐ BioNumerics – Go to 29
☐ SeqSphere – Go to 29
☐ BIGSdb-*Lm* – Go to 29
☐ Other – Go to 28

28. If another tool is used please enter here

29. Please indicate allele calling method

(State one answer only)

☐ Assembly based and mapping based – Go to 30
☐ Only assembly based – Go to 30
☐ Only mapping based – Go to 31

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

31. Please select scheme used for the allele analysis

(State one answer only)

☐ Applied Math (wgMLST) – Go to 33
☐ Applied Math (cgMLST/Pasteur) – Go to 33
☐ Pasteur (cgMLST) – Go to 33
☐ Ruppitsch (cgMLST) – Go to 33
☐ Other – Go to 32

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

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

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 IDs for isolates in the cluster detected with the additional analysis.

Please fill in all the data for the strains one by one.

34. Strain 1

Report the MLST, serotype/group, part of the cluster and SNP distance/allele difference

35. (Optional) Report the serotype/group

36. Report the 7-gene MLST

(State value between 0 and 1000000)

37. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

38. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

(Please use 9999 for not analysed)

39. Strain 2

Report the MLST, serotype/group, part of the cluster and SNP distance /allele difference

40. (Optional) Report the serotype/group

41. Report the 7-gene MLST

(State value between 0 and 1000000)

42. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

43. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

(Please use 9999 for not analysed)

44. Strain 3

Report the MLST, serotype/group, part of the cluster and SNP distance /allele difference

45. (Optional) Report the serotype/group,

46. Report the 7-gene MLST

(State value between 0 and 1000000)

47. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No.

48. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

(Please use 9999 for not analysed)

49. Strain 4

Report the MLST, serotype/group, part of the cluster and SNP distance/allele difference

50. (Optional) Report the serotype/group

51. Report the 7-gene MLST

(State value between 0 and 1000000)

52. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

53. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

(Please use 9999 for not analysed)

54. Strain 5

Report the MLST, serotype/group, part of the cluster and SNP distance /allele difference

55. (Optional) Report the serotype/group

56. Report the 7-gene MLST

(State value between 0 and 1000000)

57. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

58. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

(Please use 9999 for not analysed)

59. Strain 6

Report the MLST, serotype/group, part of the cluster and SNP distance/allele difference.

60. (Optional) Report the serotype/group

61. Report the 7-gene MLST

(State value between 0 and 1000000)

62. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

63. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

(Please use 9999 for not analysed)

64. Strain 7

Report the MLST, serotype/group, part of the cluster and SNP distance/allele difference.

65. (Optional) Report the serotype/group

66. Report the 7-gene MLST

(State value between 0 and 1000000)

67. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

68. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

(Please use 9999 for not analysed)

69. Strain 0008 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

70. QC observations

Please evaluate the QC results of the strain and explain what you observed.

71. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Unacceptable quality – strain not analysed – Go to 77

72. Strain 0008 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance allele difference.

73. (Optional) Report the serotype/group

74. Report the 7-gene MLST

(State value between 0 and 1000000)

75. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

76. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed.

(State value)

77. Strain 0009 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

78. QC observations

Please evaluate the QC results of the strain and explain what you observe

79. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Unacceptable quality – strain not analysed – Go to 85

80. Strain 0009 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

81. (Optional) Report the serotype/group

82. Report the 7-gene MLST

(State value between 0 and 1000000)

83. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

84. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed.

(State value)

85. Strain 0010 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

86. QC observations

Please evaluate the QC results of the strain and explain what you observed.

87. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Unacceptable quality – strain not analysed – Go to 93

88. Strain 0010 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

89. (Optional) Report the serotype/group

90. Report the 7-gene MLST

(State value between 0 and 1000000)

91. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

92. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

93. Strain 0011 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

94. QC observations

Please evaluate the QC results of the strain and explain what you observe

95. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Unacceptable quality – strain not analysed – Go to 101

96. Strain 0011 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

97. (Optional) Report the serotype/group

98. Report the 7-gene MLST

(State value between 0 and 1000000)

99. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

100. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

101. Strain 0012 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

102. QC observations

Please evaluate the QC results of the strain and explain what you observed.

103. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Unacceptable quality – strain not analysed – Go to 109

104. Strain 0012 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

105. (Optional) Report the serotype/group

106. Report the 7-gene MLST

(State value between 0 and 1000000)

107. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

108. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

109. Strain 0013 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

110. QC observations

Please evaluate the QC results of the strain and explain what you observe

111. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Unacceptable quality – strain not analysed – Go to 117

112. Strain 0013 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance/allele difference.

113. (Optional) Report the serotype/group

114. Report the 7-gene MLST

(State value between 0 and 1000000)

115. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

116. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

117. Strain 0014 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

118. QC observations

Please evaluate the QC results of the strain and explain what you observe

119. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Not acceptable quality - strain not analysed - Go to 125

120. Strain 0014 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

121. (Optional) Report the serotype/group

122. Report the 7-gene MLST

(State value between 0 and 1000000)

123. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

124. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

125. Strain 0015 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

126. QC observations

Please evaluate the QC results of the strain and explain what you observe

127. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Not acceptable quality - strain not analysed - Go to 133

128. Strain 0015 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

129. (Optional) Report the serotype/group

130. Report the 7-gene MLST

(State value between 0 and 1000000)

131. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

132. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

133. Strain 0016 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

134. QC observations

Please evaluate the QC results of the strain and explain what you observe

135. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Not acceptable quality – strain not analysed – Go to 141

136. Strain 0016 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

137. (Optional) Report the serotype/group

138. Report the 7-gene MLST

(State value between 0 and 1000000)

139. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

140. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

141. Strain 0017 (as downloaded sequence)

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

142. QC observations

Please evaluate the QC results of the strain and explain what you observe

143. Please select the QC status that fits with your assessment of the strain

(State one answer only)

☐ Acceptable quality
☐ Quality only acceptable for outbreak situations (less good quality)
☐ Not acceptable quality – strain not analysed – Go to 149

144. Strain 0017 continue

Report the QC status, MLST, serotype/group, part of the cluster and SNP distance /allele difference

145. (Optional) Report the serotype/group

146. Report the 7-gene MLST

(State value between 0 and 1000000)

147. Report if this strain is a part of identified cluster

(State one answer only)

☐ Yes
☐ No

148. Report the allele difference/SNP distance to the strain 16 (as 0016 downloaded sequence)

Please use 9999 for not analysed

(State value)

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

For example, 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 150
☐ No – Go to 187

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

(State one answer only)

☐ SNP-based – Go to 152
☐ Allele-based – Go to 159
☐ Other – Go to 151

151. If another analysis is used please describe in detail your approach (including: assembler, number of loci, variant caller, read mapper or reference ID etc.) – Go to 166

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

153. Please select the approach used for the SNP analysis

(State one answer only)

☐ Reference based – Go to 154
☐ Assembly based – Go to 157

154. Reference genome used: (preferable use EQA strain 0016, downloaded sequences as reference)

Otherwise indicate Multi-locus Sequence Type (e.g. ST8) and isolate ID

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

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

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

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

159. Please select tool used for the allele analysis

(State one answer only)

☐ BioNumerics – Go to 161
☐ SeqSphere – Go to 161
☐ BIGSdb-*Lm* – Go to 161
☐ Other – Go to 160

160. If another tool is used please list here

161. Please indicate allele calling method

(State one answer only)

☐ Assembly based and mapping based – Go to 162
☐ Only assembly based – Go to 162
☐ Only mapping based – Go to 163

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

163. Please select scheme used for the allele analysis

(State one answer only)

☐ Applied Maths (wgMLST) – Go to 165
☐ Applied Maths (cgMLST/Pasteur) – Go to 165
☐ Pasteur (cgMLST) – Go to 165
☐ Ruppitsch (cgMLST) – Go to 165
☐ Other – Go to 164

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

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

166. Additional analysis on data derived from WGS.

167. Results for an additional cluster analysis.

Reporting AD/SNP distances to strain 0016 (as downloaded sequence) (e.g. SNP or allele-based)

Please use 9999 for not analysed.

Isolate	Distance/difference (e.g. SNP/allele) to the strain 0009 (downloaded sequence)
Strain 1	
Strain 2	
Strain 3	
Strain 4	
Strain 5	
Strain 6	
Strain 7	
Strain 0008 (as downloaded sequence)	
Strain 0009 (as downloaded sequence)	
Strain 0010 (as downloaded sequence)	
Strain 0011 (as downloaded sequence)	
Strain 0012 (as downloaded sequence)	
Strain 0013 (as downloaded sequence)	
Strain 0014 (as downloaded sequence)	
Strain 0015 (as downloaded sequence)	
Strain 0016 (as downloaded sequence)	
Strain 0017 (as downloaded sequence)	

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

For example, 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 169
☐ No – Go to 187

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

(State one answer only)

☐ SNP-based – Go to 171
☐ Allele-based – Go to 178
☐ Other – Go to 170

170. If another analysis is used, please describe in detail your approach (including: assembler, number of loci, variant caller, read mapper or reference ID etc.) – Go to 185

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

172. Please select the approach used for the SNP analysis

(State one answer only)

☐ Reference-based – Go to 173
☐ Assembly-based – Go to 176

173. Reference genome used: (preferable use EQA strain 0009, downloaded sequences as reference)

Otherwise indicate multilocus sequence type (e.g. ST8) and isolate ID.

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

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

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

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

178. Please select tool used for the allele analysis

(State one answer only)

☐ BioNumerics – Go to 180
☐ SeqSphere – Go to 180
☐ BIGSdb-*Lm* – Go to 180
☐ Other – Go to 179

179. If another tool is used please enter here

180. Please indicate allele calling method

(State one answer only)

☐ Assembly-based and mapping-based – Go to 181
☐ Only assembly-based – Go to 181
☐ Only mapping-based – Go to 182

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

182. Please select scheme used for the allele analysis

(State one answer only)

☐ Applied Maths (wgMLST) – Go to 184
☐ Applied Maths (cgMLST/Pasteur) – Go to 184
☐ Pasteur (cgMLST) – Go to 184
☐ Ruppitsch (cgMLST) – Go to 184
☐ Other – Go to 183

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

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

185. Third analysis on data derived from WGS

186. Results for the third cluster analysis

Reporting AD/SNP distances to strain 0016 (as downloaded sequence) (e.g. SNP or allele-based)

Please use 9999 for not analysed.

Isolate	Distance/difference (e.g. SNP/allele) to the strain 0009 (downloaded sequence)
Strain 1	
Strain 2	
Strain 3	
Strain 4	
Strain 5	
Strain 6	
Strain 7	
Strain 0008 (as downloaded sequence)	
Strain 0009 (as downloaded sequence)	
Strain 0010 (as downloaded sequence)	
Strain 0011 (as downloaded sequence)	
Strain 0012 (as downloaded sequence)	
Strain 0013 (as downloaded sequence)	
Strain 0014 (as downloaded sequence)	
Strain 0015 (as downloaded sequence)	
Strain 0016 (as downloaded sequence)	
Strain 0017 (as downloaded sequence)	

187. Additional questions for the WGS part

188. Where was the sequencing performed?

(State one answer only)

☐ In own laboratory
☐ Externally

189. Protocol used to prepare the library for sequencing

(State one answer only)

☐ Commercial kits – Go to 190
☐ Non-commercial kits – Go to 192

190. Please indicate name of commercial kit

191. If relevant, please list deviation from commercial kit briefly in a few bullet points – Go to 198

192. For non-commercial kit please indicate a short summary of the protocol

193. The sequencing platform used

(State one answer only)

☐ Ion Torrent PGM – Go to 195
☐ Ion Torrent Proton – Go to 195
☐ Genome Sequencer Junior System (454) – Go to 195
☐ Genome Sequencer FLX System (454) – Go to 195
☐ Genome Sequencer FLX+ System (454) – Go to 195
☐ PacBio RS – Go to 195
☐ PacBio RS II – Go to 195
☐ HiScanSQ – Go to 195
☐ HiSeq 1000 – Go to 195
☐ HiSeq 1500 – Go to 195
☐ HiSeq 2000 – Go to 195
☐ HiSeq 2500 – Go to 195
☐ HiSeq 4000 – Go to 195
☐ Genome Analyzer lix – Go to 195
☐ MiSeq – Go to 195
☐ MiSeq Dx – Go to 195
☐ MiSeq FGx – Go to 195
☐ ABI SOLiD – Go to 195
☐ NextSeq – Go to 195
☐ MinION (ONT) – Go to 195
☐ Other – Go to 194

194. If another platform is used please list here

195. 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 five selected criteria which were the most frequently reported by in previous EQAs.

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

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

196. Did you use confirmation of species to evaluate the quality of sequence data?

☐ Yes
☐ No – Go to 198

197. Procedure used to evaluate confirmation of genus

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

(State one answer only)

☐ Yes
☐ No – Go to 200

199. Procedure or threshold used for coverage

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

(State one answer only)

☐ Yes
☐ No – Go to 202

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

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

(State one answer only)

☐ Yes
☐ No – Go to 203

202. Procedure or threshold used for genome size:

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

(State one answer only)

☐ Yes
☐ No – Go to 205

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

205. ONLY list additional information related to other criteria used to evaluate the quality of sequence data

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

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

207. Threshold or procedure used to evaluate the additional criteria 1

208. Other criteria used to evaluate the quality of sequence data – additional criteria 2

209. Threshold or procedure used to evaluate the additional criteria 2

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

211. Threshold or procedure used to evaluate the additional criteria 3

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

213. Threshold or procedure used to evaluate the additional criteria 4

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

215. Threshold or procedure used to evaluate the additional criteria 5

216. Comment(s)

For example, remarks on the submission, the data analyses or the laboratory methods.

Thank you for your participation

Thank you for filling out the submission form for the *Listeria* EQA-10.

For questions, please email: list.eqa@ssi.dk or telephone: +45 3268 8341.

We highly recommend documenting 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
ECDC.info@ecdc.europa.eu

www.ecdc.europa.eu

Follow ECDC on social media
Twitter: @ECDC_EU
Facebook: www.facebook.com/ECDC.EU
Linkedin: www.linkedin.com/company/ecdc/

Publications Office
of the European Union