Proficiency Tests on the identification and typing of pathogenic *E. coli* strains

The use of Whole Genome Sequencing

EU-RL VTEC

Istituto Superiore di Sanità, Dep. of Food Safety, Nutrition and Veterinary Public Health
European Union and National Reference Laboratory for *E. coli*, Rome, Italy

Proficiency Tests on WGS

PT WGS1 (2017): voluntary inter-laboratory study; only raw data collected

PT23 (2018): characterization by classical methods or WGS; cluster analysis through PFGE

PT26 (2019): characterization by classical methods or WGS; cluster analysis through PFGE or WGS

PT28 (2020)

PT31 (2021)

characterization by classical methods or WGS; cluster analysis through WGS

PT35 (2022)

...

Voluntary inter-laboratories study

The objectives of the study were:
- to evaluate the quality of the sequences produced and the effect on the WGS-based characterisation of STEC
- to evaluate the inter laboratory and platform variability

Methods:
- **Wet-lab:** Shipment of 6 STEC strains belonging to the same serogroup
- **Collection of results**: collection of raw reads through IRIDA platform
- **Results analysis**: Analysis of data at EURL VTEC through ARIES Galaxy platform: serotyping, virulotyping, MLST, reference-free wgSNPs, cgMLST, wgMLST
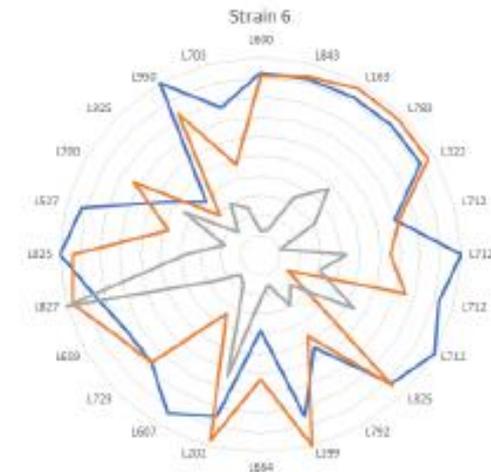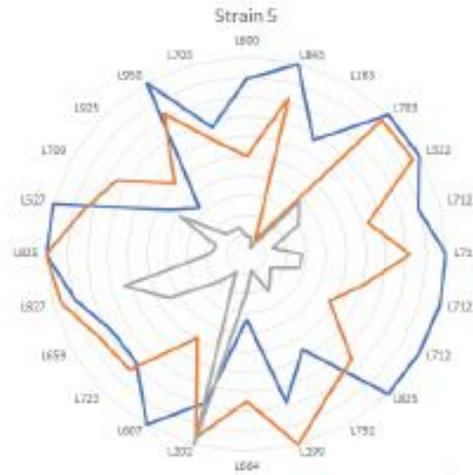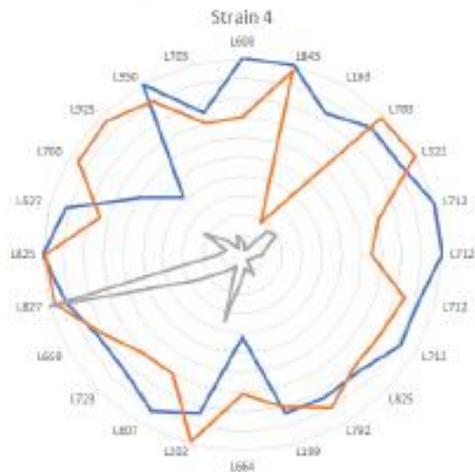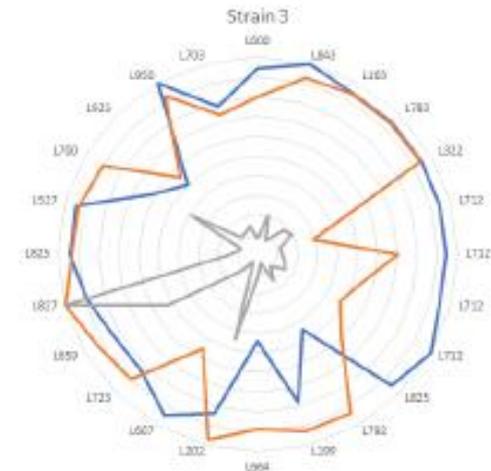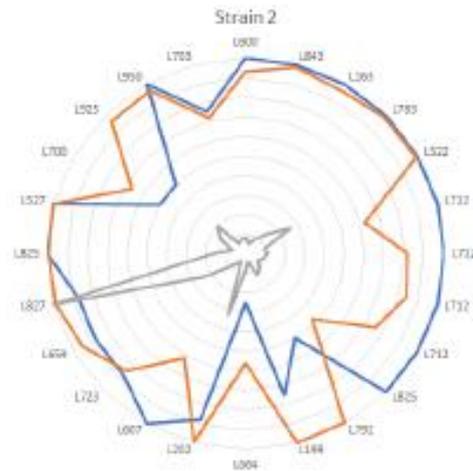
Results PT WGS 1 (2017-2018)

Legend:
- Assembly Coverage
- N50
- Depth
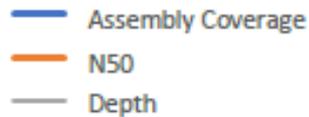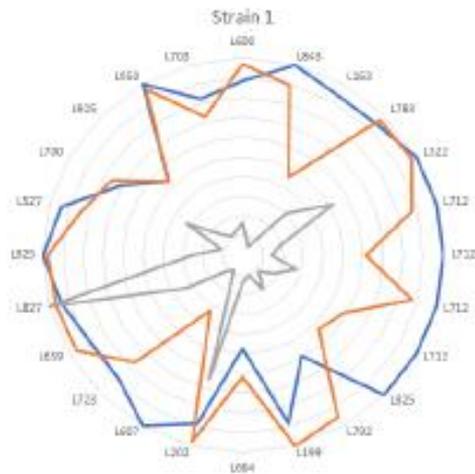
21 participants (including 3 Italian Official Laboratories)
No performance assessment
Results used to evaluate the overall quality of the sequences

PTs on NGS *E. coli* 2018 and 2019 (PT23 and PT26)

The objectives of the studies were:

- The detection of the main virulence genes (for STEC, EAEC and ETEC)
- The detection of a range of relevant serogroups
- *stx* subtyping
- cluster analysis <u>only through PFGE</u> in PT23 (2018) and <u>through PFGE or WGS (wgSNPs or cgMLST)</u> in PT26 (2019)

The participating Laboratories could submit either the results obtained by applying the laboratory procedures available at the EURL *E. coli* website, based on **conventional and Real Time PCR, or the WGS-based characterization** results.

Methods:

- **Wet-lab and dry-lab**: Shipment of 6-7 strains, typing results collected
- **Collection of results:** through a <u>Restricted Area</u> on EURL VTEC website. <u>Manual extraction and analysis</u> of results
- **Performance assessment:** Same penalty points applied for errors either obtained through conventional methods or through WGS

PTs on NGS *E. coli* 2020-2022 (PT28, PT31 and PT35)

The objectives of the studies were:

- The detection of the main virulence genes (for STEC, EAEC and ETEC)
- The detection of a range of relevant serogroups
- *stx* subtyping
- cluster analysis <u>only through WGS (wgSNPs or cgMLST)</u>

The participating Laboratories could submit either the results obtained by applying the laboratory procedures available at the EURL *E. coli* website, based on **conventional and Real Time PCR, or the WGS-based characterization** results.

Methods:

- **Wet-lab and dry-lab**: Shipment of 8 strains, typing results collected
- **Collection of results**: through <u>Microsoft Forms</u>. <u>Automatic extraction of results, automatic analysis</u> (and individual reports generation from 2021)
- **Performance assessment:** Same penalty points applied for errors either obtained through conventional methods or through WGS

WGS cluster analysis results reporting

- In PT26 (2019): Distance from each of the strains to one of the test strains reported

- From PT28 (2020): only information on the interpretation of results requested:

Did you perform cluster analysis? Y/N

47. Cluster analysis *
Indicate the method used for the cluster analysis

☐ cgMLST

☐ SNP

48. Indicate the blind codes of the strains that belong to a cluster, separated by commas *
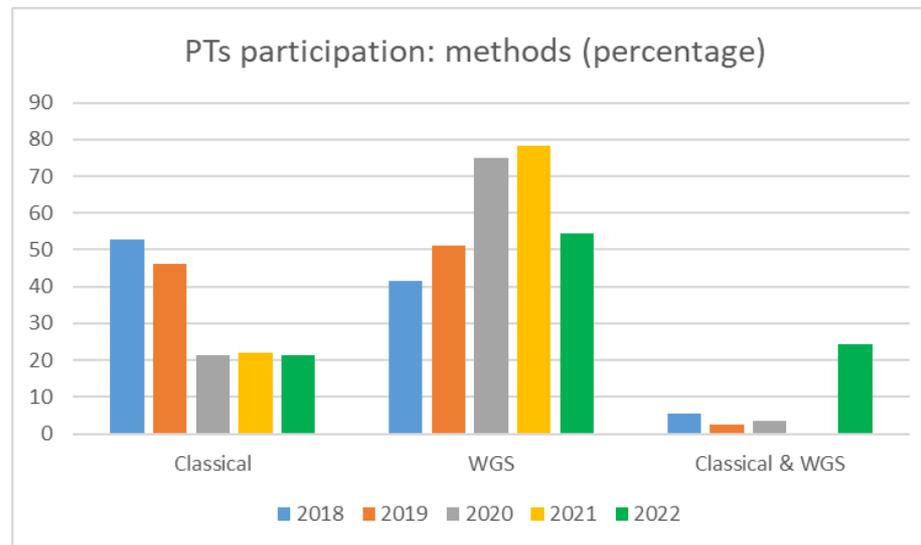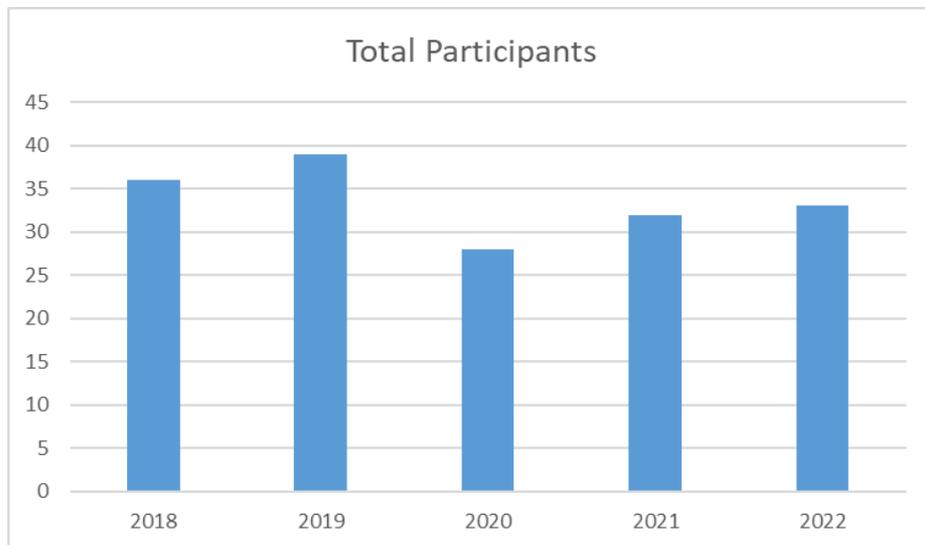E.g. 6654, 5278 etc...

> Inserisci la risposta

49. Cluster analysis: Indicate distances *
Indicate the range of differences identified among the samples composing the cluster (e.g. 0-5 allelic differences or 0-15 SNPs)
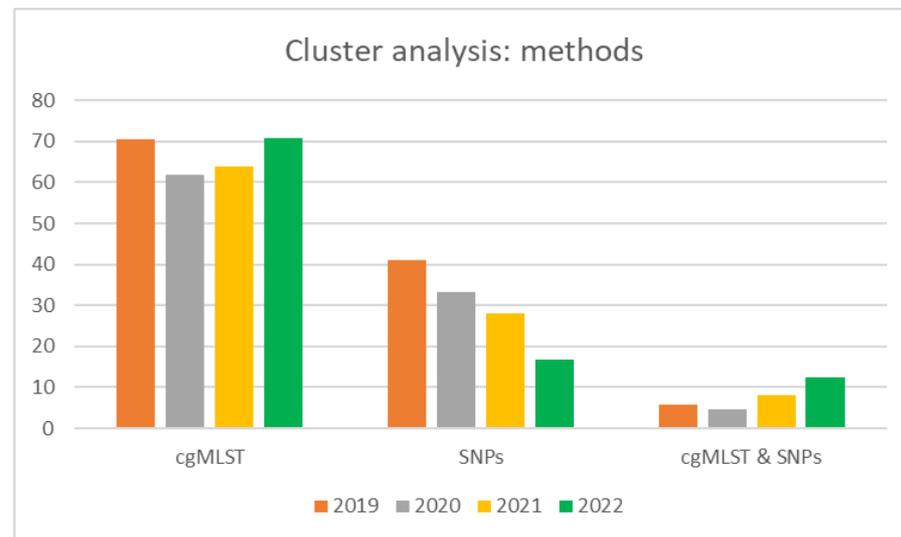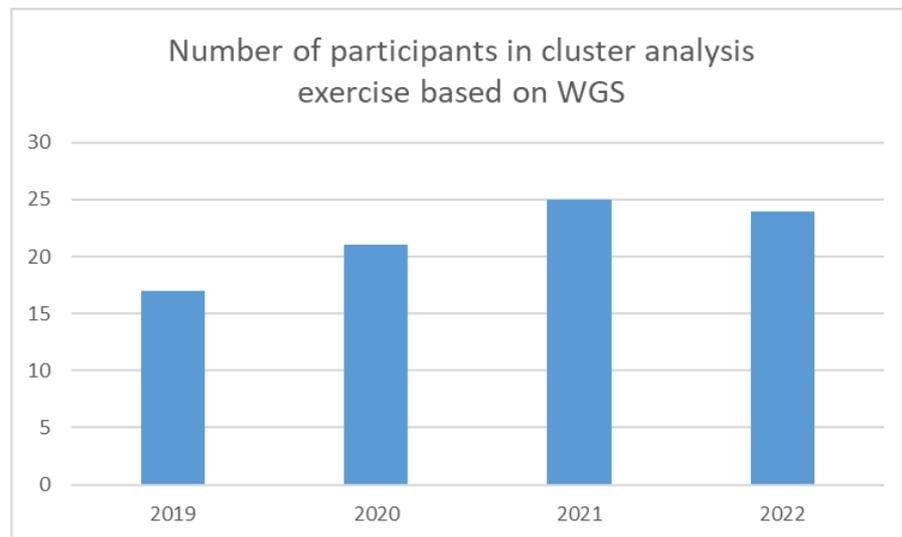
> Inserisci la risposta

Participation trend 2018-2022

Increase in the use of WGS vs classical methods

Increase in the use of cgMLST

Lessons learnt

- Collection of raw data requires automatic methods for collection and analysis

- The collection of characterization results allows to avoid intense data treatment and allows to evaluate the interpretation of results in the laboratories

- The use of web forms for results collection allows automatic data analysis, reducing the time needed for reporting

- The collection or results obtained through classical and/or WGS methods allows to compare the performance of the two approaches, with better results for WGS