European Union Reference Laboratory for *E. coli*
Department of Food Safety, Nutrition and Veterinary Public Health
Microbiological food safety and food-borne diseases Unit
Istituto Superiore di Sanità

Report of the 28th inter-laboratory study (PT28) on the identification and typing of Shiga toxin-producing *E. coli* (STEC) 2020

Edited by:

Margherita Montalbano Di Filippo, Silvia Arancia, Arianna Boni, Paola Chiani, Clarissa Ferreri, Guendalina Fornari Luswergh, Federica Gigliucci, Valeria Michelacci, Fabio Minelli, Gaia Scavia, Rosangela Tozzoli, Stefano Morabito

1. OBJECTIVES OF THE STUDY

The objectives of this study were to assess the proficiency of the NRLs for *E. coli* network in:

1. The detection of the main STEC virulence genes (*eae* and *stx* genes).
2. The identification of a range of relevant STEC serogroups (at least the 13 serogroups indicated in the EURL-VTEC_Method_003).
3. The subtyping of Shiga Toxins (Stx)-coding genes.
4. The identification of clusters of isolates based on genomic analysis.

This document represents the evaluation report of the PT28.

2. DESIGN OF THE STUDY

The study was designed according to the International Standard ISO/IEC 17043:2010 "Conformity assessment – General requirements for proficiency testing".

PT28 was conducted on a set of eight STEC strains and consisted of the following sections:

1. The identification of the Shiga toxin-producing *E. coli* main virulence genes by PCR amplification. Participants were requested to detect the following targets:

- *stx1* group, *stx2* group and the intimin-coding *eae* gene.

2. Determination of the serogroups of the strains. Participants were requested to identify the serogroup of the test strains assaying at least the following 13 serogroups, selected on the basis of their epidemiologic or regulatory importance:

- O26, O103, O111, O145 and O157: the top-5 STEC serogroups, most involved in severe human infections worldwide.
- O45 and O121: epidemiologically relevant and considered as adulterants in beef in the USA.
- O104: relevant after the 2011 German outbreak.
- O55, O91, O113, O128, O146: selected on the basis of their prevalence in human infections in Europe in the last years, according to the data collected by the European Centre for Disease Prevention and Control (ECDC).

3. Subtyping of the s*tx* genes present in the STEC strains. Participants were requested to identify the subtypes of the s*tx1* gene group (s*tx1a, stx1c* and *stx1d*) and s*tx2* gene group (from s*tx2a* to s*tx2g*).

4. The comparison of the genomic signatures of the isolates for the identification of the genomes belonging to a cluster. The participants were requested to characterise the isolates by whole genome sequencing and to determinate the relatedness between genomes using cgMLST or SNPs-based methods

3. PARTICIPANTS

Twenty-eight NRLs, representing 24 EU Member States, as well as Norway, Russia, Switzerland and UK participated in the study. Each NRL received its own individual laboratory numerical code, indicating the participant in the result tables.

The NRLs participating in the study were:

- Austria, *Institut für Medizinische Mikrobiologie und Hygiene,* AGES
- Belgium, Scientific Directorate Infectious Diseases in Humans (SCIENSANO)
- Bulgaria, NDRVMI, BFSA
- Cyprus, Laboratory for the Control of Foods of Animal Origin (LCFAO), Cyprus Veterinary Services
- Denmark, Microbiological Laboratory Ringsted
- Estonia, Veterinary and Food Laboratory
- Finland, Finnish Food Authority, Research and Laboratory Services Dept., Microbiology Research Unit, Helsinki
- France, *VetAgroSup,* LMAP/LNR/ *Equipe BPOE de l'UMR 5557 Ecologie Microbienne*
- Germany, Federal Institute for Risk Assessment (BfR), Department Biological Safety
- Hungary, Microbiological National Reference Laboratory, National Food Chain Safety Office, Food and Feed Safety Directorate
- Ireland, Department of Agriculture, Food and the Marine
- Italy, *Istituto Superiore di Sanità*
- Latvia, Research institute of Food Safety, Animal Health and Environment "BIOR"
- Norway, Norwegian Veterinary Institute
- Poland, National Institute of Public Health-National Institute of Hygiene, Dept. Food Safety

- Poland, National Veterinary research Institute (NVRI) Department of Hygiene of Food of Animal Origin
- Portugal*, Instituto Nacional de Investigação Agrária e Veterinária*, INIAV
- Romania, Institute for Hygiene and Veterinary Public Health
- Russia, International Department State Research Center for Microbiology and Biotechnology, Obolensk
- Slovakia, Dept. of Food Hygiene, Veterinary and Food Institute, Dolny Kubin
- Slovakia, National Reference Center for Environmental Microbiology, Public Health Authority, Bratislava
- Slovenia, Veterinary Faculty UL, National Veterinary Institute
- Spain, Bacteriology Department – 2, Central Veterinary Laboratory – Animal health, Ministry of Agriculture, Fisheries and Food, Algete
- Sweden, *Livsmedelsverket*/The National Food Agency
- Sweden, National Veterinary Institute (SVA)
- Switzerland, AGROSCOPE
- The Netherlands, Centre for Zoonoses and Environmental Microbiology (Z&O), National Institute for Public Health and the Environment (RIVM)
- United Kingdom, Public Health England FWEM Laboratory

4. MATERIALS AND METHODS

4.1. Sample preparation

Eight *E. coli* strains (samples 1 to 8) selected among those present in the EURL-VTEC reference collections and checked for the presence of all the required genetic and/or phenotypic features were sent to the NRLs. The characteristics of the strains reported in Table 1a were considered as the gold standard. Table 1b reports the virulence genes detected by WGS-based virulotyping performed at the EURL VTEC. The test strains were prepared on October 27th 2020 as fresh bacterial cultures seeded into soft (0.3 %) nutrient agar in borosilicate vials. The cultures were incubated 18 hours at 37°C ± 1°C and labeled with randomly generated numerical codes (3 or 4 digits), different for each set of strains sent to the NRLs. Previous data produced by the EURL-VTEC indicate that bacterial cultures prepared in this way are stable at least up to one month. On October 28th 2020, a homogeneity test was performed on six randomly selected sets of strain. The remaining test samples were stored at room temperature until November 9th 2020, when the parcels were shipped to the participating laboratories by courier. Twenty-seven laboratories received the parcel containing the test material within 24 h from the shipment, one laboratory (L825) received the parcel after 48 h.

Table 1a: Characteristics of the STEC strains included in the study

Strain	Serotype	ST	Target virulence genes		
			stx1	stx2	eae
1	O80:H2	301	-	*stx2f*	+
2	O80:H2	301	-	*stx2a*	+
3	O80:H2	301	-	*stx2a*	+
4	O80:H2	301	-	*stx2a*	+
5	O80:H2	301	-	*stx2a*	+
6	O26:H11	21	*stx1a*	-	+
7	O146:H21	442	*stx1c*	*stx2b*	-
8	O104:H7	2283	*stx1c*	-	-

Table 1b: Virulence genes detected in the test strains by WGS-based virulotyping

Strain	Virulence Genes
1	*cea, cma, cvaC, eae, efa1, ehxA, espA, espB, espF, espP, gad, hlyF, hra, ironN, iss, mchF, nleA, nleB, nleC, ompT, sitA, stx2, tir*
2	*cma , cvaC, eae, efa1, ehxA espA, espB, espP, gad, hlyF, iha, iroN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, stx2, sitA, tir*
3	*cvaC, eae, efa1, ehxa, espA, espB, espF, espP, etsC, gad, hlyF, iha, ironN, iss, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx2, tir*
4	*eae, efa1, ehxA, espA, espB, espF, espP, gad, hlyF, hra, iha, iroN, iss, mcbA, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx2, tir*
5	*eae, efa1, ehxA, espA, espB, espF, espP, gad, hlyF, hra, iha, iroN, iss, mcbA, mchB, mchC, mchF, nleA, nleB, nleC, ompT, sitA, stx2, tir*
6	*astA, cia, cib, cif, eae, efa1, ehxA, espA, espB, espF, espJ, espP, gad, iha, iss, iucC, iutA, katP, lpfA, nleA, nleB, stx1, terC, tir, toxB, traT*
7	*cea, ehxA, epeA, espI, focC, gad, iha, ireA, iroN, iss, kpsE, lpfA, mcbA, mchB, mchC, mchF, mcmA, sfaD, stx1, stx2, subA, tia*
8	*aaiC, celB, epeA, gad, ireA, katP, lpfA, neuC, orf3, stx1*

4.2. Laboratory methods

The laboratories were requested to identify the main STEC virulence genes by PCR (end-point or Real Time PCR) using any method applied in the routine testing. Methods for all the assays were also available in the EURL-VTEC website.

The participating Laboratories were also allowed to submit results obtained with WGS.

As far as the determination of the serogroups is concerned, participants were requested to identify the O-group of the STEC strains by testing at least for the following 13 serogroups: O26, O45, O55, O91, O103, O104, O111, O113, O121, O128, O145, O146, and O157. Participating labs could choose to apply any serological or molecular method in use in their laboratories, including WGS. However, procedures based on end-point or Real Time PCR for detecting the genes associated with the serogroups that were in the scope of the PT were available in the EURL website.

As for the *stx* genes subtyping, an end point PCR method for the identification of the *stx* gene subtypes of the STEC strains, based on the method described by Scheutz *et al.* (*J. Clin. Microbiol. 2012; 50: 2951-63*), was made available in the EURL-VTEC website. The

participating laboratories could choose to characterize the strains through WGS as well, and to report the results obtained with such a technique.
Finally, an exercise for the phylogenetic analysis of the isolates was carried out as part of PT28. The correlation between the test strains could be assessed by SNPs/wg/cgMLST analysis: in particular, the laboratories were requested to submit the number of SNPs or allelic differences observed between each of the strains assayed and one of the test strains selected as reference. In addition, the laboratories were requested to interpret their own results by indicating which strains were part of the same cluster.

4.3. Collection and elaboration of the NRLs' results

The results were submitted through an on-line form prearranged by the EURL for *E. coli*. The instruction on how to report the results and the link to access the form was sent by E-mail to all the participants' laboratories.

4.4. Analysis of the NRLs' results

4.4.1. Evaluation of the NRLs performance in the identification of the STEC virulence genes and the serogroups

The performance of each NRL in the identification of the virulence genes of STEC was evaluated by assigning penalty points for each incorrect result in the STEC virulence genes detection according to the following scheme:

- **4 penalty points** to each incorrect or missing result concerning the identification of the *stx* genes;
- **2 penalty points** to each incorrect or missing result concerning the identification of the *eae* gene;
- **2 penalty points** to each incorrect result concerning the identification of the serogroups.
- **1 penalty point** when the results of the serogroup identification were not uploaded ("null" field) or reported as "Not Done". No penalty points were instead assigned to the laboratories reporting the serogroup of the STEC isolated strain as not typeable (ONT).
- **1 penalty points** to each incorrect result concerning the identification of the *stx* genes subtypes.

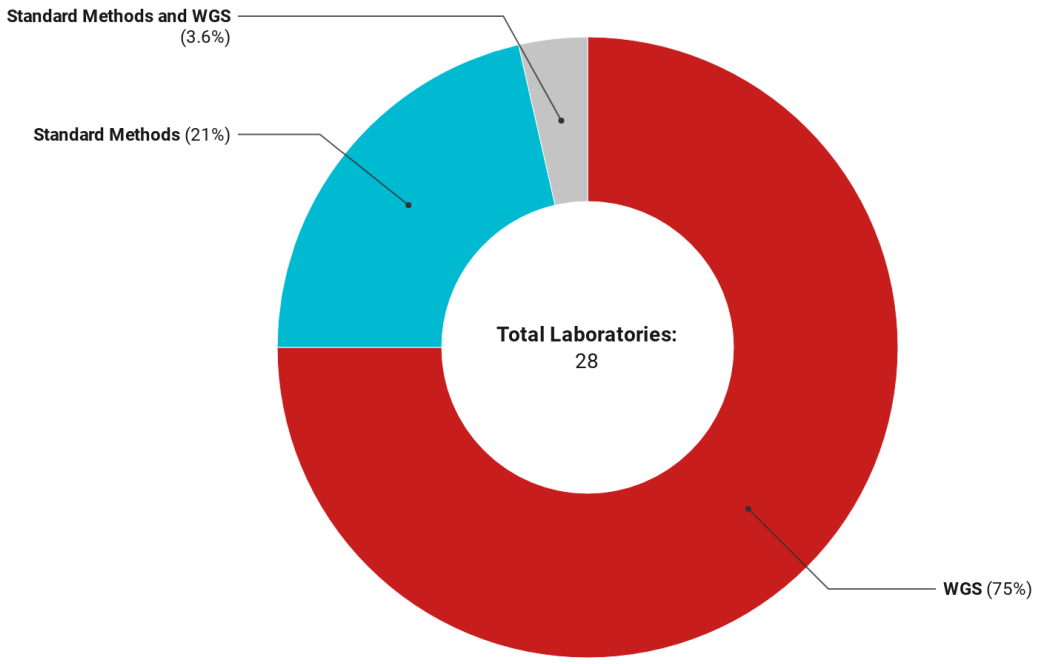
The sum of the penalty points received was used to assess the proficiency of the NRLs. A threshold of eight points was set and the laboratories presenting a higher score were considered as under-performant.

Penalties accumulated following incorrect or missing results reported for the *stx* genes subtyping were not considered for the assessment of the laboratories' proficiency but rather as indicators to identify areas where the method should be improved or the action of the EURL-VTEC in support to the NRLs enhanced.

5. RESULTS

Results were submitted by 28 Laboratories. Seven provided results obtained by conventional methods, 20 by WGS and one (L734) carried out both the conventional and WGS-based methods. **Figure 1** shows the number of participating laboratories aggregated according to the methods used to characterize the isolates.

Figure 1: Number of Laboratories reporting results obtained with the different analytical methods



5.1. Identification of the *E. coli* virulence genes

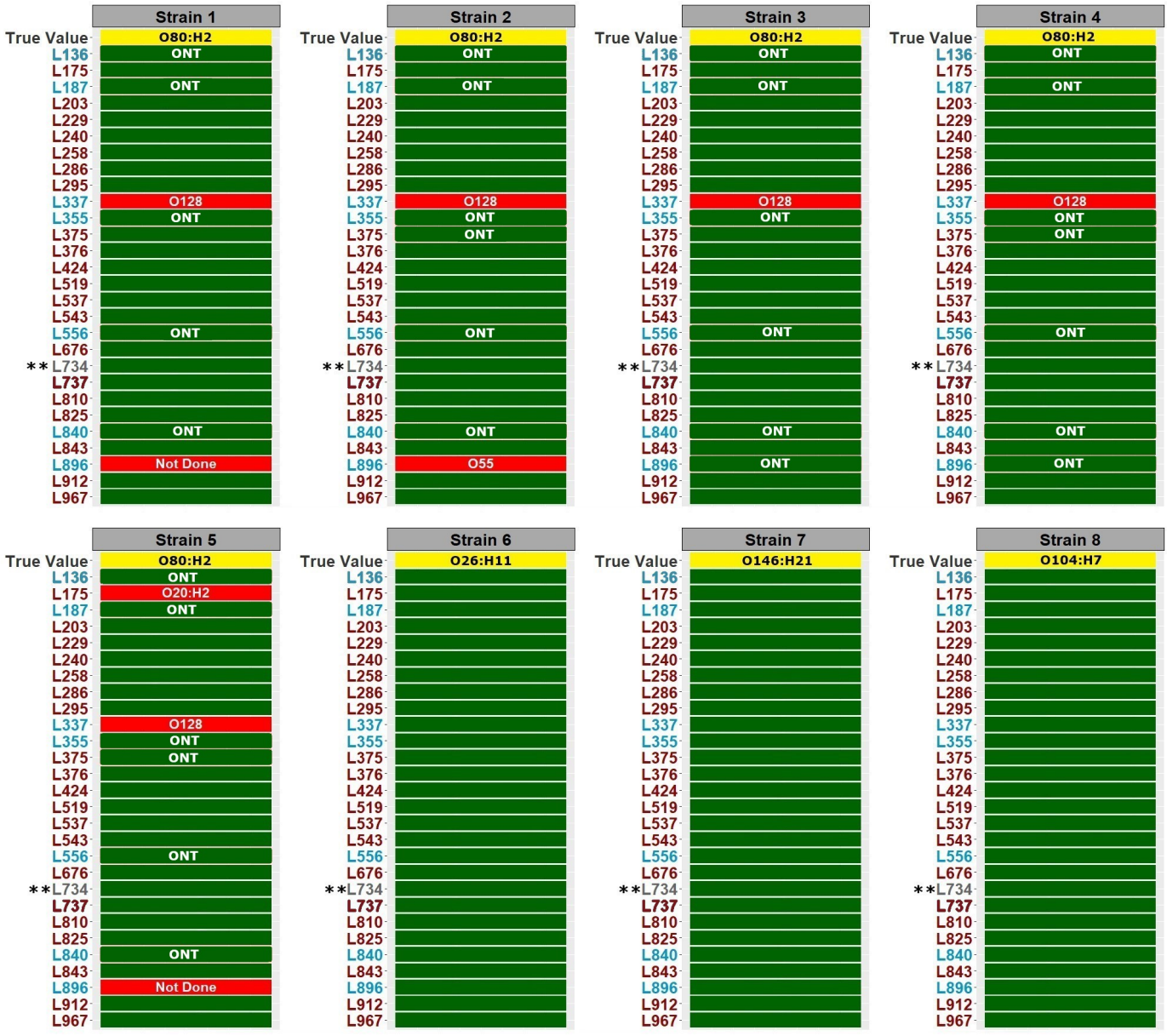
The individual results concerning the identification of the *stx1*, *stx2* and *eae* genes are reported in **Figure 2**.

Figure 2: Identification of the *E. coli* virulence genes. Green boxes: correct results, red boxes: incorrect results. The gold standards (True Values) are reported in yellow. Laboratories indicated with blue codes reported results obtained by standard methods, while the red ones reported the results obtained with WGS. The symbols ** indicate the laboratory (L734) that carried out both the conventional and WGS-based methods.

5.2. Identification of the serogroups of the test strains

The results of the identification of the O-groups of the eight test strains are shown in **Figure 3**.

Figure 3: Identification of the serogroups of the test strains. Green boxes: correct results, red boxes: incorrect results (ONT: the serogroup was not identified). The gold standards (True Values) are reported in yellow. Laboratories indicated with blue codes reported results obtained by standard methods, while the red ones reported the results obtained with WGS. The symbols **** indicate the laboratory (L734) that carried out both the conventional and WGS-based methods.



5.3. Subtyping of *stx* genes in the test strains

The results of the *stx* genes subtyping are shown in **Figure 4**.

Figure 4: Subtyping of the *stx* genes. Green boxes: correct results, red boxes: incorrect results. The gold standards (True Values) are reported in yellow. Laboratories indicated with blue codes reported results obtained by standard methods, while the red ones reported the results obtained with WGS. The symbols ** indicate the laboratory (L734) that carried out both the conventional and WGS-based methods.

5.4 Cluster analysis

A total of 21 NRLs participated in the cluster analysis exercise and performed the phylogenetic analysis on the strains received for PT28. The majority (13 labs) used cgMLST and seven NRLs calculated the SNPs differences in the whole genome to infer on the relatedness among strains. One laboratory applied both the strategies (**Figure 5a**). The results of the cluster analysis (see **Figure 5b**) showed that the majority of laboratories (17) could identify the correlation among strains 4 and 5. Four laboratories reported errors in identifying the cluster: L295, L375 and L810 wrongly reported strain 1, strain 2 and strain 3 as part of the cluster applying different strategy (see the "Distance" table in **Figure 5b**), while L258, applying alleles-based cgMLST, wrongly reported strain 1, and strain 3 as part of the cluster.

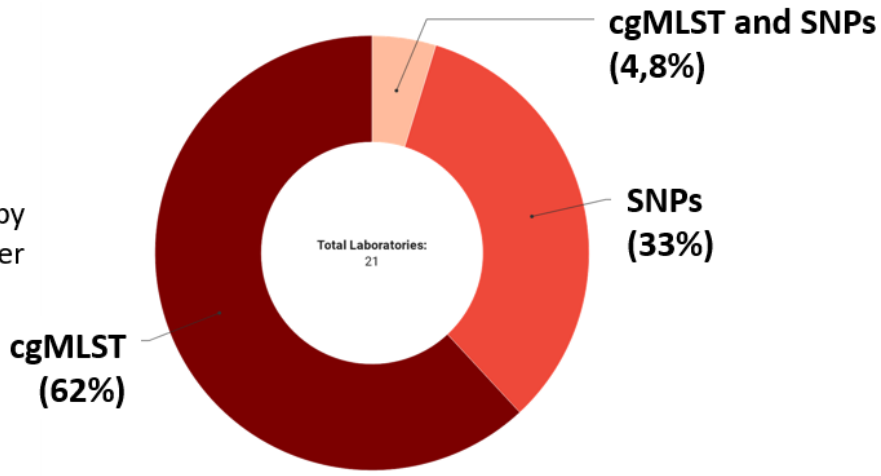
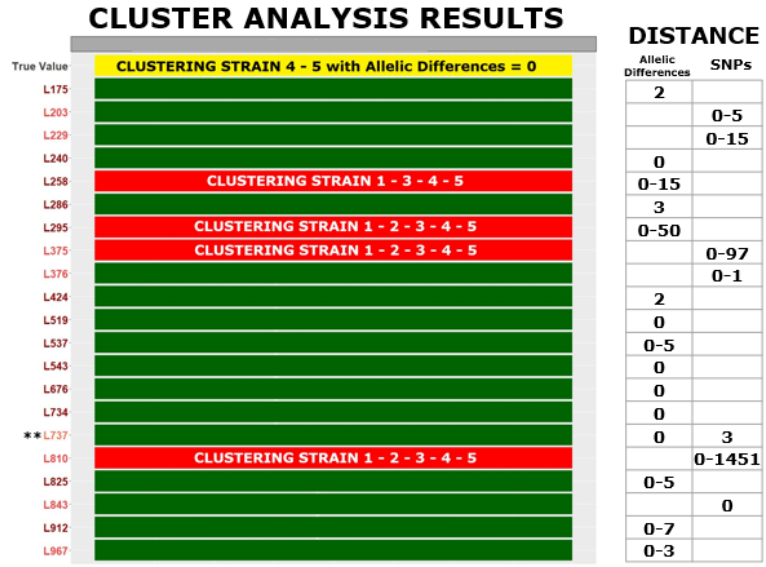


Figure 5a: WGS-based methods used by the NRLs participating in the cluster analysis exercise



DISTANCE

Allelic Differences	SNPs
2	
	0-5
	0-15
0	
0-15	
3	
0-50	
	0-97
	0-1
2	
0	
0-5	
0	
0	
0	
0	3
	0-1451
0-5	
	0
0-7	
0-3	

Figure 5b: Results obtained by NRLs for the cluster analysis exercise.
Green boxes: correct results, red boxes: incorrect results. The gold standards (True Values) are reported in yellow. Laboratories indicated with dark red label reported results obtained by cgMLST strategy, while the pink ones reported the results obtained with SNPs.
The symbols ** indicate the laboratory (L737) carried out both strategies.

The table showed the allelic differences/SNPs from the cluster strains reported by NRLs for cluster investigation.

6. Evaluation of the proficiency of the participating Laboratories

The proficiency of the Laboratories has been assessed using the results on the identification of the *stx* and *eae* genes, as well as on the determination of the serogroups. The participating laboratories presenting a score higher than eight were considered as under-performant (red bars in **Figure 6**).

Figure 6

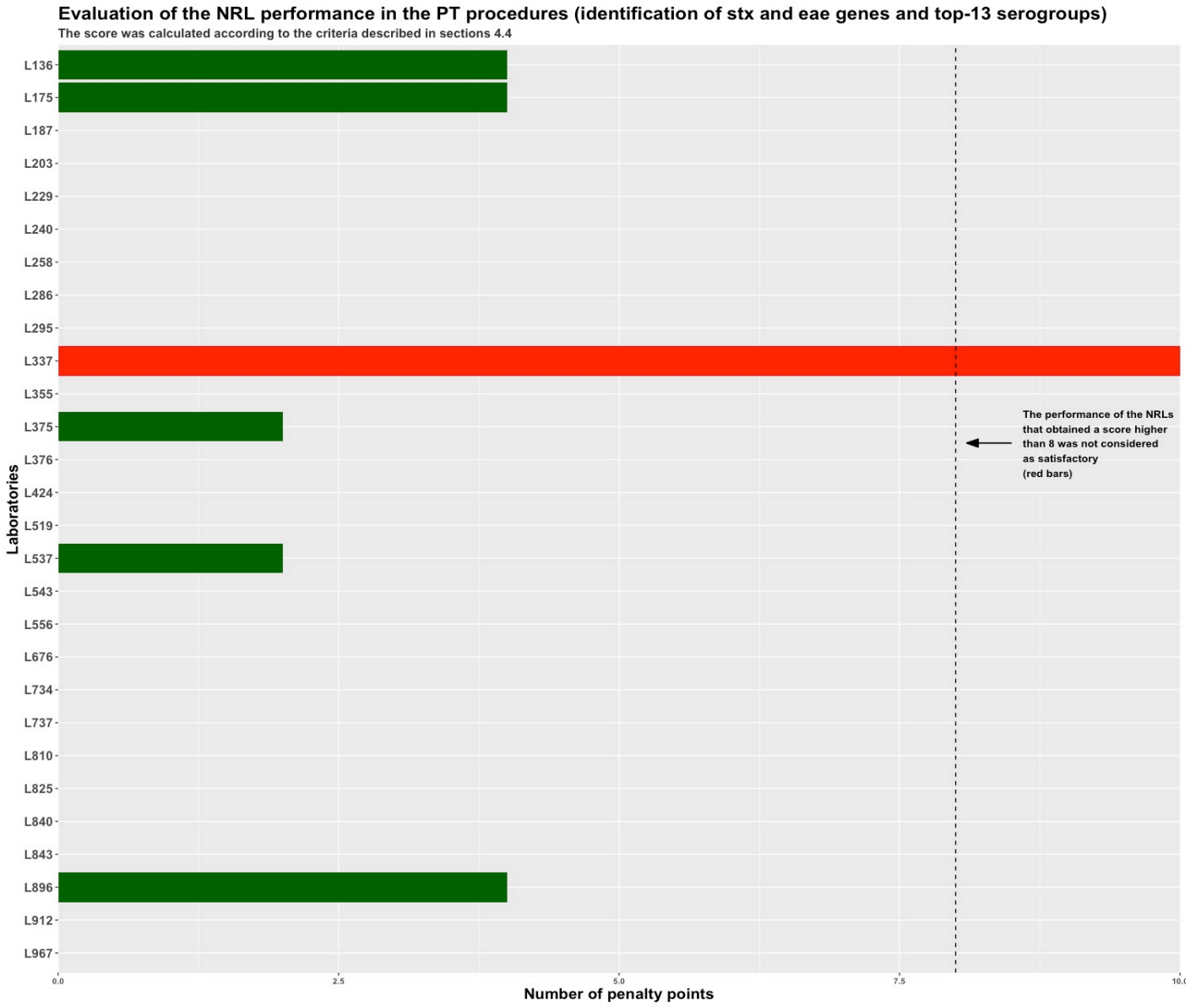
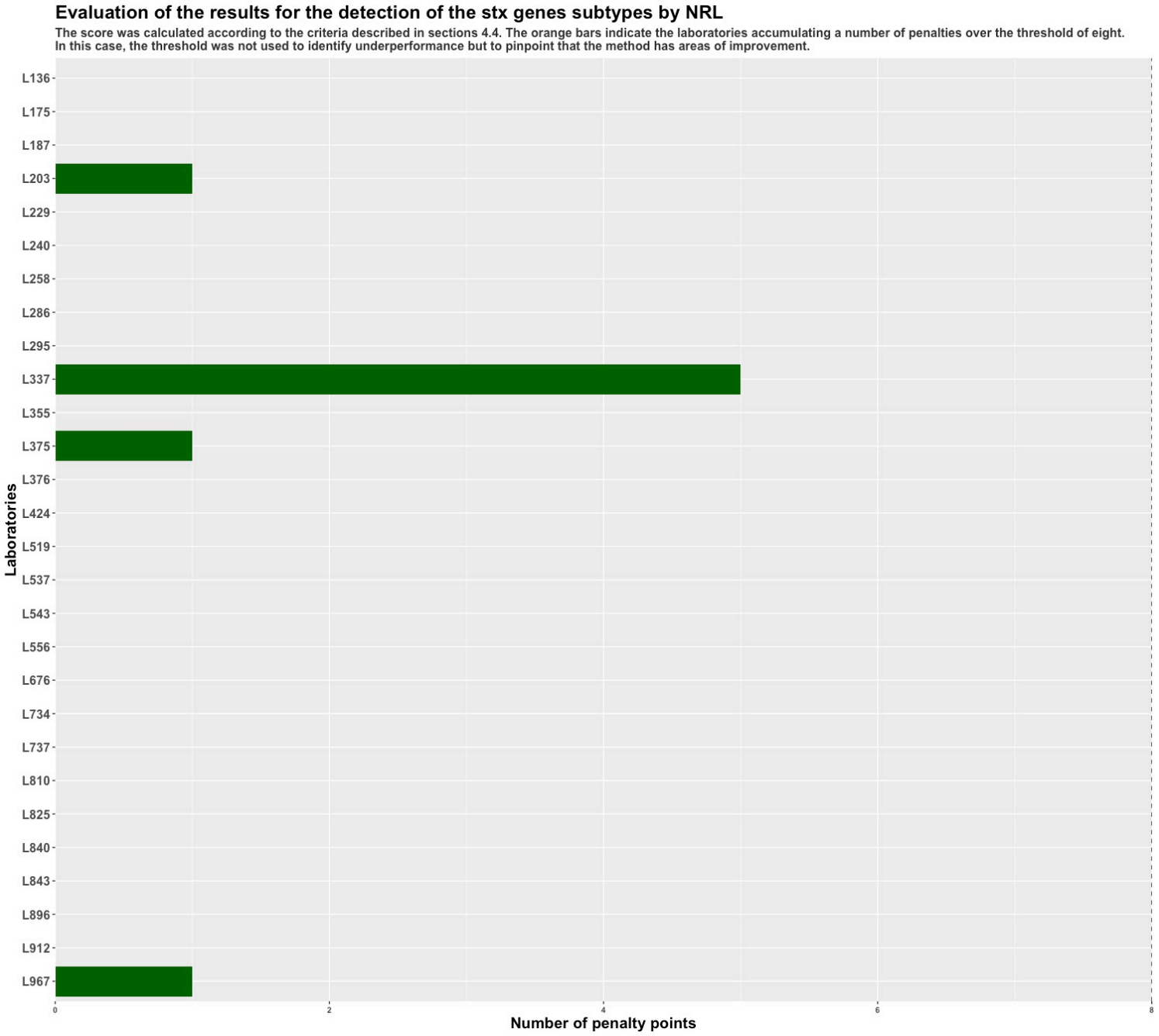


Figure 7: Evaluation of the results for the detection of the *stx* genes subtypes by NRLs



7. CONCLUDING REMARKS

1. A lower participation was recorded for PT28 compared with the previous rounds of PTs organized by the EURL for *E. coli*. However, considering the COVID-19 pandemic, the level of participation observed confirmed the eagerness of the network to collaborate in the EURL initiatives. More than 50 % of the laboratories performed WGS, which displayed an excellent performance in both the characterization and subtyping of the STEC isolates.
2. Only one laboratory underperformed in the characterization of STEC strains (detection of *stx*, *eae* and serogroups). This result indicates that the capability of the NRLs in identifying the main STEC virulence genes is highly satisfactory (more than 95% of laboratories with correct results), but still identifies an area where the technical skill of the network must be improved through provision of advice and training.
3. Although the majority of serogroups of the STEC strains fell outside the group of the most searched ones and out of the scope of the standard method ISO TS 13136:2012, the participating Laboratories performed well in their identification when WGS-based method was used.
4. Most of the laboratories submitted results for the *stx* subtyping, indicating that this assay is becoming widely adopted among the network of NRLs for *E. coli*, with a good performance on average. Apart from known criticalities of the typing method (*e.g.* discrimination between *stx2a* and *stx2c* genes in the PCR assay), the network responded well to the *stx* subtyping exercise, particularly when the results were obtained through WGS. This represents undoubtedly an area where the EURL for *E. coli* must devote efforts in refining the PCR method and providing training. Additionally, stimulating the adoption of WGS and developing easy to interrogate bioinformatics procedure will be one focus of the EURL-VTEC action.
5. 75% of the laboratories performed the cluster analysis exercise, representing all those that characterized the strains through WGS. The results showed that the majority of laboratories performed well in this part of the PT28 regardless the method used (cgMLST or SNP analysis). Among the points to be highlighted, the need to define a threshold to call the clusters using the SNP analysis should be defined.