ecdc
EUROPEAN CENTRE FOR DISEASE PREVENTION AND CONTROL

TECHNICAL REPORT

Ninth external quality assessment scheme for typing of Shiga toxin-producing *Escherichia coli*

www.ecdc.europa.eu

ECDC TECHNICAL REPORT

Ninth external quality assessment scheme for typing of Shiga toxin-producing *Escherichia coli*

This report was commissioned by the European Centre for Disease Prevention and Control (ECDC), coordinated by Taina Niskanen (ECDC) and produced by Susanne Schjørring, Gitte Sørensen, Kristoffer Kiil, Louise Gade Dahl, Flemming Scheutz and Eva Møller Nielsen of the Foodborne Infections Unit at Statens Serum Institut, Copenhagen, Denmark.

Suggested citation: European Centre for Disease Prevention and Control. Ninth external quality assessment scheme for typing of Shiga toxin-producing *Escherichia coli.* Stockholm: ECDC; 2020.

Stockholm, November 2020

ISBN 978-92-9498-524-8
DOI 10.2900/14830
Catalogue number TQ-02-20-918-EN-N

Contents

Figures

Tables

Abbreviations

aaiC	Chromosomal gene marker for enteroaggregative *E. coli*
AEEC	Attaching and effacing *E. coli*
aggR	Gene encoding the master regulator in enteroaggregative *E. coli*
BN	BioNumerics
bp	Base pairs
cgMLST	Core genome Multilocus Sequence Typing
eae	The intimin gene
EAEC	Enteroaggregative *E. coli*
EFSA	European Food Safety Authority
EQA	External quality assessment
EURL	European Union Reference Laboratory
FWD	Food- and waterborne diseases
FWD-Net	Food- and Waterborne Diseases and Zoonoses Network
GFN	Global Foodborne Infections Network, Food Safety
HUS	Haemolytic uraemic syndrome
ND	Not done
NPHRL	National Public Health Reference Laboratories
NSF	Non-sorbitol fermenter
NT	Non-typeable
PFGE	Pulsed-field gel electrophoresis
QC	Qualitative control
SNP	Single nucleotide polymorphism
SOP	Standard operating procedure
SSI	Statens Serum Institut
ST	Sequence type
STEC	Shiga toxin-producing *E. coli* (synonymous with verocytotoxin-producing *E. coli;* VTEC)
Stx1	Shiga toxin 1
stx1	Gene encoding Shiga toxin 1
Stx2	Shiga toxin 2
stx2	Gene encoding Shiga toxin 2
TESSy	The European Surveillance System
wgMLST	Whole genome Multilocus Sequence Typing
WGS	Whole genome sequencing

Executive summary

This report presents the results of the ninth round of the external quality assessment (EQA-9) scheme for typing of Shiga toxin-producing *Escherichia coli* (STEC). This EQA was organised for National Public Health Reference Laboratories (NPHRLs) to provide data to the Food- and Waterborne Diseases and Zoonoses Network (FWD-Net) managed by the European Centre for Disease Prevention and Control (ECDC). Since 2012, the unit of Foodborne Infections at Statens Serum Institut (SSI) in Denmark has arranged this EQA under a framework contract with ECDC. EQA-9 contains serotyping, detection of virulence genes and molecular typing-based cluster analysis.

Human STEC infection is a zoonotic disease. In 2018, the disease had an EU notification rate of 2.3 cases per 100 000 population which represented an increase after a stable period from 2014 to 2017. The most commonly reported STEC O group was O157 (34.5% of cases with known serogroup).

Since 2007, ECDC has been responsible for EU-wide surveillance of STEC, including facilitating, detecting and investigating food-borne outbreaks. The surveillance system relies on the capacity of NPHRLs in FWD-Net to produce and report comparable typing results to The European Surveillance System (TESSy). In order to ensure the EQA is linked to the development of surveillance methods used by NPHRLs, EQA-8 as well as EQA-9 contains a molecular typing-based cluster analysis using either pulsed-field gel electrophoresis (PFGE) and/or whole genome sequencing (WGS)-derived data, while the quality assessment of PFGE as performed in EQA-4 to-7 has been excluded.

The objectives of the EQA are to assess the quality and comparability of typing data reported by NPHRLs participating in FWD-Net. Test isolates for the EQA were selected to cover strains currently relevant to public health in Europe and represent a broad range of clinically relevant types for STEC. Two separate sets of 12 test isolates were selected for serotyping/virulence profile determination and molecular typing-based cluster analysis.

Twenty-six laboratories registered and 24 completed the exercise, representing a decrease in participation of 20% from EQA-7. This decrease in the number of participants may have been caused by adding molecular typing-based cluster analysis (using PFGE and/or WGS without a standard protocol) or removing two independent analytical steps that were covered in previous assessments: quality assessments of PFGE and phenotypic analysis.

The full O:H serotyping was performed by 54% (13/24) of participating laboratories, with an average score of 92%. In general, the more common European serotypes generated the highest scores, e.g. 100% for O157:H7 isolates, while the less frequent O187:H28 obtained an average score of 85%. Notably, not all laboratories demonstrated the capacity to determine all O groups and H types and participation in H typing was low (13/24). A shift towards the WGS-based method was observed, 50% (10/20) of the participants used WGS-based serotyping in the EQA-9 compared to 26% (6/23) in the EQA-8.

The quality of the virulence profile determination results was generally good, with high average scores for *eae* (99%), *stx1* (100%) and *stx2* (99%), similar to previous EQAs. The participants identified the enteroaggregative *E. coli* (EAEC) isolate by correctly reporting the presence of the *aaiC* and/or *aggR.* The average scores were 89% for *aaiC* and 95% for *aggR*. Subtyping of *stx1* and *stx2* obtained a combined average score of 89%, which is slightly below the range (90-92%) of the previous EQA results when excluding the results of EQA-8 (77%).

Out of the 24 laboratories participating in EQA-9, 17 (71%) performed molecular typing-based cluster analysis using any method. The aim of the cluster analysis part of the EQA was to assess the NPHRL's ability to identify a cluster of genetically closely related isolates given that a multitude of different laboratory methods and analytical methods are used as the primary cluster detection approach in Member States. The aim of this part of the EQA was to assess the participants' ability to reach the correct conclusion, i.e. to correctly categorise the cluster test isolates, instead of the ability to follow a specific procedure.

The cluster of closely related isolates contained five ST17 isolates that could be identified by both PFGE and WGS-derived data. The expected cluster was based on a predefined categorisation by the organiser. Notably, just below half the laboratories (8/17) used PFGE for cluster analysis and three also reported cluster analysis based on WGS data.

Twelve laboratories performed cluster analysis using WGS-derived data. Performance was high, with 11 (92%) participants correctly identifying the cluster of closely related isolates using WGS. In this EQA, participants were free to choose their preferred analytical method for the WGS-based cluster identification. An allele-based method was preferred, since 83% (10/12) used core genome Multi Locus Sequence Type (cgMLST) compared to 17% (2/12) using single nucleotide polymorphism (SNP) for the reported cluster analysis.

Allele- and SNP-based methods seemed equally suitable for cluster identification. In general, for cgMLST, the reported results were at a comparable level despite analysis with different approaches and allelic calling methods. Only two laboratories reported SNP results, however they showed more variability primarily outside the cluster. For inter-laboratory comparability and communication about cluster definitions, cgMLST using a standard scheme (e.g. Enterobase) gives a very high degree of homogeneity in the results, while the use of non-standardised SNP analysis may be more challenging for comparison and communication between laboratories. This issue is further complicated as many laboratories still use PFGE and will probably not switch to WGS in the near future. In this EQA, 39% (5/17) of participants in cluster analysis only used PFGE and three did not identify the correct cluster.

1. Introduction

1.1 Background

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

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

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

The main purposes of EQA schemes are to:

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

Since 2012, the unit of Foodborne Infections at Statens Serum Institut (SSI), Denmark has been the EQA provider for the three lots covering typing of *Salmonella enterica* ssp. *enterica*, Shiga toxin/verocytotoxin-producing *Escherichia coli (E. coli;* STEC/VTEC) and *L. monocytogenes*. In 2016, SSI was granted the new round of tenders (2017–2020) for all three lots. For lot 2 (STEC) as in EQA-8, the EQA scheme no longer covers assessment of PFGE quality. However, it still covers serotyping, virulence profile determination and molecular typing-based cluster analysis. This report presents the results of the ninth EQA scheme (STEC EQA-9).

1.2 Surveillance of STEC infections

STEC is a group of *E. coli* characterised by the ability to produce Shiga toxins (Stxs).

Human pathogenic STEC often harbour additional virulence factors important to the development of the disease. A large number of serotypes of *E. coli* have been recognised as Stx producers. Notably, the majority of reported human STEC infections are sporadic cases. Symptoms associated with STEC infection in humans vary from mild diarrhoea to life-threatening haemolytic uraemic syndrome (HUS), which is clinically defined by a combination of haemolytic anaemia, thrombocytopenia and acute renal failure.

In 2018, the EU notification rate of STEC infections was 2.3 cases per 100 000 population. The total number of confirmed STEC infection cases was 8 161, an increase of 37% from 2017 (n=5 958). Eleven deaths due to STEC infection were reported, resulting in an EU case fatality of 0.2% among cases with a reported outcome. As in previous years, the most commonly reported STEC O group was O157 (34.5% of cases with known serogroup). O group O157 was followed by O26 [3].

One of the key objectives for ECDC was to improve and harmonise the surveillance system in the EU to increase scientific knowledge of aetiology, risk factors and burden of FWDs and zoonoses. Surveillance data, including some basic typing parameters for the isolated pathogen, are reported by Member States to TESSy. In addition to the basic characterisation of the pathogens isolated from infections, there is public health value to use more discriminatory typing techniques in the surveillance of food-borne infections. Since 2012, ECDC has enhanced EU surveillance by incorporating molecular typing data ('molecular surveillance') through isolate-based reporting. Three selected FWD pathogens were included: *Salmonella enterica* ssp. *enterica, L. monocytogenes* and STEC/VTEC. The overall aims of integrating molecular typing into EU level surveillance are to:

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

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

EQA schemes have targeted NPHRLs performing molecular typing-enhanced surveillance or those who implement it to their surveillance at the national level.

1.3 STEC characterisation

State-of-the-art characterisation of STEC includes O:H serotyping in combination with a few selected virulence genes, i.e. the two genes for production of Shiga toxin Stx1 (*stx1*) and Stx2 (*stx2*) and the intimin *(eae)* gene associated with attaching and effacing lesion of enterocytes, also seen in attaching and effacing non-STEC *E. coli* (AEEC), including enteropathogenic *E. coli* (EPEC). The combination of virulence genes and subtypes of toxin genes is clinically relevant. The *stx2a* in *eae*-positive STEC and the activable *stx2d* subtype in *eae*-negative STEC appear to be highly associated with the serious sequela HUS [4–6]. Other specific subtypes of Stx1 and Stx2 are primarily associated with milder course of disease without HUS [4–6].

Understanding the epidemiology of *stx* subtypes is therefore important to reduce the risk of STEC infection and for the surveillance of STEC.

The recommended method for *stx* subtyping is a specific PCR [7]. STEC serotype O157:[H7] may be divided into two groups: non-sorbitol fermenters (NSF) and a highly virulent variant of O157-fermenting sorbitol (SF). STEC EQA-9 included O:H serotyping, detection of virulence genes (*eae*, *stx1* and *stx2,* including subtyping of *stx* genes) and genes for enteroaggregative *E. coli* (EAEC) and molecular typing-based cluster analysis.

1.4 Objectives

1.4.1 Serotyping

The objectives of STEC EQA-9 were to assess the ability to assign correct O groups and H types by using either serological (somatic 'O' and flagellar 'H' antigens) or molecular typing methods (PCR or WGS).

1.4.2 Virulence profile determination

The objectives of the virulence gene determination of STEC EQA-9 were to assess the ability to assign the correct virulence profile.

The presence/absence of *stx1*, *stx2*, *eae, aaiC* and *aggR* genes and subtyping of stx genes (*stx1a, stx1c* and *stx1d* and *stx2a* to *stx2g*).

1.4.3 Molecular typing-based cluster analysis

The objective of the molecular typing-based cluster analysis of STEC EQA-9 was to assess the ability to detect a cluster of closely related isolates. Laboratories could perform the analyses using PFGE or derived data from WGS.

2. Study design

2.1 Organisation

STEC EQA-9 was funded by ECDC and arranged by SSI following ISO/IEC 17043:2010 [8]. EQA-9 included serotyping, virulence gene determination and a molecular typing-based cluster analysis and was carried out between November 2018 and April 2019.

Invitations were emailed to ECDC contact points in FWD-Net (27 countries, which nominated laboratories to participate in the EQA rounds 2017-2020) by 12 December 2018 with a deadline to respond by 08 January 2019. In addition, invitations were sent to EU candidate and potential candidate countries Turkey and Republic of Macedonia, which signed to the STEC EQA rounds in 2017–2020. Each laboratory was asked to fill in the reason for participating or non-participating.

Twenty-six NPHRLs in EU/EEA and EU candidate countries accepted the invitation to participate and 24 submitted results (Annex 1). EQA test isolates were sent to participants on 5 February 2019. In Annex 2, participation details in EQA-8 and EQA-9 are listed to give an overview of the trend in the number of participants. In addition Annex 3 contains each laboratory's reason for participating or not participating. Fourteen self-funded laboratories were invited to participate in the EQA. Participants were asked to submit their results to an SFTP-site and complete the online form by 15 April 2019 (Annex 15). Two laboratories were asked to submit the missing raw reads, and three laboratories were asked to re-submit a few sequences as some had been interrupted during submission.

The EQA submission protocol, invitation letter and an empty submission form were available on the online site.

2.2 Selection of test isolates

Thirty-five test isolates were selected to fulfil the following criteria:

- represent commonly reported strains in Europe
- remain stable during the preliminary test period at the organising laboratory
- include genetically closely related isolates.

The 35 selected isolates were analysed using the methods used in the EQA before and after having been recultured 10 times. All candidate isolates remained stable using these methods and the final test isolates were selected. The 12 test isolates (Table 1) for serotyping/detection of virulence gene were selected to cover different serotypes and *stx* subtypes relevant for the current epidemiological situation in Europe (Annexes 4–5). The 12 test isolates for cluster analysis were selected to include isolates with different or varying relatedness and different sequence types (ST17, 20 and 386). A set of technical triplicates was included in the cluster test isolates. (Annexes 6–9, 11–12). Using either PFGE or WGS-derived data, the cluster of closely related isolates consisted of five STEC ST17 isolates (A technical triplicate). The characteristics of all the STEC test isolates are listed as 'Original/REF' in Annexes 4–12.

Table 1. Characterisation of test isolates

Parts	Number of test isolates	Characterisation	Annexes
Serotyping	12[#]	O55:H7, O76:H7/H-, O91:H14, O91:H21, O111;H8/H-, O121:H19, O126:H27/H-, O128:H2, O145:H28/H-, O154:H31, O157:H7/H-, O187:H28	4
Virulence profile determination	12[#]	*eae stx1a, eae stx2a, stx1a stx2b, stx2d, eae stx1a stx2a, eae stx2a, aaiC aggR* (x1), *stx1c stx2b, eae stx2a, stx1d, eae stx2c, stx2g*	5
Cluster analysis	12	ST17 (x9) (O103:H2 *stx1a*), ST20 (x2) and ST386	6–14

#: same 12 isolates.

2.3 Carriage of isolates

All test isolates were blinded and shipped on 5 February 2019. Letters stating the unique isolate IDs were included in the packages and distributed individually to the participants by email on 5 February 2019 as an extra precaution. Seventeen participants received the isolates within one day, eight within three days and one 14 days after shipment respectively. No participants reported damage to the shipment or errors in the unique isolate IDs.

In February 2019, instructions for the submission of results procedure were emailed to the participants. This included the links to the online uploading site and submission form.

2.4 Testing

In the serotyping part, 12 STEC isolates were tested to assess the participants' ability to obtain the correct serotype. The participants could perform conventional serological methods according to suggested protocol [9] or molecular-based serotyping (PCR or WGS). The serotypes were submitted in the online form.

The same set of isolates for serotyping analysis was used for the virulence profile determination. The analyses were designed to assess the participants' ability to obtain the correct virulence profile. The participants could choose to perform detection of *aaiC* and *aggR* (two genes related to EAEC), *eae* and *stx1* and *stx2,* as well as subtyping of *stx* genes *stx1 (stx1a, stx1c or stx1d)* and *stx2 (stx2a - stx2g),* according to suggested protocol [10]. The results were submitted in the online form.

In the molecular typing-based cluster analysis part, participants could perform the laboratory part using PFGE or WGS-derived data. Participants were instructed to report the IDs of isolates included in the cluster of closely related isolates by method. If PFGE analysis was conducted, the participant reported the total number of bands and number of shared bands with a selected cluster representative isolate.

Laboratories performing WGS could use their own analysis pipeline for cluster analysis, e.g. single nucleotide polymorphism analysis (SNP-based) or whole genome Multi Locus Sequence Typing (wgMLST)/cgMLST (allele-based) and were asked to submit the isolates identified as a cluster of closely related isolates based on the analysis used. Laboratories could report results from up to three analyses (1 main and 0–2 additional), but the detected cluster had to be based on results from the main analysis. The laboratories reported SNP distance or allelic differences between a selected cluster isolate and each test isolate and uploaded the raw reads (FASTQ files) to an SFTP site.

2.5 Data analysis

As the participating laboratories submitted their results, the serotype, virulence profile and cluster analysis results, as well as the participants' uploaded raw reads, were imported to a dedicated STEC EQA-9 BioNumerics (BN) database. If errors in the submission process were identified, the EQA provider reported this to participants, thereby obtaining analysable results. The EQA provider was in contact with five participants in order to ensure sequences were uploaded to the SFTP site.

Serotyping results were evaluated according to the percentage of correct results, generating a score from 0–100% for O group, H type and OH serotype.

The virulence profile determination results were evaluated according to the percentage of correct results, generating a score from 0–100% for *eae*, *aaiC*, *aggR*, *stx1*, *stx2*, subtyping of *stx1* and *stx2* and combined subtype.

Molecular typing-based cluster analysis was evaluated according to correct or incorrect identification of the expected cluster of closely related isolates based on a pre-defined categorisation by the organiser. The EQA provider's PFGE results were based on *Xba*I profiles [11]. The EQA provider's WGS-derived cluster analysis was based on allele-based cgMLST [12] and SNP analysis (NASP) [13]. The correct number of closely related STEC isolates (5) could be identified by both PFGE and WGS-derived data. The cluster contained five ST17 isolates: REF13, REF14, REF16, REF22 and REF24 (REF14, REF16 and REF22 were technical triplicates). The EQA provider found at most two allele differences or four SNPs between any two isolates in the cluster. The rest of the cluster test isolates were an additional four ST17s, two ST20s, and one ST386.

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

3. Results

3.1 Participation

Laboratories could either participate in the full EQA scheme or one part only (serotyping, virulence profile determination or molecular typing-based cluster analysis). Of the 26 participants who signed up, 24 completed and submitted their results. The majority of participants (63%; 15/24) completed the EQA with analysis in each of the three parts. In total, 20 (83%) participants participated in serotyping, 23 (96%) participated in the detection of one or more of the virulence genes and 17 (71%) in cluster analysis. (Table 2).

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

	Serotyping[1]	Virulence profile determination[2]	Cluster analysis[3]
Number of participants	20	23	17
% of participants	83*	96*	71*

[1]: O grouping and/or H typing
[2]: detection of at least one gene (aaiC, aggR, eae, stx1 and stx2) and/or subtyping of stx1 *and* stx2
[3]: molecular typing-based cluster analyses based on PFGE or WGS-derived data
**: percentage of the total number (24) of participating laboratories.*

O grouping results were provided by 20 participants (83%) and H typing results were provided by 13 (54%). The majority 10/20 (50%) used WGS-based serotyping (Annex 4). Almost all participants (96%, 23/24) participated in the detection of virulence genes (*stx1* and *stx2*). Slightly fewer 92% (22/24) participated in the detection of *eae*. Detection of enteroaggregative genes *aaiC* and *aggR* were reported by 75-79%, (18-19/24) and 79% (19/24) participated in the *stx* subtyping. Most participants (50%, 12/24) reported cluster analysis using WGS-derived data, while eight (33%) reported using PFGE data. Three submitted cluster data based on both PFGE and WGS (Table 3). In all parts of the EQA, laboratories mainly reported 'Laboratory policy to enhance the typing quality' as the reason for participating, with accreditation needs, and institute/national policy also reported (Annex 3).

Table 3. Detailed participation information for the parts of serotyping, virulence profile determination and molecular typing-based cluster analysis

	Serotyping n=17		Virulence profile determination n=23					Cluster analysis n=17		
	O group	H type	*aaiC*	*aggR*	*eae*	*stx1* and *stx2*	*stx* subtyping	PFGE	WGS	Both
Number of participants	20[#]	13Δ	18	19	22	23	19	5	9	3
Percentage of participants^	100%	65%	78%	83%	96%	100%	83%	30%	53%	18%
Percentage of participants *	83%	54%	75%	79%	92%	96%	79%	21%	38%	13%

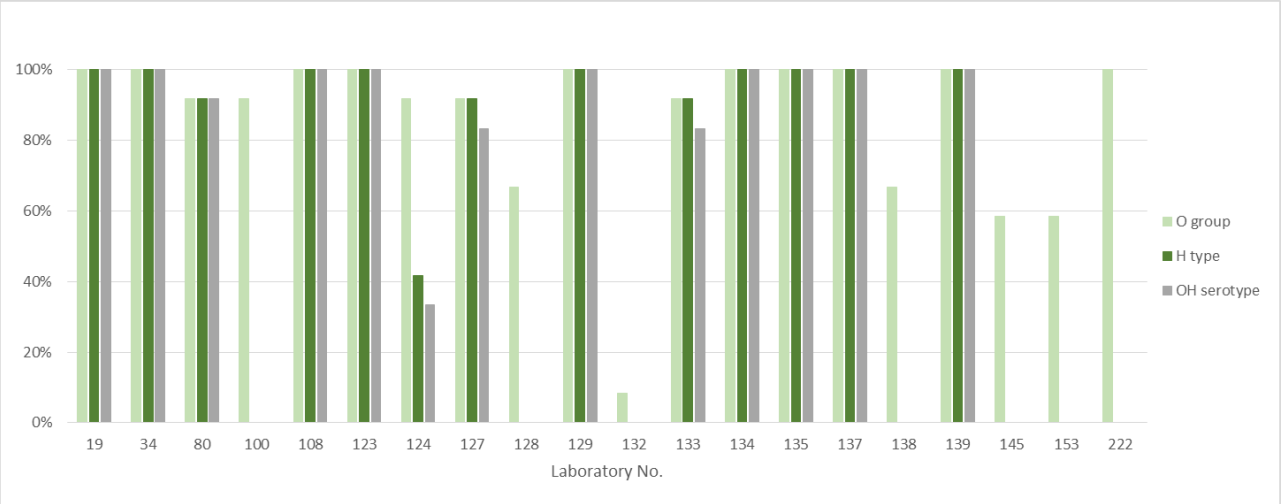
^: percentage of participants in respective part of EQA
**: percentage of total number of participating laboratories (24)*
[#]: phenotypic (n=9)/PCR-based (n=1)/WGS-based (n=10)
Δ: phenotypic (n=2)/PCR-based (n=1)/WGS-based (n=10).

3.2 Serotyping

Twenty (83%) laboratories performed O grouping and ten (50%) of the 20 were able to type all 12 test isolates correctly, giving an average score of 86% (Figure 1). Thirteen laboratories (65%) reported the correct O group for the rare O group O187 (isolate REF2) and 14 (70%) correctly reported O76 (isolate REF4) (Figure 2). The highest performances were displayed for the O157 isolates (100%), O55, O121, O128 and O145 (95%; Figure 2), some included in the minimum requirements of ECDC [14]. One laboratory (132) detected O157 only, generating incorrect (non-O157) results for the 11 other isolates (Annex 4, Figure 1).

Thirteen (54%) laboratories performed H typing. The general performance for H typing was higher than O grouping, with nine (69%) of participants correctly H typing all 12 test isolates, resulting in an average score of 94% (Figure 1). Only two laboratories reported incorrect H type, while seven of the eight additional incorrect results were reported as NT and H by one laboratory (89%; 8/9; Annex 4).

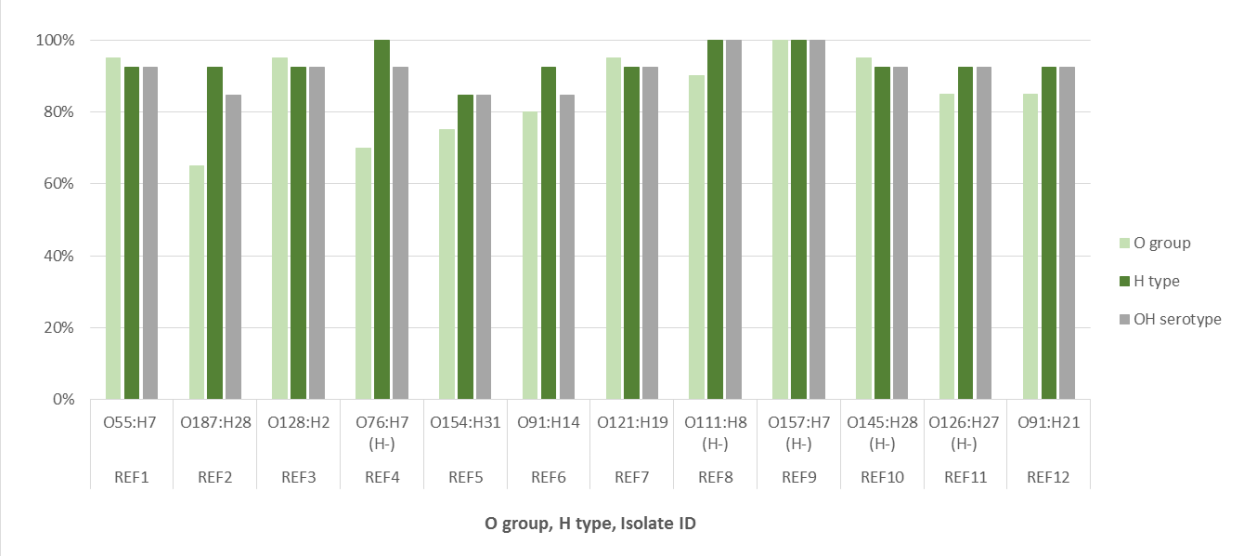
Figure 1. Participant percentage scores for O grouping and H typing



Arbitrary numbers represent participating laboratories.
Bars represent the percentage of correctly assigning O groups (light green), n= 20 participants, H types (dark green), n=13 participants, Combined O:H serotypes (grey), n=13 participants.

Complete O:H serotyping was performed by 13 (54%) participants with an average score of 92%, ranging from 85% (11/13) for O187:H28 to 100% (13/13) for isolate O157:H7 (REF9) and O111:H8 (REF8) of the participants reporting the correct serotype (Figure 2).

Figure 2. Average percentage test isolate score for serotyping of O and H



Bars represent the percentage of laboratories correctly assigning O groups (light green): n=20 participants.
H types (dark green): n=13 participants.
Combined O:H serotypes (grey): n=13 participants.
Average scores: O group, 86%; H type, 94% and combined O:H serotype, 92%.

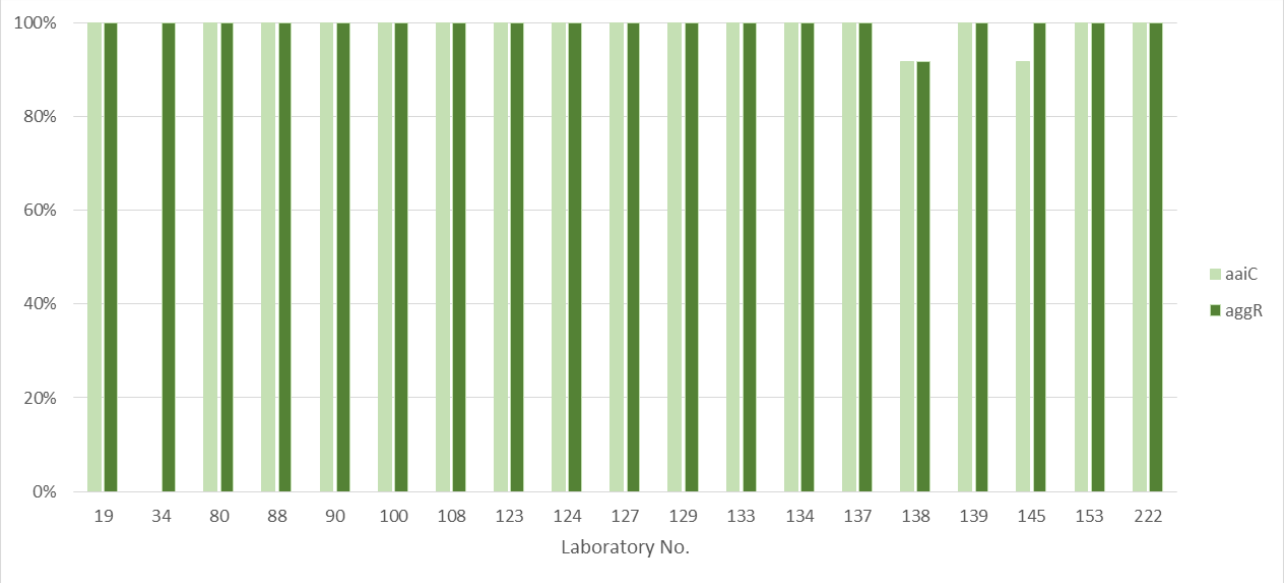
3.3 Virulence profile determination

Between 18–23 laboratories submitted results for each of the virulence genes, consisting of detection of EAEC *(aaiC* and *aggR*) and virulence genes (*eae*, *stx1* and *stx2*) and subtyping of *stx1* and *stx2* genes. Twenty-two participants submitted results for *eae* and twenty-three *stx* genes. Nineteen laboratories submitted subtyping results of *stx1* and *stx2* genes and the EAEC genes *aggR*. Eighteen reported results for the *aaiC.*

3.3.1 Detection of EAEC genes (*aaiC* and *aggR*)

The performance of the laboratories reporting correctly genotyping results for EAEC, *aaiC* (89%; 16/18) and *aggR* (95%; 18/19) was high (Figure 3). The average scores were 99% (*aaiC*) and almost 100% (*aggR*) respectively (Annex 5). One laboratory only reported *aggR* results but not *aaiC*.

Figure 3. Participant percentage scores for genotyping of *aaiC* and *aggR*

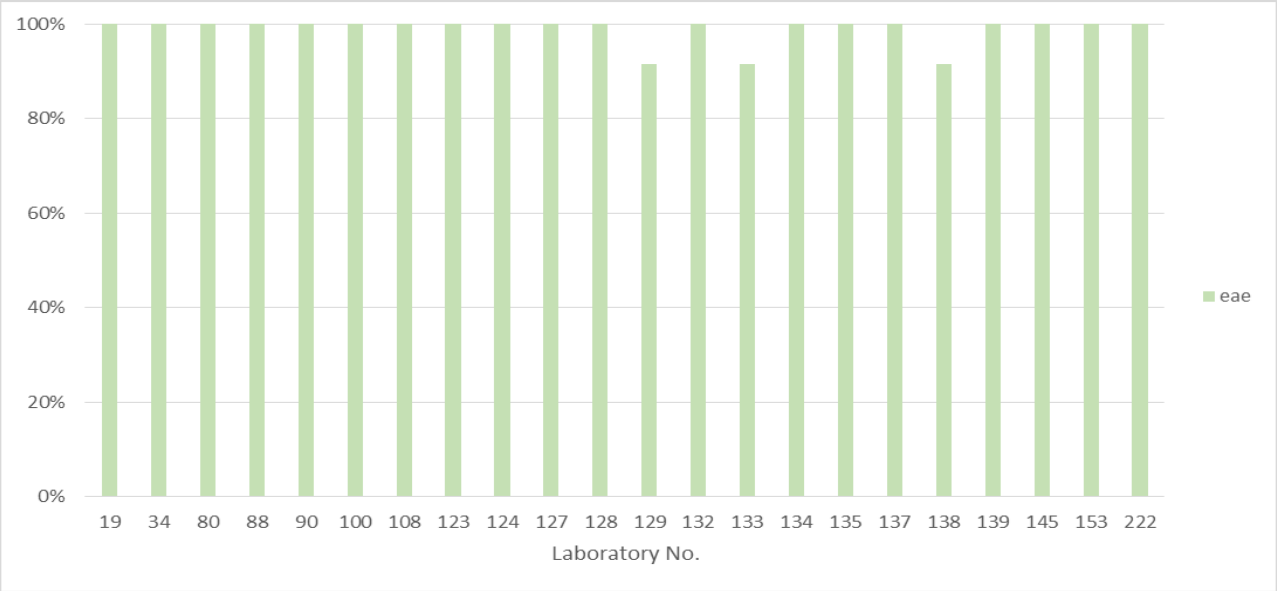


Arbitrary numbers represent participating laboratories.
Bars represent percentage of correct genotyping of aaiC *(light green) n=18 participants and* aggR *(dark green): n=19 participants.*

3.3.2 Detection of virulence genes *eae, stx1* and *stx2*

Detection of virulence genes *eae*, *stx1* and *stx2* was performed by 22-23 (92-96%) laboratories with a generally high performance (Figures 4–5). For *eae* detection, 19 (86%) laboratories obtained a 100% score (Figure 4). Three laboratories (129, 133 and 138) reported incorrect *eae* results for one isolate each (not the same isolate). In total, *eae* was only misidentified by one false negative for REF4 and two false positive for REF2 and REF11 respectively (Annex 5).

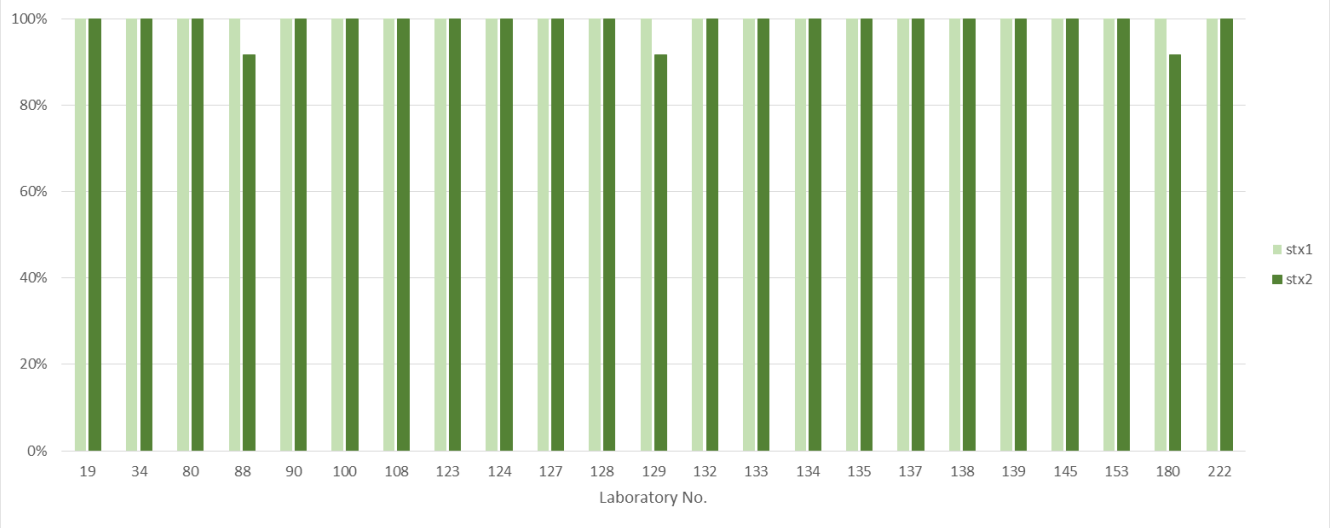
Figure 4. Participant percentage scores for genotyping of *eae*



Arbitrary numbers represent participating laboratories.
Bars represent percentage of correct genotyping of eae *(light green): n=22 participants.*

The performance of detection of *stx*1 and *stx*2 genes was high; all 23 (100%) laboratories reported 100% correct *stx1* results and 20 (87%) laboratories reported 100% correct *stx2* results (Figure 5). The three incorrect *stx2* results were reported by three different laboratories in three different isolates. (Annex 5).

Figure 5. Participant percentage scores for detection of *stx1* and *stx2*



Arbitrary numbers represent participating laboratories.
Bars represent percentage of correct genotyping of stx1 *(light green) and* stx2 *(dark green): n=23 participants.*
Average scores: stx1, *100%;* stx2, *99%.*

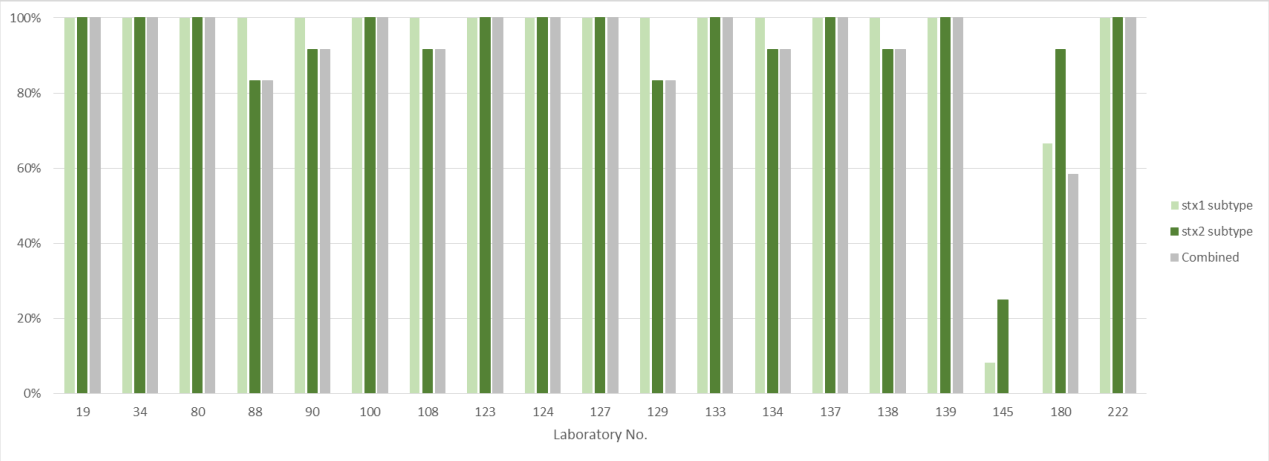
3.3.3 Subtyping of stx1 and stx2

Subtyping of *stx1* and *stx2* was performed by 19 laboratories. Seventeen (89%) subtyped *stx1* correctly and eleven (58%) reported correct *stx2* subtype for all ten test isolates (Figure 6). (Annex 5).

Only two laboratories (148, 180 - 11%) reported an incorrect subtype of *stx1* for three or more isolates. Laboratories were not allowed to only report results for selected test isolates for a particular test, so reporting ND was considered as an incorrect result if the laboratory reported results of other isolates for that test. In total, the average score was 93% (Figure 6). Laboratory 145 incorrectly reported *stx1a* and *stx1c* for four isolates which only should have been reported as *stx1c* for REF3 and *stx1a* for REF1, REF6 and REF8. Laboratory 180 incorrectly reported *stx1a* and *stx1c* for two isolates which only should have been reported as *stx1c* for REF3 and *stx1a* for REF8. In addition laboratory 180 also incorrectly reported *stx1a* and *stx1d* for one isolates which only should have been reported as *stx1d* (REF5).

Eight laboratories (42%) reported an incorrect subtyping of *stx2* for one or more isolates, primarily by reporting double *stx* subtypes for isolates which only had one. In total, the average score was 92% (Figure 6). The number of instances of mis-subtyping *stx2* was 18, of which 13 consisted of reporting too many *stx* subtypes per isolate. Laboratory 145 incorrectly reported result for nine of 12 isolates in *stx2* subtyping and 11 of 12 isolates in *stx1* subtyping.

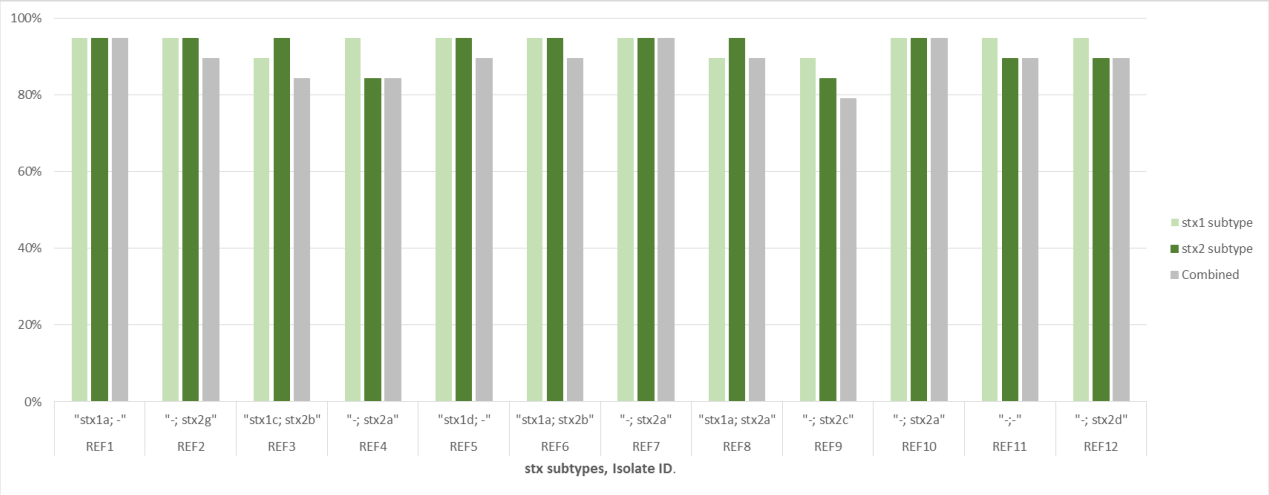
Figure 6. Participant percentage scores for subtyping of *stx1* and *stx2*



Arbitrary numbers represent participating laboratories.
Bars represent percentage of correct subtyping of stx1 (light green), stx2 (dark green), combined stx1 and stx2 (grey), n=19 participants. Reporting ND (not done) evaluated as incorrect.

Around 58% (11/19) of the participants were able to correctly *stx2* subtype all 12 test isolates (Figure 6). Among the *stx2* subtypes, REF4 with only *stx2a* was difficult to type using WGS, as a mapping analysis could identify both *stx2a* and/or *stx2c*. When analysing the protein sequence of the B subunit, an 'EDD' motif was identified and *stx2a* was the correct result [7].

Figure 7. Average percentage test isolate score for subtyping of *stx1* and *stx2*



Bars represent percentage of laboratories correctly subtyping stx1 (light green), stx2 (dark green) and combined stx1 and stx2 (grey), n=19
Average scores: stx1, 93%; stx2, 92% and combined stx1 and stx2, 89%.

Incorrect *stx2* subtype results were reported 18 times, the majority is no longer due to reporting ND instead of negative, (3/18) as in EQA-8. The incorrect results of *stx2* subtyping shown in Table 4 are divided into three categories: false negatives (1/18), incorrect subtype of *stx2* (15/18) or ND (3/18). Laboratory 145 reported six of the incorrect *stx2* subtypes and three of the ND results. Laboratory 129 and 88 reported two incorrect *stx2* subtyping results each.

Table 4. Incorrect *stx2* subtype results

		Incorrect subtype results			
Isolate ID	**EQA provider**	**False negative**	**Incorrect**	**Total true errors**	**Errors by reporting ND#**
REF1	-				1
REF2	*stx2g*		*stx2a+ stx2g* (1)	1	
REF3	*stx2b*		*stax2a + stx2b* (1)	1	
REF4	*stx2a*		*stx2c* (1) *stx2a + stx2c* (2)	3	
REF5	-				1
REF6	*stx2b*	1		1	
REF7	*stx2a*		*stx2a + stx2c + stx2d* (1)	1	
REF8	*stx2a*		*stx2a + stx2c + stx2d* (1)	1	
REF9	*stx2c*		*stx2c + stx2d* (2) *stx2a + stx2c* (1)	3	
REF10	*stx2a*		*stx2a + stx2c + stx2d* (1)	1	
REF11	-		*stx2a + stx2c* (1)	1	1
REF12	*stx2d*		*stx2c + stx2d* (1) *stx2a + stx2c + stx2d* (1)	2	
Total				**15**	**3**

ND#: not done.

3.4 Molecular typing-based cluster analysis

In this part of the EQA, participants should have correctly identified a cluster of closely related isolates among 12 test isolates by using either PFGE and/or WGS-derived data. The cluster test isolates were pre-categorised by the EQA provider.

The EQA provider's PFGE results were based on an *XbaI* profile. The EQA provider's cluster analysis of WGS-derived data was based on allele-based (cgMLST [12]) and SNP analysis (NASP [13]). The correct number of closely related isolates could be identified by both PFGE and WGS-derived data. The cluster contained five O103:H2 (*stx1a*), ST17 isolates: REF13, REF14, REF16, REF22 and REF24 (REF14, REF16 and REF22 were technical triplicates). The EQA provider found at most two allele differences or three SNPs between any two isolates in the cluster. The rest of the cluster test isolates were an additional four ST17s, two ST20s, one ST386 (Annexes 6–14).

3.4.1 PFGE-derived data

Of the 24 participants in the EQA, eight (33%) performed cluster analysis using PFGE-derived data. Five (63%) correctly identified the cluster of closely related isolates defined by a pre-categorisation from the EQA provider among the 12 cluster test isolates. Table 5 shows the overview of the isolates each participant included or excluded in cluster identification. Laboratory 130 missed one of the technical triplicated isolates (REF16), and included additional four isolates (ST17) in their cluster of closely related isolates. Laboratory 132 also include additional four isolates of ST17. Laboratory 222 only included one additional ST17 (REF21) in the cluster of closely relates isolates.

Table 5. Results of cluster identification based on PFGE-derived data

		Laboratory number							
Isolate ID	**ST**	**19**	**90**	**123**	**124**	**127**	**130**	**132**	**222**
REF13‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF14‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF15	386	No	No	No	No	No	No	No	No
REF16‡#	17	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
REF17	20	No	No	No	No	No	No	No	No
REF18	20	No	No	No	No	No	No	No	No
REF19	17	No	No	No	No	No	Yes	Yes	No
REF20	17	No	No	No	No	No	Yes	Yes	No
REF21	17	No	No	No	No	No	Yes	Yes	Yes
REF22‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF23	17	No	No	No	No	No	Yes	Yes	No
REF24‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster-identified conclusion		Yes	Yes	Yes	Yes	Yes	No	No	No

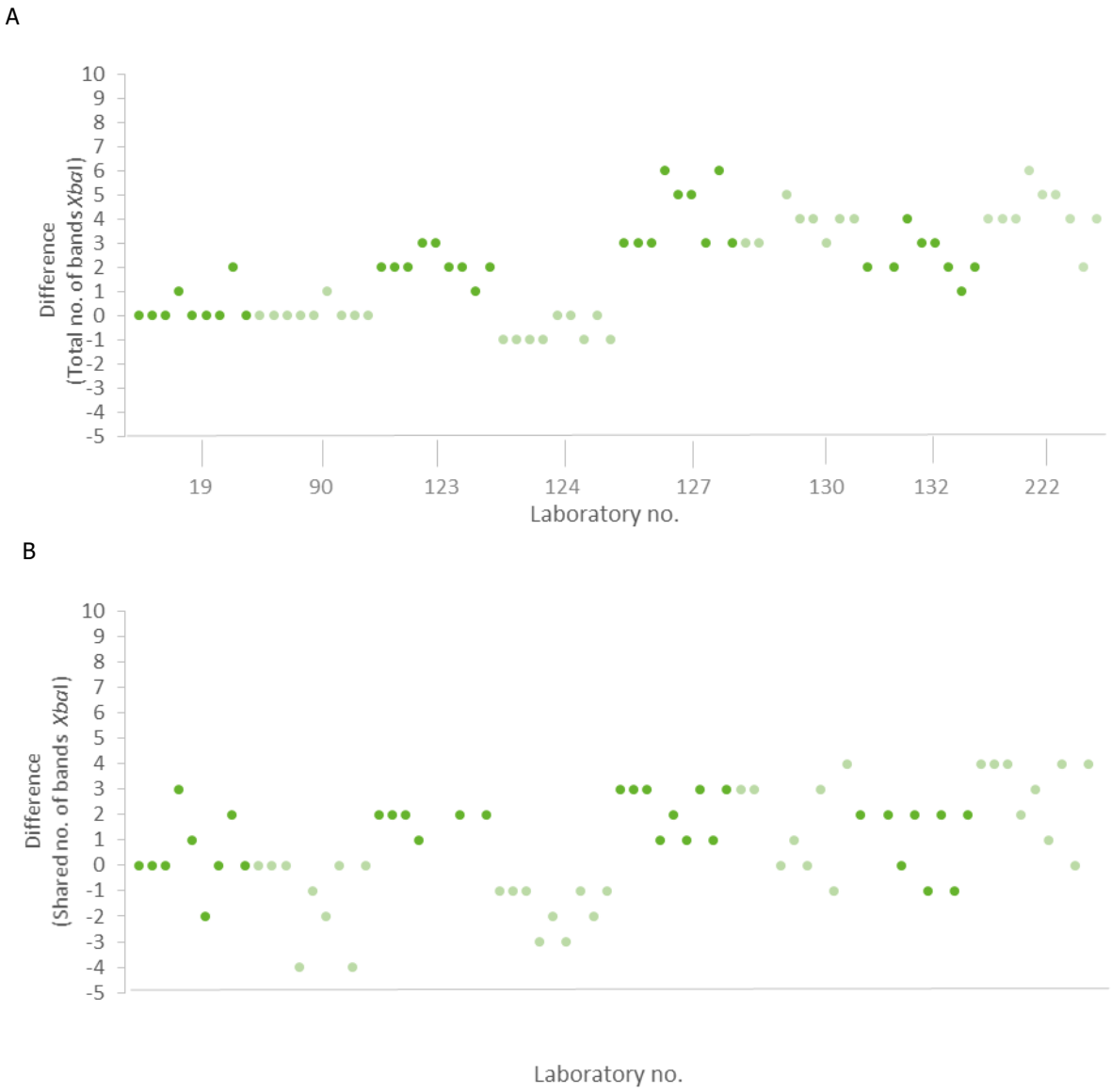
‡: closely related isolates (in grey)
#: technical triplicates isolates (in bold)
(Annex 8).

For each isolate, participants were instructed to report the total number of bands in the *Xba*I profile. The number of bands shared between each test isolate and the selected cluster representative was also reported (Figure 8, Annex 9).

Figure 8A shows the difference between the number of bands reported by the participants and the number observed by the EQA provider for *Xba*I. The PFGE profile of *E. coli* contains a large number of bands within the region of 200–350 kb, which make the cluster analysis based on PFGE harder to interpret compared with other species such as *Salmonella* or *Listeria*. This is illustrated in Figure 8, which shows a high number of variations. The reported total number of bands in the cluster profile varied from 15-20 bands (Annex 9). The three laboratories (130, 132 and 222) not identifying the correct cluster reported a low number of bands in the cluster profile, however laboratory 123 also reported a low number of bands but were able to correctly identify the cluster.

Figure 8B shows the difference between the participants' reported number of shared bands with a selected cluster representative and the number observed by the EQA provider for *Xba*I. Again high number of variations in the number of reported shared bands. Laboratories 130 included profiles in their cluster with a 2 bands difference (based on the reported number of shared bands), and missed to report REF16 as a part of the cluster. Laboratory 132 included profiles in their cluster with serval band difference (based on the reported number of shared bands). Laboratory 222 did not miss any of the cluster isolates but included on additional isolates (REF21), which only has one band difference accordingly to the EQA provider profiles).

Figure 8. Difference between reported number of bands (A) and shared bands (B) for each isolate to selected isolate



Data from all nine O103:H2, ST17 isolates: REF13, REF14, REF16, REF19, REF20, REF21, REF22, REF23 and REF24. Laboratory 130 reported zero in the total number of bands and zero for shared bands for REF16 (excluded in the figure 8), laboratory 132 did not report any number of bands for REF13 and laboratory 123 only reported data for six of the nine isolates (Annex 9).

3.4.2 WGS-derived data

3.4.2.1 Reported results from participants

Twelve participants (50%) performed cluster analysis using WGS-derived data. Only one laboratory reported using external assistance for sequencing. Different sequencing platforms were listed among the participants: 2 MiniSeq, 5 MiSeq, 1 HiSeq, 1 NextSeq, 1 Novaseq, 2 Ion Torrent (Ion GeneStudio S5 System and Ion Torrent). All reported using commercial kits for library preparation. Out of the 12 participants, nine (75%) used Illumina's Nextera kit. Three participants reported volume changes from the manufacturer protocol (Annex 10).

Performance was high in cluster analysis with WGS-derived data. Ten participants (83%) correctly identified the cluster of closely related isolates defined by pre-categorisation from the EQA provider among the 12 test isolates (Table 6). Two laboratories did not include an ST17 isolate in the cluster of closely related isolates, respectively REF13 or REF16, however laboratory 108 left a comment where the isolate was included in the cluster, listed as a (No) for REF13 and (Yes) for the cluster-identified.

Table 6. Results of cluster identification based on WGS-derived data

	ST	Laboratory number											
Isolate ID	**ST**	**19**	**34**	**80**	**100**	**108**	**123**	**133**	**134**	**135**	**137**	**139**	**222**
REF13‡	17	Yes	Yes	Yes	Yes	(No)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF14‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF15	386	No	No	No	No	No	No	No	No	No	No	No	No
REF16‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
REF17	20	No	No	No	No	No	No	No	No	No	No	No	No
REF18	20	No	No	No	No	No	No	No	No	No	No	No	No
REF19	17	No	No	No	No	No	No	No	No	No	No	No	No
REF20	17	No	No	No	No	No	No	No	No	No	No	No	No
REF21	17	No	No	No	No	No	No	No	No	No	No	No	No
REF22‡#	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
REF23	17	No	No	No	No	No	No	No	No	No	No	No	No
REF24‡	17	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main analysis		Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	SNP	Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	Allele (cgMLST)	SNP	Allele (cgMLST)	Allele (cgMLST)
Additional analysis													
Cluster-identified		Yes	Yes	Yes	Yes	(Yes)	Yes	Yes	Yes	Yes	Yes	Yes	No

‡: closely related isolates (in grey)
#: technical triplicates isolates (in bold)
ST: sequence type
Allele: allele-based analysis.
SNP: single-nucleotide polymorphism
(Annex 11).

Laboratories were instructed to report the data analysis used for cluster identification and select a representative isolate in the cluster for reporting SNP distance or allelic differences between the selected isolate and each test isolate included in analysis. Laboratories could report results from up to three analyses (one main and up to two additional), but the detected cluster had to be based on results from the main analysis.

The two participants using SNP both used a reference-based approach with different ST17 isolates as a reference. One used Burrows-Wheeler Aligner (BWA) and the other used CLC as the read mapper, and also selected to use different variant callers (Table 7).

Table 7. Results of SNP-based cluster analysis

Lab	SNP-based					
	Approach	**Reference**	**Read mapper**	**Variant caller**	**Distance within cluster**	**Distance outside cluster¤**
Provider	Reference-based	ST17 (REF14)	BWA	GATK	0–3	50-153 (913-4665)
		Recombination				
108	Reference based	ST17 9201	CLC assembly cell v4.4.2	CLC assembly cell v4.4.2	0–2	4-591 (not reported)
137	Reference based	CC17 ST17 12009	BWA v0.7.12	GATK v2.6.5	0–4	59-160 (510-1521)

¤: reported distance to ST17 (non-ST17) isolates (Annex 12).

All ten participants that used allele-based analysis selected this method as the main analysis for cluster detection – none reported additional analysis (Table 8). Six of the ten (60%) used an assembly-based allele calling method and the other four laboratories used both mapping- and assembly-based allele calling (Table 8).

Table 8. Results of allele-based cluster analysis

Laboratory	Allele-based analysis						
	Approach	Allelic calling method	Assembler	Scheme	Number of loci	Difference within cluster	Difference outside cluster¤
Provider	BioNumerics	Assembly- and mapping-based	SPAdes	Applied Maths (cgMLST/Enterobase)	2513	0–2	23-54 (179-557)
19	BioNumerics	Assembly and mapping based	SPAdes	Applied Math (cgMLST/Enterobase)	2513	0-2	22-47 (179-553)
34	SeqPhere	Assembly- based and mapping-based	SPAdes	Enterobase (cgMLST)	2971	0-3	30-2815 (247-750) Δ 30-75 (247-750)
80	SeqPhere	Only assembly-based	Velvet 1.1.04	Enterobase (cgMLST)	2513	0-2	21-49 (177-547)
100	SeqPhere	Only assembly-based	Velvet	Enterobase (cgMLST)	2513	0-2	21-98 (268-631)
123	SeqPhere	Only assembly-based	SPAdes	Enterobase (cgMLST)	2513	0-2	21-48 (not reported)
133	BioNumerics	Assembly- and mapping-based	SPAdes	Applied Math (cgMLST/Enterobase)	2513	0-2	11-20 (150-184)
134	SeqPhere	Only assembly-based	Velvet	Enterobase (cgMLST)	2513	0-2	21-100 (276-652)
135	SeqPhere	Only assembly-based	SPAdes	Enterobase (cgMLST)	2513	0-2	21-101 (277-651)
139	Enterobase	Assembly-and mapping-based	Enterobase	Enterobase (cgMLST)	2598	0-5	50 (200-1100)
222	chewBBACA on ARIES webserver	Only assembly based	SPAdes 3.11 followed by assembly optimization with Pilon through ARIES webserver	cgMLST scheme developed by INNUENDO EFSA-funded project through the curation of Enterobase scheme and finally including 2360 loci.	2360	0-9	23-62 (193-573)

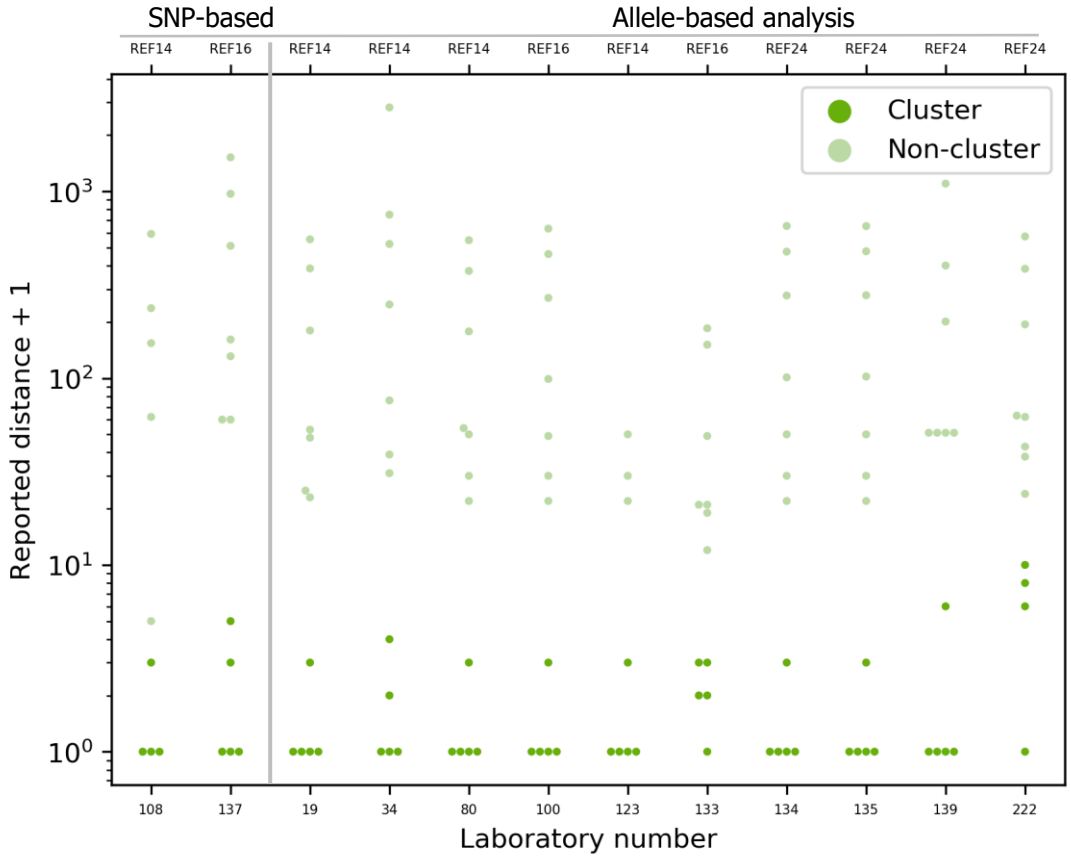
¤: reported differences to ST17 (non-ST17; Annex 12)
Lab 34, has sequenced an incorrect isolate as REF19 (Δ is without this isolate).

Out of the ten laboratories using allele-based methods, nine (90%) identified the correct cluster of five closely related isolates (Table 8–9). These nine laboratories performed cgMLST using the same scheme as the EQA provider (cgMLST/Enterobase [12]) and only one, laboratory 222 used the INNUENDO cgMLST scheme 2360 loci (and did not identify the correct cluster). Laboratory 34 sequenced one isolate which is not an EQA test isolate (REF 19).

The nine laboratories that identified the correct cluster reported allele differences of 0–5 within the cluster of closely related isolates (Figure 9, Table 8). The laboratory (222) not identifying the correct cluster reported allele difference of 0-9.

Four other test isolates (REF19, REF20, REF21, and REF23) were also ST17, but not pre-defined by the EQA provider as part of the cluster. Based on cgMLST, nine laboratories reported allele differences to the selected cluster isolate at 11–101 for this group of isolates, if the data on REF19 from laboratory 34 is discarded. Only one of the laboratories (123) did not report results for the three non-ST17 isolates (REF15, REF17 and REF18). Based on cgMLST, the reported differences were 150–1100 for non-ST17 isolates (Table 8, Annex 12).

Figure 9. Reported SNP distances or allelic differences for each test isolate to selected cluster representative isolate



SNP: single nucleotide polymorphism

Selected cluster representative marked as REF.
Dark green: reported cluster of closely related isolates, Light green: not reported as part of cluster.

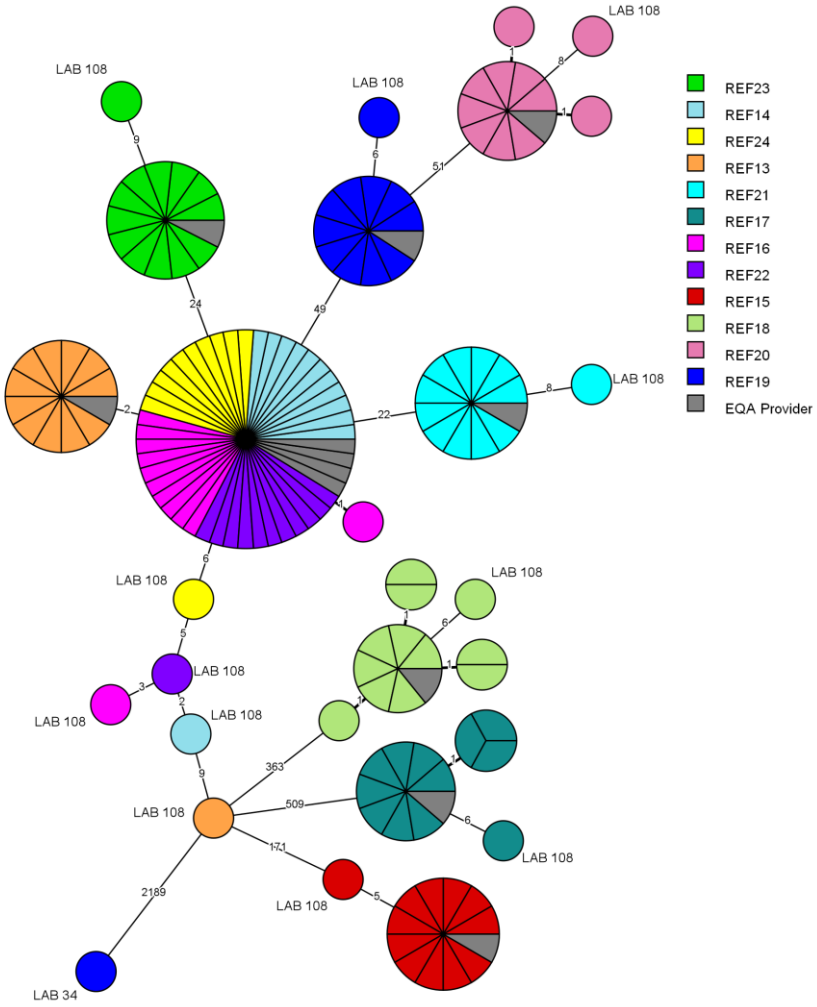
Of the two laboratories performing SNP analysis, only laboratory 137 identified the correct cluster of closely related isolates and reported SNP distances within the cluster from 0–4. Laboratory 108 did not include REF13 in the reported cluster, however they made a comment at the end of the submission stating that 'REF13 were included in the cluster with 4 SNPs' (Figure 9).

3.4.2.2 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 (Enterobase) [12] and evaluated by the EQA provider's in-house quality control (QC) pipeline [15].

The overall cgMLST analysis, shown in the minimum spanning tree (MST) based on submitted raw reads (FASTQ files) from 12 laboratories reveals clear clustering of the results for each test isolate (Figure 10). Laboratory 34 appears to have sequenced and included an isolate which is not an EQA isolates instead of their REF19. Two results from laboratory 133 were omitted from the analysis presented in Figure 10, sequences from one isolate were not uploaded to the EQA provider and one was not included due to low quality.

Figure 10. Minimum spanning tree of core genome Multi Locus Sequence Typing participant FASTQ files

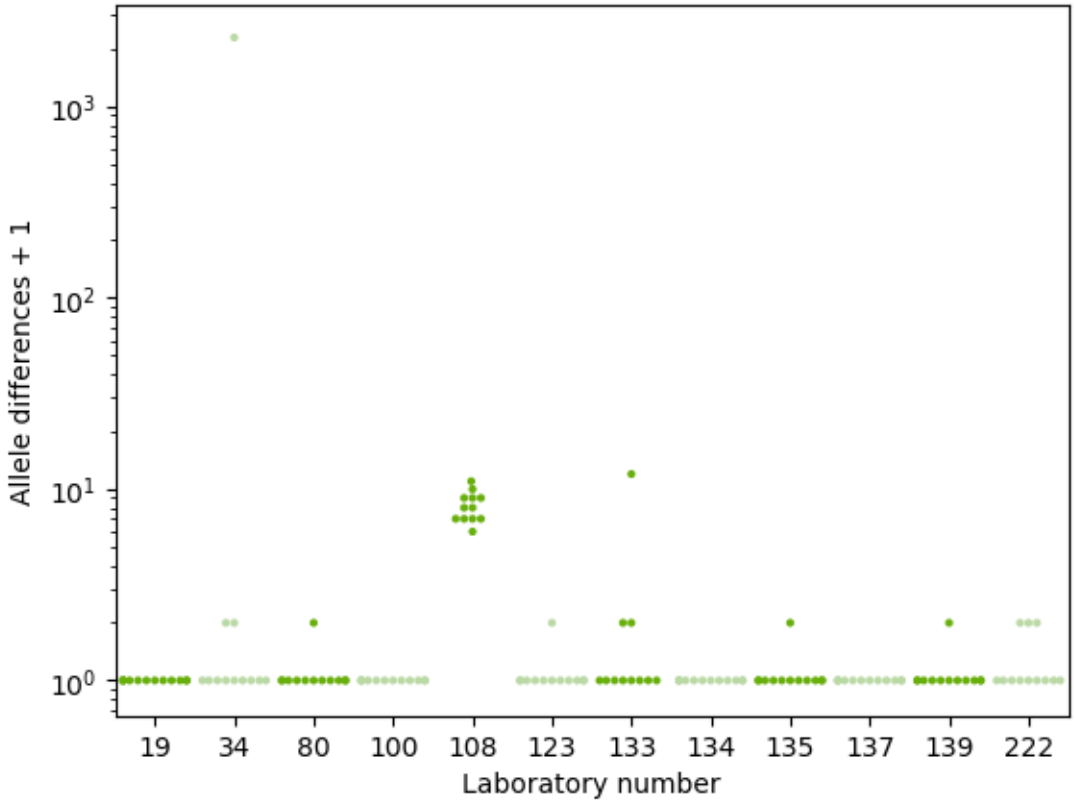


Minimum spanning tree (MST) in log scale of core genome Multi Locus Sequence Typing (cgMLST) [12] based on submitted raw reads (FASTQ files).
Each of the REF1–12 test isolates have a different colour. REF results from the EQA-provider are in grey.
Laboratory 133 is missing two isolates (REF20) were not uploaded, and REF 24 were discarded due to low quality. Results from laboratory 108 and 222 were run in CE (using Ion Torrent setup for allele calling).

The allele differences in Figure 10 do not exactly match those illustrated in the individual reports and consequently those in Figure 11, where the same data are used. This discrepancy is caused by loci being dropped if they did not pass QC for all isolates in the analysis. Joint analysis accordingly contains fewer loci.

For each laboratory, cgMLST was performed on the submitted raw reads (FASTQ files), applying Applied Maths allele calling with the Enterobase scheme [12]. A hierarchical single linkage clustering was performed on the submitted data for each laboratory along with the EQA provider's reference isolates. Figure 11 shows the allele differences between each submitted sequence and the corresponding reference.

Figure 11. Participant allele difference from reference result (EQA-provider) for each test isolate



Allele difference from corresponding REF isolates (EQA provider) based on submitted raw reads (FASTQ files) and analysed by EQA provider.

For 119 of 143 results (83%), no allele difference was identified. For eleven results (8%) a difference of 1 allele from the REF isolate was calculated and for thirteen results (9%) a difference of 5–10 alleles was observed, 12 out of 13 results were reported by laboratory 108. In addition, results from laboratory 34 identified a difference of 2300 alleles, caused by the laboratory submitting and analysing of an isolate which is not an EQA isolate instead of REF19.

Separately, the laboratories listed quantitative and qualitative QC parameters used to evaluate their data. As seen in Table 9, confirmation of genus and coverage with acceptance thresholds ranging from 20–50X was the most widely used QC parameter, hereafter genome size and number of good cgMLST loci were also listed as an important parameters for QC. Refer to the full list of QC parameters reported by the participants in Annex 13.

Table 9. Summary of selected QC parameters reported by participants

Laboratory	Confirmation of genus	Coverage	Q score (Phred)	Genome size	No. of good cgMLST loci
19	Cut off: max 5% of another genus	50x	No	4,6-5,6 Mbp	core percentage called
34	BIOLOG GEN III	100% coverage of reference sequence	Fast QC, Trimmomatic	No	Varscan
80	KRAKEN	No	No	No	>90% good targets
100	PhiX control	>40	FastQC	length of assembled genome (app 5.2 Mbp)	No
108	Species total match size average similarity	> 20 x	No	1	Similarity, Coverage >10 times
123	high percentage of cgMLST targets (E. coli scheme) is required.	>50	No	>5.000	>98%
133	No	50X	No	3.9-6.5mb	No
134	No	No	No	length of contigs assembled < ref genome + 10%	cgMLST alleles found and called >95 %
135	KrakenBracken	>30	No	between 4,2-4,8Mbp	No
137	Kmer ID	Average depth coverage (SNP-typing)	No	No	Average coverage of all alleles (Achtman 7 gene MLST)
139	Enterobase	40	No	N50 > 70 000	No
222	No mismatches 7 MLST panel	Mentalist tool was used to perform MLST. 100% coverage in lenght for all the seven genes of the scheme and at least 30x average depth of coverage among all of the seven genes were considered as threashold values.	No	No	No
% of laboratories using the QC parameter	**83%**	**83%**	**17%**	**67%**	**58%**

Text adjusted for purpose of the table, as the questions (Q score (Phred)) had been misunderstood by a large number of participants (details is added to the Annex 13).

For each laboratory, the submitted raw reads (FASTQ files) were evaluated by the EQA provider's in-house quality control pipeline [15]. For the full QC evaluation of all isolates, see Annex 14.

According to the QC parameters, sequencing quality was uniformly good. Coverage was sufficient expect for one isolate submitted by laboratory 133 (ID 9357) which was discarded from the analyses in this report.

Table 10. Results of raw reads submitted by participants evaluated by EQA provider QC pipeline summarised by laboratory

		Laboratory number											
Parameters	Ranges*	19	34	80	100	108	123	133	134	135	137	139	222
Detected species	{Ec} or {Sf}	Ec	Ec/ Sf	Ec	Ec	Ec	Ec/ Sf	Ec	Ec	Ec/ Sf	Ec	Ec	Ec
Species 1 (%)		83.1-92.8	56.2-98.2	76.8-94.9	87.7-96.1	90.8-96.0	83.0-95.6	74.7-93.9	86.8-96.2	82.0-95.4	83.0-92.4	85.4-92.4	89.6-96.9
Species 2 (%)		0.9-2.7	0.1-13.0	0.8-3.4	0.8-2.3	0.9-1.6	0.7-7.7	0.9-3.2	0.8-2.8	0.7-8.8	1.0-2.5	1.0-2.5	0.8-1.5
Unclassified reads (%)	{<100}	5.1-12.1	1.5-41.0	1.9-18.9	1.8-7.9	1.9-5.6	2.2-13.3	4.0-20.7	1.8-8.8	2.1-14.0	5.9-12.9	5.9-12.9	0.8-6.6
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.0-5.4	5.1-5.5	4.7-5.4	5.1-5.5	4.5-5.0	5.1-5.5	0.9-5.5	5.1-5.5	5.1-5.4	4.9-5.2	5.0-5.3	5.1-5.4
Length [0–25] x min. coverage (kbp)	{<250}	2.7-18.7	0.0-15.5	0.0-673.4	0.0-99.9	0.0-20.0	0.0-1.2	0.0-3843.4	0.0-2.1	9.9-62.4	0.0-251.4	0.0-126.8.0	0.0-2.8
Number of contigs at 25 x min. coverage	{>0}	341-534	158-290	124-211	149-225	2417-5182	210-316	185-553	163-247	153-247	829-1008	194-272	343-2122
Number of contigs [0–25] x min. coverage	{<1000}	5-32	0-14	0-74	0-14	0-62	0-3	0-3130	0-3	6-38	0-236	0-54	0-12
Average coverage	{>50}	88-168	40-88	37-130	46-145	25,5-95,5	85-136	21-135	36-96	77-108	36-204	36-205	52-266,5
Number of reads (x1000)		3332-6194	973-2240	741-3094	1087-3397	490-1752	1775-2982	774-3857	1267-3397	2795-3949	2047-10851	2047-10852	963,5-4602
Average read length		134-141	204-243	252-288	229-243	272-294	232-259	153-238	151-151	142-148	98.5-99.5	99	290-326
Average insert size		211-246	265-359	286-418	284-372	NA	255-313	152-269	298-380	342-418	302-355	303-357	NA
N50 (kbp)		30-48	54-166	86-135	86-140	1-3	73-116	2-127	88-140	58-135	18-26	54-109	4-40

Quality assessment made by the EQA-provider in-house quality control pipeline.
**: indicative QC range*
Ec: E. coli, *Sf:* Shigella flexneri *(listed if >5%), NA: Not Applicable*

4. Discussion

4.1 Serotyping

Twenty (83%) laboratories participated in the serotyping part of the EQA-9, of which nine participants (45%) provided phenotypic serotyping results and 11 (55%) provided molecular serotyping results (one by PCR and ten by WGS). This was an increase from six to ten participants using WGS based serotyping compared with EQA-8.

4.1.1 O group

A decrease in participation in O group typing was observed in this EQA-9. From EQA-4 through EQA-8, 93% of participants performed O grouping (26/28; 26/29; 26/29; 27/30) whereas 20 out of 24 participants performed it in EQA-9.

The performance of O grouping was higher in EQA-9 than in EQA-8. Ten participants (50%) reported the correct O group for all 12 test isolates compared with only 35% in EQA-8. O group O187 is not common in Europe and was reported as O74 by two laboratories and O121 by one other laboratories. O187 has strong cross-reactions with both O74, O103 and O175. The incorrect O grouping of O187 as O74 can therefore be explained for two of the three incorrect reported results as they used phenotypic methodology. False reactions against O74 should have been removed by absorption. One laboratory reported O121 instead of O187 by the use of PCR, and it is of note that the published PCR method for O grouping will not detect O187 [16]). The performance of this isolate increased from 35% (EQA-8) to 65% (EQA-9). Additional incorrect O groups: O154 as O127 (using PCR) or O126 as O157 (using WGS) have not been seen before.

Some of the more common O groups, also included in the minimum requirements of ECDC, generated the highest performances (O157: 100%, O121 and O145: 95%). Again this year the average score was slightly higher in EQA-9 (85%) compared with EQA-8 (79%) and EQA-7 (67%). There was also a shift observed (from 26% to 50%) from phenotypic serotyping towards WGS based analysis.

4.1.2 H type

A reduction in H typing participation from 18 laboratories in EQA-4 to 13 in EQA-9 was also detected. The general performance for H typing was higher than O grouping, but fewer performed H typing. The majority of participants (69%, 9/13) correctly reported the H type of all 12 test isolates (Figure 5). One laboratory reported 70% (7/10) of the errors. Compared with the previous EQA, the average score of 94% correct results was an improvement to the previous EQA-8 (92%) and 81% in EQA-7.

4.1.2 OH serotyping

Thirteen participants performed both O group and H typing of which nine (69%) correctly serotyped all 12 test isolates. The O:H serotyping results ranged from 100% for isolates O157:H7 to only 85% (11/13) for O187:H28. The average percentage of correct O:H serotyping in the EQA-9 was higher (92%) compared with EQA-8 (86%), EQA-7 (71%), and EQA-6 (78%). In general, the less common European serotypes generated the lowest scores and vice versa.
In addition to O grouping, H typing is crucial for outbreak detection, epidemiological surveillance, taxonomic differentiation of *E. coli* and detection of pathogenic serotypes. As such, it remains a main challenge to enable more PH NRLs to perform complete and reliable O:H serotyping, particular H typing, however with the use of WGS, this might be more feasible for some countries in the future.

4.2 Virulence profile determination

Twenty-three laboratories participated in the detection of the virulence profile with the participation rate and performance varying substantially between the different tests. As in previous EQAs, the participation rate was highest for the genotypic detection of the *stx* genes (96%) and detection of *eae* (92%) and lowest for the detection of *aaiC/aggR* (75-79%) and subtyping of stx genes (79%).

4.2.1 Detection of *aaiC* and *aggR*

The performance of detection of the two EAEC genes was high, with 89% and 95% of the participants respectively detecting *aaiC* and/or *aggR* correctly. In this EQA, only one EAEC isolate were included and one laboratory (138) did not report any positive EAEC isolates among the test isolates. The performance of detecting *aaiC* / *aggR* in EAEC isolates has been high through the four EQAs including an EAEC isolate (EQAs-4 to -9).

4.2.2 Detection of *eae*

Genotyping of *eae* had a high participation rate (92%) and performance; 19 (86%) laboratories obtained a 100% score, giving an average score of 99%. The average correct score has been fairly unchanged through the EQAs (EQA-4, 96%; EQA-5, 98%; EQA-6, 97%; EQA-7, 98%, EQA-8, 99%).

4.2.3 Detection of *stx1* and *stx2*

Both the participation (96%) and performance rates were high for genotyping of *stx1* (100%) and *stx2* genes (99%), similar to previous EQAs.

4.2.4 Subtyping of *stx1* and *stx2*

The average scores of correct subtyping of *stx1* and *stx2* were 93% and 92% respectively, which is an increase compared with EQA-8 (84% and 87%) but also a slightly higher performance compared with previous EQAs (EQA-4, 90%; EQA-5, 92%; EQA-6, 91%; EQA-7, 90%). The unexpected reporting[i] of 'not done' results which was an issue in EQA-8 was only reported by one laboratory in the EQA-9. The EQA-provider specified in the invitation letter and in the submission protocol of EQA-9 that when a participant signs up for a test and subsequently participates, all isolates must be tested using this test.

In the current EQA, the true errors ('not done' results excluded) were seven incorrect *stx1* subtyping results. All errors were reported by two laboratories, reporting 2 x *stx1* genes in one isolate multiple times. No isolates have ever been described with two copies of the *stx1* genes.

Of the true errors of incorrect *stx2* (n=15), REF4 and REF9 were reported incorrectly three times. The serotyping of REF4 (*stx2a*) is complicated. Using mapping, it could be mistaken for either an isolate with only *stx2c* or an isolate with both *stx2a* and *stx2c*. Translation of the B-subunit into amino acid showed that the sequence contained an 'EDD' (related to *stx2a*) motif instead of the 'END' (related to *stx2c*) [7]. REF9 (*stx2c*) was incorrectly reported as *stx2c* and *stx2d* twice or *stx2a* and *stx2c* once. The suggested protocol using PCR for subtyping [7] would have provide the correct result.

4.3 Molecular typing-based cluster analysis

Since EQA-8 the EQA scheme of STEC no longer covers PFGE as an independent part, but contains a cluster identification using either PFGE and/or WGS-derived data. Twelve laboratories participated in cluster identification using WGS-derived data, eight participated using PFGE-derived data and three of the 17 laboratories participated in cluster identification using both methods. Four out of the five laboratories that previously participated in PFGE (EQA-7), but did not participate in cluster identification in EQA-8, were still not participating in the cluster identification part of the EQA. The present cluster designed by the EQA provider allowed the participants to detect the same number of closely related isolates by both PFGE and WGS.

4.3.1 PFGE-derived data

Of the 24 laboratories, eight (33%) performed cluster analysis using PFGE-derived data, three of them also performed WGS analyses. Out of eight participants, five (63%) correctly identified the cluster of closely related isolates.

Compared with PFGE analysis of *Salmonella* or *Listeria*, the PFGE profile of *E. coli* contains a large number of bands within the region of 200–350 kb, which makes cluster analysis based on PFGE harder to interpret. The PFGE gel needs to be of a very good quality in order to correctly assign all bands in this region. One of the three laboratories which did not identify the correct cluster, missed one of the technical triplicates (REF16) and included additional four isolates in their cluster accepting more than two band difference. The second laboratory only included one additional isolate (with on band difference). The third laboratory incorrectly included nine of twelve isolates in their cluster.

This performance is comparable with EQA-8, where 60% correctly identified the cluster using PFGE. It also highlights the challenge of using PFGE for inter-laboratory comparisons and shows PFGE can be a problematic method for cluster analysis of STEC.

4.3.2 WGS-derived data

Twelve of 24 laboratories (50%) performed cluster analysis using WGS-derived data. Only one reported the use of external assistance for sequencing and the majority (10/12) reported using an Illumina platform. All reported using commercial kits for preparing the library.

[i] In EQA-8, some laboratories wrongly omitted performing the subtyping test on isolates already found negative in the initial screening (detection for stx1 and stx2).

Performance was very high, with 11 (92%) correctly identifying the cluster of closely related isolates, when accepting laboratory 108 comment regarding REF13 as being part of the cluster, but not listing the ID in the cluster field. Performance was comparable (91%) to EQA-8. Out of 12 laboratories, ten (83%) reported using an allele-based method as the main analysis and two (17%) reported using SNP analysis. Compared with EQA-8 this is an increase in the use of allele-based analysis, where 73% reported using an allele-based method for the main analysis and 27% reported using SNP analysis.

The one laboratory that did not identify the correct cluster used allele based analysis. If only evaluating the main analysis of the laboratories reporting the correct cluster, the distances reported using SNP-based analyses were 0–4 inside the cluster and the number of allele differences using cgMLST were 0–5 inside the cluster.

The two approaches to analyse WGS-derived data (allele- and SNP-based analysis) showed comparable results. One exception was the results from the laboratory not identifying the correct cluster using allelic based analysis. A higher number of allelic difference were observed by this laboratory (222) when using their schema of a reduced number of loci (2360) compared with the Enterobase scheme (2513 loci). Laboratory 222 reported 5, 9 and 23 allelic differences for their three triplicated isolates to the selected cluster representative (REF24), and they only reported isolates with allelic differences of nine and below to be a part of the cluster, missing REF16. However, the reasoning based on the data is understandable and led to setting a 'cut-off' between 9 and 23 instead of between 23 and 37 allelic difference.

When testing the submitted raw reads from laboratory 222 (IonTorret data), in SNP pipeline (NASP [13]) and in BioNumerics (Enterobase) scheme, the correct cluster were easily identified in both, and no difference were observed between the technical triplicates. As troubleshooting, the participants were asked to test EQA-provider sets of raw reads in their analysing tool and the result were similar to the EQA provider's analysis. This led to the conclusion that the analysing tool, despite the lower number of loci, worked on illumine data. The most likely explanation is that the ChewBBACA pipeline only uses assembly based mapping, and the Ion Torrent data is not assembled correctly leading to the many incorrect allele calls. In addition, the same laboratory (222) included one isolate too many in the cluster based on their PFGE results.

High similarity was seen for the reported cgMLST results based on Enterobase (0–2/3 allele differences within the cluster). The laboratory (139) using Enterobase scheme in Enterobase reported five allelic difference within the cluster, a slight increase compared with other laboratories using the same scheme but Hierarchical clustering in Enterobase. Particularly for the isolates with different ST´s (REF15, REF17 and REF18), the allelic differences were highest when using Enterobase (150-750) or Enterobase scheme in a different programme (200-1100). The use of a standard scheme does give a more comparable result.

Non-standardised SNP analyses can provide valid cluster detection at the national level and can be used for communication about cluster definitions, however, few laboratories performed SNP analysis in the EQA-9. Recombination had no importance for cluster identification, however some discrepancies were observed depending on which software was used. Highest variations in the SNP distances were outside the cluster, depending on the SNP pipeline and additional analysis performed.

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

In order to compare the quality of the raw data, the EQA provider analysed the submitted raw reads to obtain selected QC parameters. All laboratories submitted sequences of fine quality, only one laboratory submitted one sequence which were discarded for the EQA provider analysis. The EQA provider's analysis of the submitted raw data showed that when using a standardised cgMLST analysis, it is not unlikely to observe a random variation of one allele, even with high coverage (Figure 10). However, one participant (108) deviated systematically. This is likely due to a combination of sequencing technology and allele-calling software.

5. Conclusions

Twenty-four laboratories participated in the EQA-9 scheme, with 20 (83%) performing the serotyping part, 23 (96%) the virulence profile determination part and 17 (71%) the cluster identification part. As in EQA-8, EQA-9 contained the inclusion of molecular typing-based cluster analysis using either PFGE and/or WGS-derived data. This adjustment of the EQA seemed to be well accepted by most Member States, thus not all laboratories performing PFGE (EQA-7) signed up for molecular typing-based cluster analysis.

A decrease in the number of participants was seen compared with previous years for both serotyping and virulence profile determination. Laboratories reporting phenotypic serotyping decreased from 57% to 45% and the laboratories reporting WGS-based serotyping increased from 26% to 50% in the EQA-9 compared to the EQA-8.

The full O:H serotyping was only performed by 54% (13/24) of the participants, with an average score of 92%. As in previous EQAs, participation in the O grouping was higher than in H typing. Notably, not all laboratories demonstrated the capacity to determine all O groups and H types. In general, the more common European serotypes generated the highest scores. Serotype O187:H28 generated the lowest scores, correctly reported by eleven laboratories.

The participation and performance of virulence profile determination has been high through the EQAs. EQA-9 demonstrated a high performance for *aaiC* and *aggR,* with 89% and 95% average scores, respectively. Similarly the detection of *eae, stx1* and *stx2* had high participation rates and average scores above 99%. Subtyping of *stx1* and *stx2* is highly valuable since specific subtypes have been associated with HUS. The participation rate of 79% is therefore encouraging. The average score for subtyping of *stx1* and *stx2* was 93% and *stx2* 92%, respectively, and the combined average score of 89%. The incorrect results were mainly due to reporting two subtypes in the same isolate.

Incorporating the molecular typing-based cluster analysis in the EQA is up to date with the development of surveillance methods used by PH NRLs in Europe. Twelve laboratories performed cluster analysis using WGS-derived data, which was one laboratory more than in EQA-8. Performance was high, with 11 (92%) of participants correctly identifying the cluster of five closely related isolates. Five participants in the cluster analysis used only PFGE and two (40%) did not identify the correct cluster. This is highly related to the difficulties to interpret of the region of large number of bands in STEC PFGE profiles. The higher performance among WGS participants compared with PFGE emphasises the advantage of using WGS instead of PFGE for cluster analysis of STEC.

An allele-based method was preferred in the cluster analysis, as most laboratories (10/12; 83%) used cgMLST compared to 17% (2/12) using SNP analyses. The use of a standard cgMLST scheme (e.g. Enterobase) gives a very high degree of homogeneity in the results, and allele-based methods seem to be useful for inter-laboratory comparability and communication about cluster definitions. SNP analyses can also provide valid cluster detection at the national level and can be used for communication about cluster definitions, however, few laboratories performed SNP analysis in the EQA-9.

The current EQA scheme for typing of STEC is the ninth organised for laboratories in FWD-Net. The molecular surveillance system implemented as part of TESSy relies on the capacity of FWD-Net laboratories to produce analysable and comparable typing results in a central database. WGS-based typing for surveillance is increasingly used in EU.

6. Recommendations

6.1 Laboratories

Participants are encouraged to assign sufficient resources to repeat failed analysis if required to meet the deadline of submission.

Laboratories are expected to employ each method as an individual test irrespective of results obtained in the screening and detection or any other test. Therefore, when a participant signs up for a test and subsequently participates, all isolates must be tested using this test, e.g. subtyping of *stx*.

6.2 ECDC and FWD-Net

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

ECDC is encouraging more participants to take part in the new molecular typing-based cluster analysis, also participants who have not previously participated in the PFGE gel analysis part.

6.3 EQA provider

In the coming EQA round the EQA provider will evaluate the possibility to modify the cluster analysis to mimic a more realistic microbiological investigation by including genome sequences for the WGS analysis. This part is designed to be a simulation of an outbreak situation in a country to detect genetically closely related isolates and to compare the original cluster with genomes produced in other laboratories, which might be using different procedures and equipment.

7. References

1. European Parliament and European Council. Regulation (EC) No 851/2004 of the European Parliament and of the Council of 21 April 2004 establishing a European centre for disease prevention and control – Article 5.3. Strasbourg: European Parliament and European Council; 2004. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32004R0851.
2. European Parliament and European Council. Decision No 1082/2013/EU of the European Parliament and of the Council of 22 October 2013 on serious cross-border threats to health and repealing Decision No 2119/98/EC (Text with EEA relevance). Strasbourg: European Parliament and European Council; 2013. Available from: http://publications.europa.eu/en/publication-detail/-/publication/8d817a1f-45fa-11e3-ae03-01aa75ed71a1.
3. European Food Safety Authority and European Centre for Disease Prevention and Control. The European Union One Health 2018 Zoonoses Report EFSA Journal. 2019; 17(12):5926. Available from: https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/j.efsa.2019.5926
4. Bielaszewska M, Friedrich AW, Aldick T, Schürk-Bulgrin R, Karch H. Shiga Toxin Activatable by Intestinal Mucus in *Escherichia coli* Isolated from Humans: Predictor for a Severe Clinical Outcome. Clin Infect Dis. 2006 Nov 1;43(9):1160-7.
5. Friedrich AW, Bielaszewska M, Zhang WL, Pulz M, Kuczius T, Ammon A, et al. *Escherichia coli* Harboring Shiga Toxin 2 Gene Variants: Frequency and Association with Clinical Symptoms. J Infect Dis. 2002 Jan 1;185(1):74-84.
6. Persson S, Olsen KE, Ethelberg S, Scheutz F. Subtyping Method for *Escherichia coli* Shiga Toxin (Verocytotoxin) 2 Variants and Correlations to Clinical Manifestations. J Clin Microbiol. 2007 Jun;45(6):2020-4.
7. Scheutz F, Teel LD, Beutin L, Piérard D, Buvens G, Karch H, et al. Multicenter Evaluation of a Sequence-Based Protocol for Subtyping Shiga Toxins and Standardizing Stx Nomenclature. J Clin Microbiol. 2012 Sep;50(9):2951-63.
8. International Organization for Standardization. ISO/IEC 17043:2010 – Conformity assessment -- General requirements for proficiency testing. Vernier: ISO; 2010. Available from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=29366.
9. Scheutz F, Fruth A, Cheasty T, Tschäpe H. Appendix 1 – O Grouping: Standard Operation Procedure (O SOP) and Appendix 2: and H Determination: Standard Operation Procedure (H SOP) – *Escherichia coli* O antigen grouping and H antigen determination. Copenhagen: Statens Serum Institut; 2002. Available from: http://www.ssi.dk/English/HealthdataandICT/National%20Reference%20Laboratories/Bacteria/~/media/49802860CB5E44D6A373E6116ABBDC0D.ashx.
10. Scheutz F, Morabito S, Tozzoli R, Caprioli A. Identification of three vtx1 and seven vtx2 subtypes of verocytotoxin encoding genes of *Escherichia coli* by conventional PCR amplification. Copenhagen: Statens Serum Institut; 2002.
11. PulseNet International. Standard operating procedure for PulseNet PFGE of *Escherichia coli* O157:H7, *Salmonella* serotypes, *Shigella sonnei*, and *Shigella flexneri*. Atlanta: PulseNet International; 2013. Available at: http://www.pulsenetinternational.org/assets/PulseNet/uploads/pfge/PNL05_Ec-Sal-ShigPFGEprotocol.pdf.
12. Warwick Medical School. EnteroBase [Internet]. Coventry: University of Warwick; 2018 [cited 21 August 2018]. Available from: http://enterobase.warwick.ac.uk.
13. 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.
14. European Centre for Disease Prevention and Control. Surveillance of National Reference Laboratory (NRL) capacity for six food- and waterborne diseases in EU/EEA countries – Campylobacteriosis, listeriosis, salmonellosis, Shiga toxin/ verocytotoxin–producing *Escherichia coli* (STEC/VTEC), shigellosis and yersiniosis. Stockholm: ECDC; 2012. Available from: http://ecdc.europa.eu/publications-data/survey-national-reference-laboratory-capacity-six-fwd-eueea-countries.
15. Statens Serum Institut. Bifrost_QC [Internet; software package]. Copenhagen: Statens Serum Institut; 2019. Available from: https://github.com/ssi-dk/bifrost.
16. Iguchi A, Iyoda S, Seto K, Morita-Ishihara T, Scheutz F, Ohnishi M, et al. *Escherichia coli* O-Genotyping PCR; a Comprehensive and Practical Platform for Molecular O Serogrouping. J.Clin.Microbiol. 2015 Aug;53(8):2427-32.

Annex 1. List of participants

Country	Laboratory	National institute
Austria	Nationale Referenzzentrale für Escherichia coli einschließlich Verotoxin bildender E. coli	Institut für Klinische Mikrobiologie und Hygiene (AGES)
Belgium	NRC STEC	Universitair Ziekenhuis Brussel
Czech Republic	NRL for E.coli and Shigella	National Institute of Public Health
Denmark	Foodborne Infections	Statens Serum Institut
Estonia	Laboratory of Communicable Diseases	Health Board
Finland	Expert Microbiology Unit	National Institute for Health and Welfare
France	CNR associé E.coli - Service de Microbiologie	Hopital Robert Debré
Germany	NRC for Salmonella	Robert Koch Institute
Greece	National Reference Centre for Salmonella, Shigella, VTEC	National School of Public Health
Iceland	Department of Clinical Microbiology	Landspítali University Hospital
Ireland	VTEC National Reference Laboratory	Public Health Laboratory – Health Service Executive
Italy	Microbiological Food Safety and Foodborne Disease Unit	Istituto Superiore di Sanità
Latvia	Infectology Centre of Latvia, National Microbiology Reference Laboratory	Riga East University Hospital
Lithuania	National Public Health Surveillance Laboratory	Budget organisation
Luxembourg	Epidemiology and Microbial Genomics	Laboratoire National de Sante
Macedonia		Food Institute
Norway	National Reference Laboratory for Enteropathogenic Bacteria	Norwegian Institute of Public Health
Poland	Department of Bacteriology	National Institute of Public Health – National Institute of Hygiene
Portugal	LNR Infeções Gastrintestinais	Instituto Nacional de Saúde Doutor Ricardo Jorge
Romania	Molecular Epidemiology Laboratory	Cantacuzino National Medico-Military Institute of Research and development
Slovenia	National Laboratory of Health, Environment and Food	Centre for Medical Microbiology
Sweden	Microbiology	Folkhälsomyndigheten (The Public Health Agency of Sweden)
The Netherlands	Department of Bacterial Surveillance and Response	National Institute for Public Health and the Environment
United Kingdom	Gastrointestinal Bacteria Reference Unit	Public Health England

Annex 2. Participation overview EQA-8/EQA-9

	2017-2018 (EQA-8)					2018-2019 (EQA-9)				
	All	Serotyping	Virulence	PFGE	WGS	All	Serotyping	Virulence	PFGE	WGS
Laboratory number				Cluster					Cluster	
19	x	x	x	x	x	x	x	x	x	x
34	x	x	x		x	x	x	x		x
80	x	x	x		x	x	x	x		x
88	x		x			x		x		
90	x		x	x		x		x	x	
94										
100	x	x	x	x		x	x	x		x
108	x	x	x		x	x	x	x		x
114										
123	x	x	x	x	x	x	x	x	x	x
124	x	x	x	x		x	x	x	x	
125										
126										
127	x	x	x	x		x	x	x	x	
128	x	x	x			x	x	x		
129	x	x	x		x	x	x	x		
130	x	x	x			x			x	
131	x	x	x							
132	x	x	x	x		x	x	x	x	
133	x	x	x		x	x	x	x		x
134	x	x	x		x	x	x	x		x
135	x	x	x		x	x	x	x		x
136	x	x	x	x						
137	x	x	x		x	x	x	x		x
138	x	x	x			x	x	x		
139	x	x	x		x	x	x	x		x
145	x	x	x			x	x	x		
153	x	x	x			x	x	x		
180						x		x		
222	x	x	x	x		x	x	x	x	x
Number of participants	**25**	**23**	**25**	**9**	**11**	**24**	**20**	**23**	**8**	**12**

Annex 3. Reason(s) for participating in EQA

Part	Reason(s)	Laboratory number																											
		19	34	80	88	90	#	108	114#	123	124	127	128	129	130	132	133	134	135	136	137	138	139	145	153	180	222	No.
Serotyping	Accreditation needs	x	x					x	x	x	x	x	x		x		x		x	x	x			x	x		x	16
	Institute policy		x	x		*e*	x	x		x				x	x					x	x			x		*a*		10
	National policy		x						x	x	x			x	x		x				x	x		x				10
	Enhance the typing quality		x				x		x		x	x	x	x	x	x	x	x	x		x		x	x			x	16
Virulence profile (Stx1 + Stx2)	Accreditation needs	x	x		x			x	x	x	x	x	x		x		x		x	x	x			x			x	16
	Institute policy		x	x			x	x		x				x	x		x			x	x			x				11
	National policy		x						x	x	x			x	x		x				x	x		x				10
	Enhance the typing quality		x			x	x		x		x	x	x	x	x	x	x	x	x		x		x	x	x	x	x	19
Virulence profile (eae)	Accreditation needs	x	x		x			x	x	x	x	x	x		x		x		x	x	x			x			x	16
	Institute policy		x	x			x	x		x				x	x		x			x	x			x				11
	National policy		x						x	x	x			x	x		x				x	x		x				10
	Enhance the typing quality		x			x	x		x		x	x	x	x	x	x	x	x	x		x		x	x	x	x	x	19
Virulence profile (*aaiC/aggR*)	Accreditation needs	x	x		x					x	x		*a*		*b* + *c*	*a*	x			x	x	*d*		x				9
	Institute policy		x	x			x	x		x							x			x	x			x				9
	National policy		x						x	x	x						x				x			x				7
	Enhance the typing quality		x			x	x		x		x	x		x			x	x			x		x	x	x	x	x	15
Virulence profile (subtyping)	Accreditation needs	x	x		x					x	*x*		*a*		*b* + *c*	*a*	x		x	x	x			x	*b* + *d*			10
	Institute policy		x	x			x	x	x	x				x			x			x	x			x				11
	National policy		x						x	x	x			x			x				x	x		x				9
	Enhance the typing quality		x			x	x		x		x	x		x			x	x	x		x		x	x		x	x	15
Cluster identification	Accreditation needs	x	x						x	x	x				x		x		x	x	x			x				10
	Institute policy		x	x			x	x		x			*a*	*a* + *b*	x		x			x	x			x	*b* + *d*	*a*		10
	National policy		x						x	x	x	x			x		x				x			x				9
	Enhance the typing quality		x			x	x		x		x	x			x	x	x	x	x		x		x	x			x	15

Reasons for NOT participating in the EQA (in gray)
a Lack of laboratory capacity
b Lack of financial means
c No national surveillance of STEC
d Method not relevant to our laboratory
e The laboratory is able to serotype only a limited number of serogroups.

Annex 4. Serotyping result scores

O group

		Laboratory number																			
Isolate	EQA	19	34	80	100	108	123	124	127	128	129	132	133	134	135	137	138	139	145	153	222
REF1	O55	55	55	55	55	55	55	55	55	55	55	ND	55	55	55	55	55	55	55	55	55
REF2	O187	187	187	187	74	187	187	187	74	ND	187	ND	187	187	187	187	121	187	NT	NT	187
REF3	O128	128	128	128	128	128	128	128	128	128	128	ND	128	128	128	128	128	128	128	128	128
REF4	O76	76	76	76	76	76	76	NT	76	ND	76	ND	76	76	76	76	NT	76	NT	NT	76
REF5	O154	154	154	154	154	154	154	154	154	127	154	ND	154	154	154	154	127	154	NT	NT	154
REF6	O91	91	91	91	91	91	91	91	91	91	91	ND	NT	91	91	91	91	91	NT	NT	91
REF7	O121	121	121	121	121	121	121	121	121	121	121	ND	121	121	121	121	121	121	121	121	121
REF8	O111	111	111	111	111	111	111	111	111	ND	111	ND	111	111	111	111	111	111	111	111	111
REF9	O157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157	157
REF10	O145	145	145	145	145	145	145	145	145	145	145	ND	145	145	145	145	145	145	145	145	145
REF11	O126	126	126	157	126	126	126	126	126	126	126	ND	126	126	126	126	NT	126	126	126	126
REF12	O91	91	91	91	91	91	91	91	91	91	91	ND	91	91	91	91	91	91	NT	NT	91
Method	**A**	**A**	**C**	**C**	**A**	**C**	**C**	**A**	**A**	**A**	**C**	**A**	**C**	**C**	**C**	**C**	**B**	**C**	**A**	**A**	**A**

n=20 participants
Purple shading: incorrect result
A: phenotypic serotyping, B: PCR-based serotyping, C: WGS-based serotyping
NT: non-typable
ND: not done.

H type

		Laboratory number												
Isolate	EQA	19	34	80	108	123	124	127	129	133	134	135	137	139
REF1	H7	7	7	7	7	7	NT	7	7	7	7	7	7	7
REF2	H28	28	28	28	28	28	NT	28	28	28	28	28	28	28
REF3	H2	2	2	2	2	2	NT	2	2	2	2	2	2	2
REF4	H-/H7	H-	7	7	7	7	H-	7	7	7	7	7	7	7
REF5	H31	31	31	31	31	31	H-	31	31	H-	31	31	31	31
REF6	H14	14	14	14	14	14	H-	14	14	14	14	14	14	14
REF7	H19	19	19	19	19	19	NT	19	19	19	19	19	19	19
REF8	H-/H8	8	8	8	8	8	H-	8	8	8	8	8	8	8
REF9	H-/H7	7	7	7	7	7	H-	7	7	7	7	7	7	7
REF10	H-/H28	H-	28	28	28	28	H-	14	28	28	28	28	28	28
REF11	H-/H27	H-	27	7	27	27	H-	27	27	27	27	27	27	27
REF12	H21	21	21	21	21	21	H-	21	21	21	21	21	21	21
Method	**A**	**A**	**C**	**C**	**C**	**C**	**A**	**B**	**C**	**C**	**C**	**C**	**C**	**C**

n=13 participants
Purple shading: incorrect results
A: phenotypic serotyping, B: PCR-based serotyping, C: WGS-based serotyping
NT: non-typable
Some H- results was accepted as correct results (REF4, 8 -11), when the EQA provider observed a tendency to be H- more than one during testing.

Annex 5. Virulence profile determination result scores

Detection of *aaiC*

		Laboratory number																	
Isolate	EQA	19	80	88	90	100	108	123	124	127	129	133	134	137	138	139	145	153	222
REF1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	-	+	+
REF12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

n=18 participants
Purple shading: incorrect results

Detection of *aggR*

		Laboratory number																			
Isolate	EQA	19	34	80	88	90	100	108	123	124	127	129	133	134	137	138	139	145	153	222
REF1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+
REF12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

n=19 participants
Purple shading: incorrect results

Detection of *eae*

		Laboratory number																						
Isolate	EQA	19	34	80	88	90	100	108	123	124	127	128	129	132	133	134	135	137	138	139	145	153	222
REF1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
REF3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+
REF5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF6	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF11	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-
REF12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

n=22 participants
Purple shading: incorrect results.

Detection of *stx1*

		Laboratory number																					
Isolate	EQA	19	34	80	88	90	100	108	123	124	127	128	129	132	133	134	135	137	138	139	145	153	180	222
REF1	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF5	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
REF9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
REF12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

n=23 participants

Detection of *stx2*

		Laboratory number																						
Isolate	EQA	19	34	80	88	90	100	108	123	124	127	128	129	132	133	134	135	137	138	139	145	153	180	222
REF1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
REF2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
REF3	+	+	+	+	-	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
REF4	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
REF5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0
REF6	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	-	1
REF7	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
REF8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
REF9	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
REF10	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1
REF11	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	0
REF12	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1

n=23 participants.
Purple shading: incorrect result

stx subtyping

stx1

		Laboratory number																		
Isolate	EQA	19	34	80	88	90	100	108	123	124	127	129	133	134	137	138	139	145	180	222
REF1	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a 1c	1a	1a
REF2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-
REF3	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1c	1a 1c	1a 1c	1c
REF4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-
REF5	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1d	1a 1d	1d
REF6	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a 1c	1a	1a
REF7	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-
REF8	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a	1a 1c	1a 1c	1a
REF9	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	1a	-
REF10	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-
REF11	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-
REF12	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-

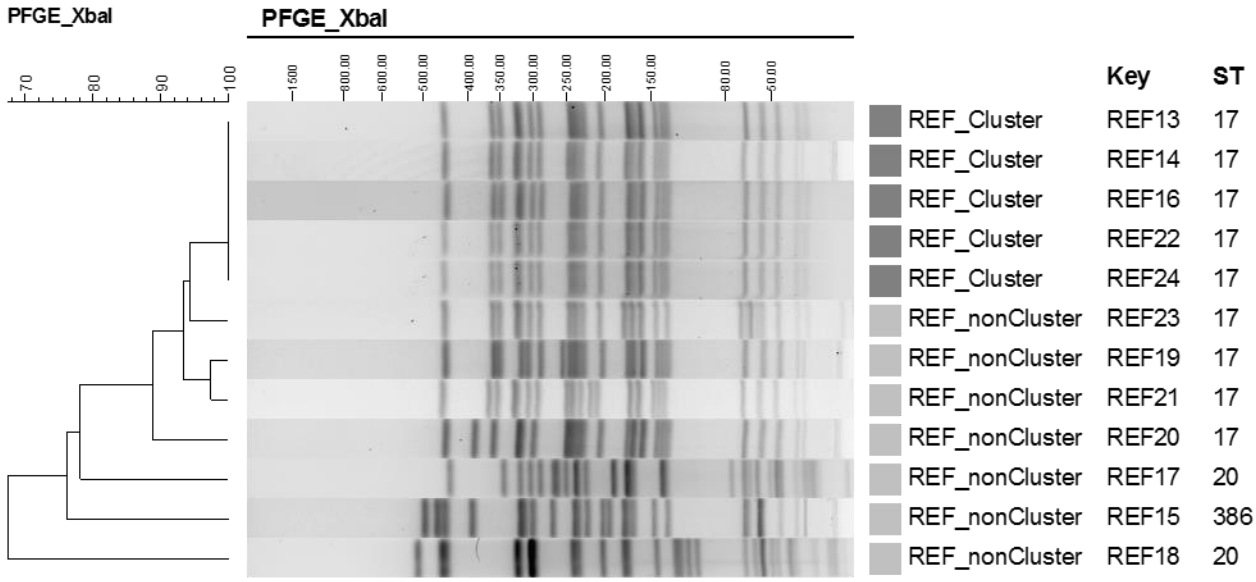
n=19 participants
Purple shading: incorrect result
ND: not done.

stx2

Isolate	EQA	19	34	80	88	90	100	108	123	124	127	129	133	134	137	138	139	145	180	222
REF1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-
REF2	2g	2g	2g	2g	2g	2g	2g	2g	2g	2g	2g	2g	2g	2g	2g	2a 2g	2g	2g	2g	2g
REF3	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2a 2b	2b	2b	2b	2b	2b	2b
REF4	2a	2a	2a	2a	2a	2a	2a	2c	2a	2a	2a	2a 2c	2a	2a	2a	2a	2a	2a 2c	2a	2a
REF5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	ND	-	-
REF6	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	2b	-	2b
REF7	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a 2c 2d	2a	2a
REF8	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a 2c 2d	2a	2a
REF9	2c	2c	2c	2c	2c 2d	2c 2d	2c	2c	2c	2c	2c	2c	2c	2c	2c	2c	2c	2a 2c	2c	2c
REF10	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a	2a 2c 2d	2a	2a
REF11	-	-	-	-	-	-	-	-	-	-	-	2a 2c	-	-	-	-	-	ND	-	-
REF12	2d	2d	2d	2d	2c 2d	2d	2d	2d	2d	2d	2d	2d	2d	2d	2d	2d	2d	2a 2c 2d	2d	2d

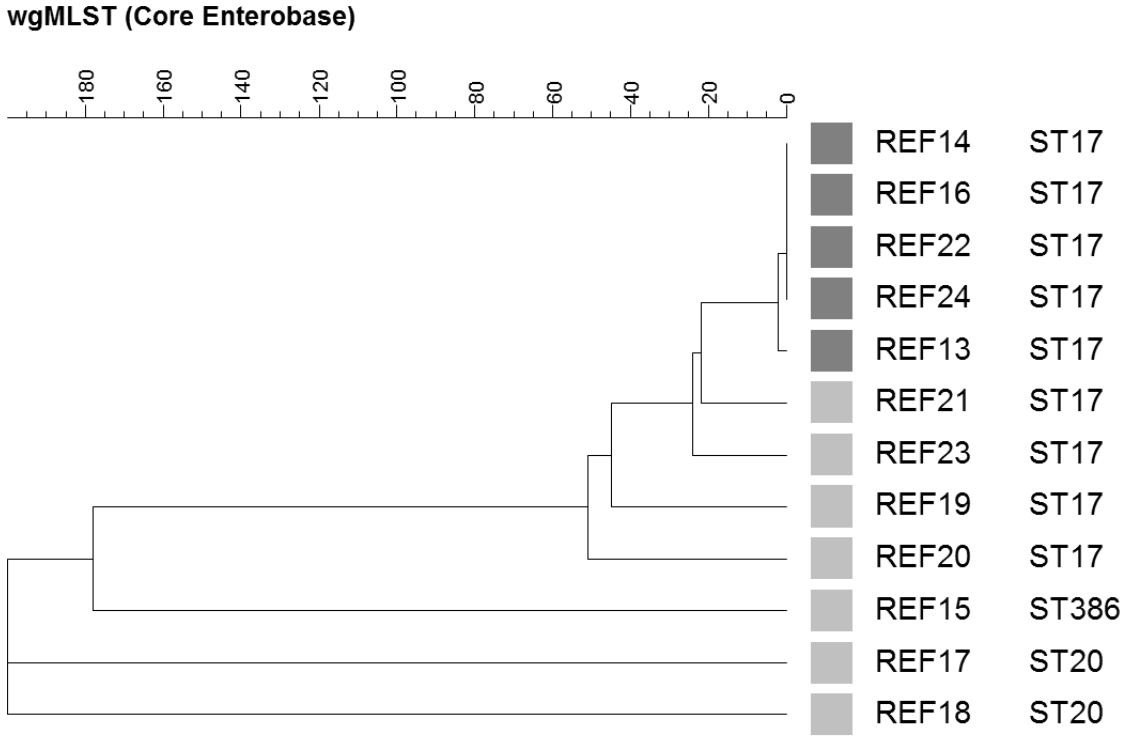
n=19 participants
Purple shading: incorrect results
ND: not done.

Annex 6. EQA provider cluster analysis based on PFGE-derived data



Cluster of closely related isolates (dark gray): O103:H2, stx1a
REF14, REF16 and REF22 are technical triplicates.

Annex 7. EQA provider cluster analysis based on WGS-derived data



Single linked dendrogram of core genome Multi Locus Sequence Typing (cgMLST) profiles of STEC EQA-9 isolates (cgMLST, EnteroBase, http://enterobase.warwick.ac.uk).
Analysed in BioNumerics: maximum distance of 200 exceeded, results clipped.
Cluster isolates: dark grey, outside cluster isolates: light grey.
REF14, REF16 and REF24 are technical triplicates.

Annex 8. Reported cluster of closely related isolates based on PFGE-derived data

Lab	Reported cluster	Corresponding REF isolates	Correct
		REF13, REF14, REF16, REF22, REF24	
19	9281, 9431, 9487, 9770, 9837	REF24, REF22, REF16, REF13, REF14	Yes
90	9098, 9369, 9390, 9647, 9717	REF22, REF13, REF24, REF16, REF14	Yes
123	9100, 9120, 9207, 9470, 9555	REF14, REF13, REF22, REF16, REF24	Yes
124	9038, 9211, 9886, 9950, 9999	REF14, REF16, REF13, REF24, REF22	Yes
127	9088, 9159, 9191, 9844, 9934	REF24, REF13, REF16, REF22, REF14	Yes
130	9171, 9624, 9062, 9347, 9383, 9052, 9067, 9060	REF19, REF24, REF20, REF22, REF21, REF14, REF13, REF23	No
132	9193, 9194, 9257#, 9290, 9345, 9525, 9759, 9896, 9900	REF24, REF14, REF13, REF21, REF20, REF23, REF16, REF22, REF19	No
222	9077, 9155, 9395, 9423, 9631, 9648	REF24, REF16, REF22, REF13, REF14, REF21	No

REF14, REF16 and REF24 are technical triplicates.

 # Incorrect ID: should have been 9275

Annex 9. Reported band differences

8			Laboratory number							
Isolate ID	ST	Total no of bands Expected *xbaI* bands	19	90	123	124	127	130	132	222
REF13‡	17	18/19*	19	19	17	20	16¤	16¤	17	15
REF14‡#	17	18/19*	19	19	17¤	20¤	16	16	9999¤	15
REF15	386	Clearly unrelated	20	19	19	19	18	16	17	16
REF16‡#	17	18/19*	19	19	17	20	16	0	17	15
REF17	20	Clearly unrelated	19	20	17	21	18	16	17	18
REF18	20	Clearly unrelated	19	21	17	19	18	15	17	17
REF19	17	19/20*	19	20	17	21	17	16	16	14
REF20	17	18/19*	19	19	16	19	15	16	16	14
REF21	17	19/20*	20	19	18	20	17	17	17	15
REF22‡#	17	18/19*	19	19¤	17	20	16	17	17	15
REF23	17	20	18	20	19	20	18	18	19	18
REF24‡	17	18/19*	19¤	19	17	20	16	15	17	15¤

8			Laboratory number							
Isolate ID	ST	Bands shared with *xba*I	19	90	123	124	127	130	132	222
REF13‡	17	18/19*	19	19	17	20	16¤	16¤	17	15
REF14‡#	17	18/19*	19	19	17¤	20¤	16	16	9999¤	15
REF15	386	Clearly unrelated	7	15	9999	10	6	12	11	11
REF16‡#	17	18/19*	19	19	17	20	16	0	17	15
REF17	20	Clearly unrelated	9	16	9999	13	6	9	9	13
REF18	20	Clearly unrelated	7	14	9999	10	5	11	10	11
REF19	17	14/15*	12	19	14	18	14	15	15	13
REF20	17	15/16*	15	17	9999	18	14	15	14	13
REF21	17	15/16*	18	18	9999	19	15	16	17	15
REF22‡#	17	18/19*	19	19	17¤	20	16	16	17	15
REF23	17	15	13	19	9999	17	14	16	16	15
REF24‡	17	18/19*	19¤	19	17	20	16	15	17	15¤

ST: sequence type
‡: cluster identification of closely related isolates (based on PFGE-derived data)
#: technical triplicates
¤: isolate used as cluster representative by participant
9999: not reported by laboratory
** one band is close to the 33kb therefore two results (for Total no. of bands and Shared* with ¤) are listed as expected.
Grey shading: cluster isolates

Annex 10. Reported sequencing details

Sequencing performed	Protocol (library prep)	Commercial kit	Sequencing platform
In own laboratory	Commercial kits	Illumina DNA flex	MiniSeq
In own laboratory	Commercial kits	Kapa Hyper Plus (Roche)	MiSeq
In own laboratory	Commercial kits	NexteraxT V3	MiSeq
In own laboratory	Commercial kits	Nextera xT	MiSeq
Externally	Commercial kits	Nextera xT	Illumina NovaSeq
In own laboratory	Commercial kits	Ion express TM Plus Fragment Library Kit	Ion Torrent S5xL
In own laboratory	Commercial kits	Nextera	HiSeq 2500
In own laboratory	Commercial kits	NEBNext® Fast DNA Fragmentation & Library Prep Set for Ion Torrent, New England Biolabs *	Ion GeneStudio S5 System
In own laboratory	Commercial kits	Nextera DNA Flex *	MiniSeq Illumina
In own laboratory	Commercial kits	NexteraxT	NextSeq
In own laboratory	Commercial kits	Nextera xT, Illumina *	MiSeq
In own laboratory	Commercial kits	Nextera DNA flex library prep Kit v2	MiSeq

**: adjusted volume of reagents.*

Annex 11. Reported cluster of closely related isolates based on WGS-derived data

Lab	Reported cluster	Corresponding to REF isolates	Correct
		REF13, REF14, REF16, REF22, REF24	
19	9487,9837,9281, 9431, 9770	REF16, REF14, REF24, REF22, REF13	Yes
34	9296, 9571, 9622, 9668, 9694	REF14, REF24, REF13, REF16, REF22	Yes
80	9200, 9466, 9471, 9729, 9907	REF14, REF22, REF24, REF13, REF16	Yes
100	9009, 9430, 9879, 9985, 9681	REF16, REF24, REF22, REF14, REF13	Yes
108	9201, 9108, 9987, 9608	REF14, REF16, REF22, REF24*	(Yes)
123	9100, 9120, 9207, 9470, 9555	REF14, REF13, REF22, REF16, REF24	Yes
133	9287, 9357, 9491, 9789, 9929	REF16, REF24, REF13, REF22, REF14	Yes
134	9346, 9582, 9855, 9991, 9804	REF14, REF24, REF16, REF22, REF13	Yes
135	9012, 9315, 9590, 9830, 9921	REF13, REF24, REF14, REF22, REF16	Yes
137	9278, 9735, 9782, 9818, 9819	REF16, REF24, REF14, REF13, REF22	Yes
139	9185, 9368, 9392, 9563, 9994	REF13, REF24, REF14, REF22, REF16	Yes
222	9077, 9395, 9423, 9631	REF24, REF22, REF13, REF14	No

REF14, REF16 and REF24 are technical triplicates.

**reported in the comments "9608 (REF24) and 9914 (REF13) are defined as part of the cluster but differs by 2 and 4 pure SNPs respectively".*

Annex 12. Reported SNP distance and allelic differences

SNP distances

			Laboratory number		
Isolate ID	ST	Provider	Provider (recombination – deleted)	108	137
REF13‡	17	3	3	4	4
REF14‡#	17	0¤	0¤	0¤	0
REF15	386	913	727	9999	510
REF16‡#	17	0	0	0	0¤
REF17	20	4665	2725	9999	1521
REF18	20	3604	2181	9999	970
REF19	17	131	125	153	130
REF20	17	127	130	236	160
REF21	17	50	50	61	59
REF22‡#	17	0	0	0	0
REF23	17	153	109	591	59
REF24‡	17	1	1	2	2

Allelic differences

					Laboratory number							
Isolate ID	ST	Provider	19	34	80	100	123	133	134	135	139	222
REF13‡	17	2	2	3	2	2	2	2	2	2	5	7
REF14‡#	17	0¤	0¤	0¤	0¤	0	0¤	2	0	0	0	5
REF15	386	179	179	247	177	268	9999	48	276	277	200	193
REF16‡#	17	0	0	0	0	0¤	0	0¤	0	0	0	23
REF17	20	557	553	750	547	631	9999	184	652	651	1100	573
REF18	20	379	386	522	374	461	9999	150	475	477	400	384
REF19	17	48	47	2815	49	48	49	20	49	49	50	62
REF20	17	54	52	75	53	98	9999	20	100	101	50	61
REF21	17	23	22	30	21	21	21	18	21	21	50	42
REF22‡#	17	0	0	0	0	0	0	1	0	0	0	9
REF23	17	25	24	38	29	29	29	11	29	29	50	37
REF24‡	17	0	0	1	0	0	0	1	0¤	0¤	0¤	0¤

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

Annex 13. Reported QC parameters

Lab no.	Criteria 1		Criteria 2		Criteria 3	
	Parameter	Threshold	Parameter	Threshold	Parameter	Threshold
19		Acceptable genome size at 1x and 10x coverage (4,6-5,6Mbp). Difference between (1x and 10x): <250000				
100	contamination for Salmonella Enteritidis (SEn)	BWA mapping with in-house SEn sequence 0 contigs assembled	contamination for Neisseria meningitidis (Nm)	BWA mapping with in-house Nm sequence 0 contigs assembled	SAV	cluster density, clusters passing filter and Q30 score were all according to Illumina recommendations
108	Contamination	genome size , MLST all genes 100%	Similarity, Length, CDS Covered	Not reported	-	-
123	N50	>50.000	%cgMLST targets	Not reported	-	-
133	No of contigs N50	<500 >60.000	read length	3.9-6.5mb	core %	0,96
137	Variant ratio Minimum depth coverage (SNP-typing)	≥ 0.9 ≥ 10x	Minimum consensus depth (Achtman 7 gene MLST)	> 0	Minimum metric yield (sequence quality)	≥ 150 Mb
139	N50 and number of contig	Not reported				
222	GC content	GC content expected at 50% along all the reads. Assessed thorugh FASTQC	Phred score	Minimum accepted Phred score = 25. This threshold was applied during trimming by using the tool	N50	>50000

Annex 14. Calculated qualitative/quantitative parameters

Quality Assessment made by the SSI in-house quality control pipeline https://github.com/ssi-dk/bifrost

		Laboratory 19											
Parameters	Ranges*	9134	9281	9367	9431	9487	9560	9632	9687	9770	9810	9837	9881
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		91.4	92.5	89.3	90.7	90.9	90.4	92.8	92.3	91.7	83.1	91.6	91.7
Species 2 (%)		0.9	1.3	2.1	1.3	1.4	1.4	1.4	1.4	1.3	2.7	1.3	1.3
Unclassified reads (%)	{<100}	5.8	5.5	6.6	7.4	7.1	7.5	5.1	5.6	6.1	12.1	6.4	6.4
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.1	5.1	5.3	5.1	5.1	5.1	5.0	5.3	5.1	5.4	5.1	5.2
Length [0–25] x min. coverage (kbp)	{<250}	5.7	9.6	15.5	8.5	18.7	15.0	2.7	6.1	3.2	7.4	5.5	14.1
Number of contigs at 25 x min. coverage	{>0}	341	456	534	471	436	515	415	381	351	451	410	417
Number of contigs [0–25] x min. coverage	{<1000}	9	16	29	14	32	31	6	10	5	14	11	22
Average coverage	{>50}	123	136	112	121	104	121	168	141	139	117	126	88
Number of reads (x1000)		4545	5073	4392	4753	4046	4721	6194	5507	5248	4677	4837	3332
Average read length		141	140	137	134	136	134	140	140	139	140	138	139
Average insert size		232	228	232	219	226	211	222	225	229	231	229	246
N50 (kbp)		44	38	32	36	34	30	43	48	48	35	41	37

		Laboratory 34										
Parameters	Ranges*	9296	9422	9476	9571	9622	9668	9679	9694	9741	9821	9927	9958
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec, Sf	Ec	Ec	Ec
Species 1 (%)		96.2	56.2	98.2	96.3	96.3	96.8	96.8	96.8	65.8	94.7	96.6	96.6
Species 2 (%)		0.9	1.4	0.1	1.0	1.1	1.0	1.3	1.0	13.0	1.4	0.9	1.0
Unclassified reads (%)	{<100}	2.5	41.0	1.5	2.3	2.3	1.9	1.5	1.9	11.2	2.8	2.2	2.1
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.2	5.5	5.3	5.2	5.2	5.2	5.4	5.2	5.3	5.1	5.2	5.2
Length [0–25] x min. coverage (kbp)	{<250}	0.0	15.5	0.0	0.0	0.0	0.0	1.3	0.0	8.3	0.4	5.1	0.0
Number of contigs at 25 x min. coverage	{>0}	214	280	212	210	210	213	204	206	290	158	219	171
Number of contigs [0–25] x min. coverage	{<1000}	0	14	0	0	0	0	1	0	8	1	5	0
Average coverage	{>50}	88	80	77	77	71	84	74	83	74	76	40	69
Number of reads (x1000)		2240	2016	1812	1935	1800	2022	1646	1966	1966	1845	973	1664
Average read length		205	219	224	209	206	218	243	222	204	214	218	219
Average insert size		275	312	328	272	265	287	359	309	270	283	317	308
N50 (kbp)		108	54	166	114	108	108	87	104	82	112	79	130

		Laboratory 80											
Parameters	Ranges*	9158	9200	9274	9285	9437	9466	9471	9497	9729	9898	9907	9944
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		94.7	94.0	94.9	92.3	94.1	93.1	89.7	94.0	94.5	93.7	94.5	76.8
Species 2 (%)		1.6	1.7	0.8	2.2	1.5	1.6	3.4	1.8	1.5	1.8	1.5	2.3
Unclassified reads (%)	{<100}	2.5	2.5	1.9	2.9	2.7	2.9	3.9	2.4	2.5	2.7	2.6	18.9
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.2	5.1	5.0	4.7	5.2	5.2	5.2	5.1	5.2	5.4	5.2	5.3
Length [0–25] x min. coverage (kbp)	{<250}	11.2	52.8	110.5	673.4	0.6	1.3	0.0	0.0	0.0	0.0	0.0	159.1
Number of contigs at 25 x min. coverage	{>0}	143	167	124	176	164	176	202	158	164	175	190	211
Number of contigs [0–25] x min. coverage	{<1000}	1	3	7	74	1	3	0	0	0	0	0	13
Average coverage	{>50}	72	69	52	37	82	102	130	89	83	91	90	72
Number of reads (x1000)		1398	1313	953	741	1613	2023	3094	1716	1594	1842	1848	1590
Average read length		275	284	288	275	275	272	252	274	279	275	263	257
Average insert size		386	411	418	381	378	369	286	379	391	382	340	332
N50 (kbp)		135	135	127	135	132	135	132	134	135	135	135	86

		Laboratory 100											
Parameters	Ranges*	9009	9066	9135	9160	9245	9355	9430	9507	9681	9879	9973	9985
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		96.1	95.4	96.0	93.9	95.3	87.7	95.2	95.4	95.8	95.1	95.8	94.7
Species 2 (%)		1.2	1.3	1.3	1.7	1.3	2.3	1.3	0.8	1.3	1.4	1.3	1.3
Unclassified reads (%)	{<100}	1.8	2.3	1.9	2.5	2.5	7.9	2.5	2.3	2.1	2.6	2.2	3.0
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.2	5.2	5.1	5.4	5.3	5.5	5.2	5.1	5.2	5.2	5.2	5.2
Length [0–25] x min. coverage (kbp)	{<250}	0.0	0.0	0.5	0.0	99.9	8.3	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25 x min. coverage	{>0}	194	177	180	225	188	217	191	149	192	210	177	204
Number of contigs [0–25] x min. coverage	{<1000}	0	0	1	0	14	6	0	0	0	0	0	0
Average coverage	{>50}	66	51	51	64	57	46	63	70	99	129	70	145
Number of reads (x1000)		1515	1141	1113	1493	1336	1087	1408	1585	2245	3055	1628	3397
Average read length		235	241	242	238	240	243	241	236	237	229	233	232
Average insert size		318	357	366	337	333	372	352	326	324	284	297	287
N50 (kbp)		127	104	127	140	127	86	101	127	112	135	127	127

Parameters	Ranges *	Laboratory 108											
		9065	9108	9201	9328	9409	9554	9568	9608	9904	9914	9987	9996
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		94.6	96.0	95.9	90.8	95.9	95.5	94.8	95.7	95.6	95.7	95.5	95.9
Species 2 (%)		1.3	0.9	0.9	1.6	1.0	0.9	1.4	0.9	1.0	0.9	1.0	0.9
Unclassified reads (%)	{<100}	2.2	1.9	1.9	5.6	1.9	2.3	2.3	2.2	2.2	2.1	2.0	1.9
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	4.5	4.8	4.8	4.9	5.0	5.0	5.0	4.9	4.9	4.5	4.9	4.8
Length [0–25] x min. coverage (kbp)	{<250}	0.0	0.0	0.2	0.0	0.3	0.3	0.2	0.0	0.2	20.0	0.0	0.2
Number of contigs at 25 x min. coverage	{>0}	4944	3503	3465	4670	3258	2417	2717	2539	2512	5182	2626	3253
Number of contigs [0–25] x min. coverage	{<1000}	0	0	1	0	1	1	1	0	1	62	0	1
Average coverage	{>50}	84	72,5	68,5	76,5	65	54,5	59	62,5	56,5	25,5	95,5	68,5
Number of reads (x1000)		1521	1308,5	1276	1485	1227	1018	1114	1222,5	1058	490	1752	1261
Average read length		291	294	285	293	292	286	289	272	282	278	289	286
Average insert size		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
N50 (kbp)		1	2	2	2	3	3	3	3	3	1	3	2

Parameters	Ranges*	Laboratory 123											
		9034	9100	9120	9128	9207	9450	9470	9473	9477	9555	9589	9899
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec, Sf	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		83.0	95.5	94.6	85.4	95.5	94.6	95.1	95.3	95.5	95.6	94.6	95.4
Species 2 (%)		2.1	1.2	1.3	7.7	1.3	1.3	1.3	1.2	1.2	1.3	0.7	1.2
Unclassified reads (%)	{<100}	13.3	2.3	3.2	3.3	2.4	3.0	2.7	2.6	2.5	2.2	2.9	2.4
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.5	5.2	5.2	5.3	5.1	5.1	5.1	5.4	5.2	5.2	5.1	5.2
Length [0–25] x min. coverage (kbp)	{<250}	0.3	0.0	0.9	0.0	0.0	0.0	0.0	0.8	0.6	1.1	0.0	1.2
Number of contigs at 25 x min. coverage	{>0}	316	210	217	286	291	211	277	252	242	255	222	241
Number of contigs [0–25] x min. coverage	{<1000}	1	0	1	0	0	0	0	3	2	2	0	3
Average coverage	{>50}	111	85	115	106	127	116	126	121	136	102	114	102
Number of reads (x1000)		2536	1775	2385	2363	2913	2478	2750	2739	2982	2146	2436	2153
Average read length		246	255	259	246	232	249	245	245	242	252	248	254
Average insert size		283	301	313	283	255	284	279	283	275	294	287	301
N50 (kbp)		73	109	87	98	109	103	112	112	109	97	116	109

		Laboratory 133										
Parameters	Ranges*	9180	9287	9357	9491	9538	9572	9789	9828	9929	9170	9173
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		86.8	92.6	82.8	92.1	74.7	85.2	93.4	93.9	92.9	90.6	90.2
Species 2 (%)		3.1	1.5	1.5	1.8	3.2	1.7	1.5	1.5	1.8	0.9	1.5
Unclassified reads (%)	{<100}	7.8	5.2	15.1	5.4	20.7	12.5	4.4	4.0	4.7	6.9	7.6
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.2	5.1	0.9	5.1	5.5	5.3	5.1	5.2	5.1	5.0	5.0
Length [0–25] x min. coverage (kbp)	{<250}	0.3	0.0	3843.4	0.5	0.2	0.0	0.0	0.6	0.0	23	18
Number of contigs at 25 x min. coverage	{>0}	553	344	323	286	283	471	262	257	279	2225	1311
Number of contigs [0–25] x min. coverage	{<1000}	1	0	3130	1	1	0	0	1	0	128	90
Average coverage	{>50}	96	107	**21**	104	92	103	135	106	108	94	87
Number of reads (x1000)		3015	2847	774	2707	2845	3857	3160	2480	2630	2597	2759
Average read length		183	210	153	216	190	157	237	238	229	206	173
Average insert size		192	223	152	232	206	159	269	268	253	230	181
N50 (kbp)		38	73	2	114	81	41	127	112	109	6	12

		Laboratory 134										
Parameters	Ranges*	9137	9197	9346	9426	9512	9517	9582	9783	9803	9804	9855	9991
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		96.0	95.8	95.8	96.0	96.0	86.8	96.1	93.4	96.2	96.1	96.0	95.8
Species 2 (%)		1.4	0.8	1.4	1.4	1.4	2.8	1.4	2.5	1.4	1.4	1.4	1.4
Unclassified reads (%)	{<100}	2.0	2.0	2.2	2.0	2.0	8.8	2.0	2.1	1.8	1.9	2.0	2.2
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.4	5.1	5.2	5.2	5.2	5.5	5.2	5.3	5.1	5.1	5.1	5.2
Length [0–25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0
Number of contigs at 25 x min. coverage	{>0}	198	171	201	191	163	221	189	247	188	201	204	192
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	0	0	0	0	0	0	3	0
Average coverage	{>50}	76	66	73	96	66	47	86	83	73	66	36	55
Number of reads (x1000)		2804	2297	2580	3397	2339	1747	3031	3030	2539	2312	1267	1935
Average read length		151	151	151	151	151	151	151	151	151	151	151	151
Average insert size		334	298	334	320	374	380	345	310	329	312	350	354
N50 (kbp)		127	118	127	132	135	88	127	140	127	127	109	112
		Laboratory 135											

Parameters	Ranges*	9012	9072	9264	9315	9391	9590	9613	9653	9830	9840	9851	9921
Detected species	{Ec} or {Sf}	Ec	Ec	Ec, Sf	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		94.6	94.4	82.8	95.0	95.3	95.2	82.0	95.4	93.6	92.9	92.8	94.8
Species 2 (%)		1.5	0.7	8.8	1.5	1.5	1.6	2.5	1.7	1.6	1.6	1.5	1.6
Unclassified reads (%)	{<100}	3.0	3.5	4.6	2.9	2.6	2.5	14.0	2.1	3.7	4.4	4.5	2.9
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.1	5.1	5.3	5.2	5.4	5.2	5.4	5.1	5.2	5.2	5.2	5.2
Length [0–25] x min. coverage (kbp)	{<250}	18.9	16.1	32.1	14.1	37.2	14.6	62.4	27.9	18.2	13.0	34.0	9.9
Number of contigs at 25 x min. coverage	{>0}	203	153	227	189	198	174	247	184	177	186	182	193
Number of contigs [0–25] x min. coverage	{<1000}	9	12	18	10	25	6	38	14	11	14	21	9
Average coverage	{>50}	83	96	103	108	98	98	77	79	99	102	94	102
Number of reads (x1000)		2979	3467	3935	3949	3661	3525	2962	2795	3675	3894	3534	3754
Average read length		147	145	145	145	147	148	146	148	145	142	144	144
Average insert size		390	376	368	365	398	408	380	418	354	342	375	357
N50 (kbp)		112	104	135	109	112	132	58	103	132	100	104	104

		Laboratory 137											
Parameters	Ranges*	9025	9111	9278	9385	9410	9474	9636	9735	9782	9818	9819	9977
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		83.0	91.3	92.4	91.1	91.8	91.9	91.7	91.6	91.7	91.7	92.4	90.4
Species 2 (%)		2.5	1.1	1.1	1.0	1.1	1.2	1.2	1.2	1.1	1.1	1.2	1.9
Unclassified reads (%)	{<100}	12.9	6.8	6.0	6.7	6.5	6.3	6.6	6.7	6.5	6.6	5.9	6.2
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.1	5.2	5.0	5.0	5.0	4.9	5.0	5.0	5.0	5.0	5.0	5.1
Length [0–25] x min. coverage (kbp)	{<250}	251.4	43.0	48.4	45.2	27.8	29.2	49.5	0.0	61.0	46.4	13.0	52.7
Number of contigs at 25 x min. coverage	{>0}	881	982	875	829	920	851	879	1075	856	883	962	1008
Number of contigs [0–25] x min. coverage	{<1000}	236	111	117	34	97	85	126	0	136	112	38	138
Average coverage	{>50}	36	48	48	56	53	52	46	204	45	48	63	46
Number of reads (x1000)		2047	2663	2515	2904	2806	2744	2470	10851	2376	2507	3305	2533
Average read length		98.5	98.9	99.3	99.4	98.8	98.9	98.8	98.7	98.7	98.9	99.3	99.5
Average insert size		343	348	328	348	313	325	313	311	302	355	327	326
N50 (kbp)		23	24	25	26	24	24	25	18	25	24	23	24

		Laboratory 139											
Parameters	Ranges*	9019	9185	9368	9378	9392	9468	9561	9563	9696	9708	9754	9994
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		95.6	95.7	95.9	92.9	95.6	95.9	95.9	96.1	85.4	95.1	96.0	95.8
Species 2 (%)		1.4	1.4	1.3	2.8	1.3	1.4	1.3	1.3	2.7	0.9	1.3	1.3
Unclassified reads (%)	{<100}	2.3	2.3	2.3	2.3	2.5	2.0	2.3	2.1	9.9	2.6	2.2	2.3
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.4	5.1	5.2	5.3	5.2	5.1	5.2	5.2	5.5	5.1	5.2	5.2
Length [0–25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Number of contigs at 25 x min. coverage	{>0}	214	204	207	255	204	186	185	204	241	153	196	201
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	0	0	0	0	0	0	0	0
Average coverage	{>50}	60	62	45	72	69	78	64	75	57	70	74	73
Number of reads (x1000)		2195	2175	1611	2616	2437	2712	2282	2640	2116	2435	2614	2585
Average read length		151	151	151	151	151	151	151	151	151	151	151	151
Average insert size		307	313	323	298	302	325	328	315	325	322	316	309
N50 (kbp)		105	112	112	140	111	132	132	112	82	116	109	112

		Laboratory 222											
Parameters	Ranges*	9077	9155	9300	9321	9395	9423	9603	9631	9648	9662	9746	9878
Detected species	{Ec} or {Sf}	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec	Ec
Species 1 (%)		96.3	96.4	95.5	89.6	96.1	95.9	95.2	96.4	96.1	96.5	96.9	96.4
Species 2 (%)		0.8	0.8	1.1	1.5	0.9	0.9	1.5	0.8	0.9	0.9	0.8	0.8
Unclassified reads (%)	{<100}	1.3	1.7	1.3	6.6	1.3	1.6	1.8	1.4	1.2	1.4	0.8	1.3
Length at 25 x min. coverage (Mbp)	{>45 ∧ <58}	5.1	5.1	5.1	5.4	5.1	5.1	5.1	5.1	5.1	5.3	5.1	5.1
Length [0–25] x min. coverage (kbp)	{<250}	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.3	0.0	0.0	0.2
Number of contigs at 25 x min. coverage	{>0}	546	583	421	543	539	430	2122	448	858	928	343	1144
Number of contigs [0–25] x min. coverage	{<1000}	0	0	0	0	0	0	12	0	1	0	0	1
Average coverage	{>50}	77,5	61,5	266,5	215,5	132	103	52	155,5	120,5	87	189,5	79,5
Number of reads (x1000)		1274,5	1078	4602	3769,5	2308,5	1820	963,5	2755,5	1948	1542	3014,5	1312
Average read length		318	299	300	322	300	295	290	296	326	308	324	317
Average insert size		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
N50 (kbp)		21	21	26	25	22	30	4	28	11	11	40	8

*Quality assessment made by the EQA-provider in-house quality control pipeline. *: indicative QC ranges; Ec: E. coli.*

Annex 15. Word format of the online form

This is a preview of all the fields and questions available.
Please keep in mind that, depending on your answers in the questionnaire, you will not necessarily have to answer all the questions (indicated by the "Go to").

1. STEC EQA-9 2018

Dear Participant
Welcome to the ninth External Quality Assessment (EQA-9) scheme for typing of STEC in 2018-2019.
Please note that most of the fields are required to be filled in before the submission can be completed.
Any comments can be written at the end of the form.
You are always welcome to contact us.
Please start by filling in your country, your Laboratory name and your LAB_ID.
Available options in this submission form include:
- Click "Options" and "Pause" to save your results and finish at a later time (using the same link)
- Click "Options" and "Print" to print your answers. This can be done at any time, but before pressing "Submit results"
- Click "Previous" to go back to the questions you have already answered
Note: After pressing "Submit results", you will not be able to review your results.

2. Country

- ❑ Austria
- ❑ Belgium
- ❑ Czech Republic
- ❑ Denmark
- ❑ Estonia
- ❑ Finland
- ❑ France
- ❑ Germany
- ❑ Greece
- ❑ Iceland
- ❑ Ireland
- ❑ Italy
- ❑ Latvia
- ❑ Lithuania
- ❑ Luxembourg
- ❑ Norway
- ❑ Poland
- ❑ Portugal
- ❑ Republic of Macedonia
- ❑ Romania
- ❑ Scotland
- ❑ Slovenia
- ❑ Spain
- ❑ Sweden
- ❑ The Netherlands
- ❑ Turkey
- ❑ UK

3. Institute name

4. Laboratory name

5. Laboratory ID

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

6. E-mail

7. Serotyping and virulence gene determination of STEC

8. Submitting results

- ❑ Submit serotyping/virulence gene determination results (please fill in the strain ID´s in the next section) - Go to 9
- ❑ Did not participate in the serotyping nor virulence determination part(s) - Go to 21

9. Serotyping/virulence isolate IDs

Please enter the strain ID (4 digits)
We recommend to print this page out! To have the overview of isolate ID´s and isolate No. 1-12, it will make the work easier.

Isolate 1 ___
Isolate 2 ___
Isolate 3 ___
Isolate 4 ___
Isolate 5 ___
Isolate 6 ___
Isolate 7 ___
Isolate 8 ___
Isolate 9 ___
Isolate 10 ___
Isolate 11 ___
Isolate 12 ___

10. Submitting results - Serotyping

(State one answer only)

- ❑ Both O group and H type - Go to 11
- ❑ Only O Group - Go to 11
- ❑ Only H type - Go to 13
- ❑ Did not participate in serotyping - Go to 15

11. Results for serotyping (O Group)

Please type the number of O Group by using (1-188)
Non Typable: 7777, Rough: 8888, Not done: 9999

O Group
Isolate 1 ___
Isolate 2 ___
Isolate 3 ___
Isolate 4 ___
Isolate 5 ___
Isolate 6 ___
Isolate 7 ___
Isolate 8 ___
Isolate 9 ___
Isolate 10 ___
Isolate 11 ___
Isolate 12 ___

12. Please specify the method used:

Phenotypic or molecular (PCR based, WGS based)

(State only one answer per question)
Method

- ❑ Phenotypic
- ❑ PCR based
- ❑ WGS based

13. Results for serotyping (H Type)

Please type the number of H Type by using (1-56)
H-: 6666, Non Typable: 7777, Not done: 9999

H type	
Isolate 1	___
Isolate 2	___
Isolate 3	___
Isolate 4	___
Isolate 5	___
Isolate 6	___
Isolate 7	___
Isolate 8	___
Isolate 9	___
Isolate 10	___
Isolate 11	___
Isolate 12	___

14. Please specify the method used:

Phenotypic or molecular (PCR based, WGS based)
(State only one answer per question)
Method

- ❑ Phenotypic
- ❑ PCR based
- ❑ WGS based

15. Submitting results - Virulence gene determination

- ❑ Submit virulence gene determination data (*eae, aaiC, aggR, stx1a, stx2* or subtyping
- ❑ Did not participate in the virulence gene determination (*eae, aaiC, aggR, stx1a, stx2* or subtyping) - Go to 21

16. Please specify the method used for the virulence gene determination (incl. subtyping):

- ❑ WGS
- ❑ Other

17. Results for virulence gene determination

Please use 1 for detected and 0 for not detected, Not done: 9999

	eae	*aaiC*	*aagR*	*stx1*	*stx2*
Isolate 1	___	___	___	___	___
Isolate 2	___	___	___	___	___
Isolate 3	___	___	___	___	___
Isolate 4	___	___	___	___	___
Isolate 5	___	___	___	___	___
Isolate 6	___	___	___	___	___
Isolate 7	___	___	___	___	___
Isolate 8	___	___	___	___	___
Isolate 9	___	___	___	___	___
Isolate 10	___	___	___	___	___
Isolate 11	___	___	___	___	___
Isolate 12	___	___	___	___	___

18. Submitting results – subtyping results

(State one answer only)

- ❑ Submit subtyping data

❑ Did not participate in subtyping - Go to 21

19. Results for subtyping

Subtyping of *stx1,* select variant (*stx1a, stx1c, stx1d*)
All isolates have to be subtyped regardless of the result of the initial screening. "Not done/ND" will by default be evaluated as an incorrect result.

	stx1a	*stx1c*	*stx1d*	*stx1a; stx1c*	*stx1a; stx1d*	*stx1c; stx1d*	Negative	ND
Isolate 1	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 2	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 3	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 4	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 5	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 6	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 7	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 8	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 9	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 10	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 11	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 12	❑	❑	❑	❑	❑	❑	❑	❑

20. Subtyping of *stx2* select variant (stx2a, stxb, stx2c, stx2d, stx2e, stx2f, stx2g)

All isolates have to be subtyped regardless of the results of the initial screening. "ND" will by default be evaluated as an incorrect result.

	stx2a	*stx2b*	*stx2c*	*stx2d*	*stx2e*	*stx2f*	*stx2g*	*stx2a stx2b*	*stx2a stx2c*	*stx2a stx2d*	*stx2b; stx2c*	*stx2b stx2d*	*stx2c stx2d*	*stx2a stx2b stx2c*	*stx2a stx2c stx2d*	*stx2b stx2c stx2d*	*stx2a stx2b stx2c stx2d*	Negative	ND
Isolate 1	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 2	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 3	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 4	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 5	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 6	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 7	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 8	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 9	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 10	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 11	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑
Isolate 12	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑	❑

21. Submitting Cluster results

❑ Cluster analyses based on PFGE and/or WGS - Go to 22

❑ Did not participate in the Cluster part - Go to 121

22. Cluster isolate ID´s

please enter the cluster isolate ID (4 digits)
We recommend to print this page out!
To have the overview of isolate ID's and isolate No. 1-12, it will make the work easier.
Cluster strain ID
Isolate 1 ___
Isolate 2 ___
Isolate 3 ___
Isolate 4 ___
Isolate 5 ___
Isolate 6 ___
Isolate 7 ___
Isolate 8 ___
Isolate 9 ___
Isolate 10 ___
Isolate 11 ___
Isolate 12 ___

23. Submitting Cluster analysis results

❑ Cluster analysis based on PFGE - Go to 24

❑ Do not wish to submit any cluster results based on PFGE analysis - Go to 29

24. Cluster analysis based on PFGE data

25. Please list the ID for the isolate included in the cluster of closely related isolates detected by PFGE results:

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

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

Indicate the isolate ID

27. xbaI - Total number of bands (>33kb) in the selected representative cluster isolate

28. Results for cluster analysis - PFGE (xbaI)

Please use 9999 for not analysed

	xbaI - Total number of bands (>33kb)	xbaI - Number of bands with same/shared position as the profile of the selected cluster isolate (>33kb)
Isolate 1		
Isolate 2		
Isolate 3		
Isolate 4		
Isolate 5		
Isolate 6		
Isolate 7		
Isolate 8		
Isolate 9		
Isolate 10		
Isolate 11		
Isolate 12		

29. Submitting Cluster results

(State one answer only)

❑ Cluster analysis based on WGS data - Go to 30

❑ Do not wish to submit any cluster results based on WGS data - Go to 121

30. Cluster analysis based on WGS data

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

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

- ❑ SNP based - Go to 33
- ❑ Allele based - Go to 40
- ❑ Other - Go to 32

32. If another analysis is used please describe your approach:

- Go to 47

33. Please report the used SNP-pipeline

Reference if publicly available or in-house pipeline

34. Please select the approach used for the SNP analysis

(State one answer only)

- ❑ Reference based - Go to 35
- ❑ Assembly based - Go to 38

35. Reference genome used

Please indicate Multi-locus Sequence Type (e.g. ST8) and isolate ID (e.g. one of the isolate from the current cluster, ID of a public reference isolate or an in-house isolate)

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

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

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

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

40. Please select tools used for the allele analysis

(State one answer only)

- ❑ BioNumerics - Go to 42
- ❑ SeqPhere - Go to 42
- ❑ Enterobase - Go to 42
- ❑ Other - Go to 41

41. If another tool is used please enter here:

42. Please indicate allele calling method:

- ❑ Assembly based and mapping based - Go to 43
- ❑ Only assembly based - Go to 43
- ❑ Only mapping based - Go to 44

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

44. Please select scheme used for the allele analysis

- ❑ Applied Math (wgMLST) - Go to 46
- ❑ Applied Math (cgMLST/Enterobase) - Go to 46
- ❑ Enterobase (cgMLST) - Go to 46
- ❑ Other - Go to 45

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

46. Pleas report the number of loci in the used allelic scheme

47. Cluster detected by analysis on data derived from WGS

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

48. Please list the ID´s for the isolates included in the cluster

Please use semicolon (;) to separate the isolate ID´s

49. Select a representative isolate in the cluster

Indicate the isolate ID

50. Results for cluster analysis (e.g. SNP or Allele based)

Please use 9999 for not analysed

	7-Multi-locus Sequence Type (ST)	Distance (e.g. SNP) to the selected cluster isolate
Isolate 1	_____	_____
Isolate 2	_____	_____
Isolate 3	_____	_____
Isolate 4	_____	_____
Isolate 5	_____	_____
Isolate 6	_____	_____
Isolate 7	_____	_____
Isolate 8	_____	_____
Isolate 9	_____	_____
Isolate 10	_____	_____
Isolate 11	_____	_____
Isolate 12	_____	_____

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

e.g. if SNP based results are submitted you can also report allele based results or results from an additional SNP analysis

(State one answer only)

- ❑ Yes - Go to 52
- ❑ No - Go to 91

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

(State one answer only)

- ❑ SNP based - Go to 54
- ❑ Allele based - Go to 61
- ❑ Other - Go to 53

53. If another analysis is used please describe your approach

- Go to 68

54. Please report the used SNP-pipeline

reference if publicly available or in-house pipeline

55. Please select the approach used for the SNP analysis

(State one answer only)

- ❑ Reference based - Go to 56
- ❑ Assembly based - Go to 59

56. Reference genome used

Please indicate Multi-locus Sequence Type (e.g. ST8) and isolate ID (e.g. one of the isolate from the current cluster, ID of a public reference isolate or an in-house isolate)

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

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

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

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

61. Please select tool used for the allele analysis

(State one answer only)

- ❑ BioNumerics - Go to 63
- ❑ SeqPhere - Go to 63
- ❑ Enterobase - Go to 63
- ❑ Other - Go to 62

62. If another tool is used please list here:

63. Please indicate allele calling method

(State one answer only)

- ❑ Assembly based and mapping based - Go to 64
- ❑ Only assembly based - Go to 64
- ❑ Only mapping based - Go to 65

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

65. Please select scheme used for the allele analysis

(State one answer only)

- ❑ Applied Math (wgMLST) - Go to 67
- ❑ Applied Math (cgMLST/Enterobase) - Go to 67
- ❑ Enterobase (cgMLST) - Go to 67
- ❑ Other - Go to 66

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

description

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

68. Additional analysis on data derived from WGS

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

Indicate the isolate ID

70. Results for the additional cluster analysis (e.g. SNP or Allele based)

Please use 9999 for not analysed

	7-Multi-locus Sequence Type (ST)	Distance (e.g. SNP) to selected cluster isolate
Isolate 1		
Isolate 2		
Isolate 3		
Isolate 4		
Isolate 5		
Isolate 6		
Isolate 7		
Isolate 8		
Isolate 9		
Isolate 10		
Isolate 11		
Isolate 12		

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

e.g. if SNP based results are submitted you can also report allele based results or results from an additional SNP analysis
(State one answer only)

- ❑ Yes - Go to 72
- ❑ No - Go to 91

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

(State one answer only)

- ❑ SNP based - Go to 74
- ❑ Allele based - Go to 81
- ❑ Other - Go to 73

73. If another analysis is used please describe your approach:

- Go to 88

74. Please report the used SNP-pipeline

reference if publicly available or in-house pipeline

75. Please select the approach used for the SNP analysis

(State one answer only)

- ❑ Reference based - Go to 76
- ❑ Assembly based - Go to 79

76. Reference genome used

Please indicate Multi-locus Sequence Type (e.g. ST8) and isolate ID (e.g. one of the isolate from the current cluster, ID of a public reference isolate or an in-house isolate)

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

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

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

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

81. Please select tool used for the allele analysis

(State one answer only)

- ❑ BioNumerics - Go to 83
- ❑ SeqPhere - Go to 83
- ❑ Enterobase - Go to 83
- ❑ Other - Go to 82

82. If another tool is used please enter here:

83. Please indicate allele calling method:

(State one answer only)

- ❑ Assembly based and mapping based - Go to 84
- ❑ Only assembly based - Go to 84
- ❑ Only mapping based - Go to 85

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

85. Please select scheme used for the allele analysis

(State one answer only)

- ❑ Applied Math (wgMLST) - Go to 87
- ❑ Applied Math (cgMLST/Enterobase) - Go to 87
- ❑ Enterobase (cgMLST) - Go to 87
- ❑ Other - Go to 86

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

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

88. Third analysis on data derived from WGS

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

Indicate the isolate ID

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

Please use 9999 for not analysed

	7-Multi-licus Sequence Type (ST)	Distance (e.g. SNP) to selected cluster isolate
Isolate 1	_____	_____
Isolate 2	_____	_____
Isolate 3	_____	_____
Isolate 4	_____	_____
Isolate 5	_____	_____
Isolate 6	_____	_____
Isolate 7	_____	_____
Isolate 8	_____	_____
Isolate 9	_____	_____
Isolate 10	_____	_____
Isolate 11	_____	_____
Isolate 12	_____	_____

91. Additional questions to the WGS part

92. Where was the sequencing performed

(State one answer only)

- ❑ In own laboratory
- ❑ Externally

93. Protocol used to prepare the library for sequencing:

(State one answer only)

- ❑ Commercial kits - Go to 94
- ❑ Non-commercial kits - Go to 96

94. Please indicate name of commercial kit:

95. If relevant please list deviation from commercial kit shortly in few bullets:

- Go to 97

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

97. The sequencing platform used

(State one answer only)

- ❑ Ion Torrent PGM - Go to 99
- ❑ Ion Torrent Proton - Go to 99
- ❑ Genome Sequencer Junior System (454) - Go to 99
- ❑ Genome Sequencer FLx System (454) - Go to 99
- ❑ Genome Sequencer FLx+ System (454) - Go to 99
- ❑ PacBio RS - Go to 99
- ❑ PacBio RS II - Go to 99
- ❑ HiScanSQ - Go to 99
- ❑ HiSeq 1000 - Go to 99
- ❑ HiSeq 1500 - Go to 99
- ❑ HiSeq 2000 - Go to 99
- ❑ HiSeq 2500 - Go to 99
- ❑ HiSeq 4000 - Go to 99

- ❑ Genome Analyzer Iix - Go to 99
- ❑ MiSeq - Go to 99
- ❑ MiSeq Dx - Go to 99
- ❑ MiSeq FGx - Go to 99
- ❑ ABI SOLiD - Go to 99
- ❑ NextSeq - Go to 99
- ❑ MinION (ONT) - Go to 99
- ❑ Other - Go to 98

98. If another platform is used please list here:

99. Criteria used to evaluate the quality of sequence data

In this section you can report criteria used to evaluate the quality of sequence data.
Please first reply on the use of 5 selected criteria, which were the most frequently reported by the participants in the STEC EQA-8 scheme, 2017-2018.
Next you will be asked to report 5 additional criteria of your own choise.
For each criteria please also report the threshold or procedure used to evaluate the current criteria.

100. Did you use confirmation of organism to evaluate the quality of sequence data?

- ❑ Yes
- ❑ No - Go to 104

101. Procedure used to evaluate confirmation of organism:

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

- ❑ Yes
- ❑ No - Go to 104

103. Procedure or threshold used for coverage:

104. Did you evaluate assembly quality?

- ❑ Yes
- ❑ No - Go to 104

105. Procedure used to evaluateassembly quality:

106. Did you use assembly length to evaluate the quality of sequence data?

(State one answer only)

- ❑ Yes
- ❑ No - Go to 108

107. Procedure or threshold used for assembly length:

108. Did you evaluate allele calling result?

- ❑ Yes
- ❑ No - Go to 110

109. Procedure used to evaluate allele calling:

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

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

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

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

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

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

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

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

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

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

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

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

121. Comment(s):

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

122. Thank you for your participation

Thank you for your submission of STEC EQA-9 results.
We highly recommend to document this Submission form by printing it. You will find the Print option after pressing the "Options" button.
Important: After pressing "Submit results" you will no longer be able to edit or print your information.
For final submission, remember to press "Submit results" after printing.

European Centre for Disease Prevention and Control (ECDC)

Gustav III:s Boulevard 40, 16973 Solna, Sweden

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

An agency of the European Union
www.europa.eu

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

Contact us
publications@ecdc.europa.eu

Follow us on Twitter
@ECDC_EU

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

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

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

Publications Office