### Inter-EURLs Working group on Next Generation Sequencing

Valeria Michelacci

4 - 5 November 2019





#### AIM:

to promote the use of WGS across the EURLs' networks, build WGS capacity within the EU and ensure liaison with the work of the EURLs and the work of EFSA and ECDC on the WGS mandate sent by the Commission

#### **MEMBERSHIP:**

EURL *E. coli* (coordinator) EURL *Listeria monocytogenes* EURL CPS EURL *Salmonella* EURL *Campylobacter* EURL Parasites EURL Parasites EURL AR EURL Foodborne viruses Observers: SANTE G4, EFSA, ECDC

### Kick off: November 14th 2017





### TASKS

# To elaborate guidance documents, reports and common strategies for:

- PTs
- WGS laboratory procedures
- Bioinformatics tools
- WGS cluster analysis
- Bench marking
- Training on NGS
- Reference and confirmatory testing using NGS
- Follow-up of ISO activities on WGS





### **Meetings**

- Twice a year
- 4 meetings already performed
- 5<sup>th</sup> meeting in December at EC
- 6<sup>th</sup> meeting in March at ISS, EURL-VTEC in occasion of

### **MVN-funded project**

To bring to the attention of all the stakeholders the state of play in the field in EU, including the activity of the Inter-EURLs WG on NGS and the programs of EC, EFSA and ECDC. **Stakeholders:** NRLs, OLs, competent Authorities of the EU, Member States, in addition to the MVN partner institutions.

To stimulate the adoption of the technology by the Member States that are currently not applying such methodologies





### **MVN-funded workshop in March 2020**



Tuesday 10 March		
09:00	Welcome and Introduction (10')	
09:15	Greetings (10')	Stefano Morabito
09:30	Vision on NGS (20')	Pamina Suzuki, EC
09:55	EFSA State of play and future perspectives (30')	Valentina Rizzi, Mirko Rossi EFSA
10:30	ECDC State of play and future perspectives (30')	Johanna Takkinen, ECDC
11:00	Coffee Break (30')	

Inter EURLs working group on NGS		
11:30	NGS-based Proficiency Testing (10')	Rene Hendriksen, EURL AMR
11:45	WGS laboratory procedures (10')	Simone Cacciò, EURL Parasites
12:00	Bioinformatics tools for WGS analysis (10')	Valeria Michelacci, EURL VTEC
12.15	WGS Cluster Analysis (10')	Joakim Skarin, EURL Campylobacter
12:30	Bench marking of bioinformatics tools (10')	Maroua Sayeb, EURL Lm
12:45	Training on NGS (10')	Deborah Merda, EURL CPS
13:00	Reference and confirmatory testing using NGS (10')	Angela Van Hoek, EURL Salmonella
13.15	Strategies for WGS on low levels of highly sequence variable viruses (10')	Magnus Simonsson, EURL Food borne viruses
13:30	Launch (1h)	
14:30	Round Table (1h, 45') Chairman: Stefano Morabito Speakers: Valentina Rizzi, Johanna Takkinen and Pamina Suzuki	
16:15	Concluding remarks	Stefano Morabito
16:30	Closure	





### **Deliverables**

Documents will be made available through a dedicated page in the website of each EURL linking the same documents, for easy update by the responsible of each task

- 1. Document on PT-WGS information previous experience and future plans
- 2. Reference genome collection
- 3. Active document with links to protocols and including protocols not yet available online
- 4. Active document with links to bioinformatics
- 5. Documents stating differences among workflows for cluster analysis
- 6. Guidance document on benchmarking of WGS data
- 7. Document with info on training modules available
- 8. Report on the two surveys
- 9. Comparison of WG survey with EFSA





#### Survey on the use of NGS in NRLs





## 87/176 do not sequence at all





#### Sequence quality check & Analytical approach/3



Q13: Do you perform cluster analysis?





### Task 3 – Bioinformatics tools - inventory

#### Aim:

To collect info on the bioinformatics tools in use, starting from what is already available and public, and spot the areas in which tools are missing and needed

> Data panel 39 NRLs

	EURL-VTEC	EURL-Campylobacter	EURL-AMR	EURL-Salmonella
Replies	7 NRLs	12 NRLs	10 NRLs	10 NRLs

- Agreement on tools for basic analysis (QC, trimming and assembly)
- Different approaches for characterization, but use of shared databases





### **Quality check**



TOOL	No. Of NRLs	Source
FastQC	28	ARIES, CGE, Other webserver, Free software, In house, Commercial





### Trimming



TOOL	No. Of NRLs	Source
Trimmomatic	20 (10 different Institutes)	ARIES, Other webserver, Free software, In house





### Assembly



TOOL	No. Of NRLs	Source
SPAdes (v 3.8 or higher)	28 (14 different Institutes)	ARIES, CGE, Other webserver, Free software, In house, Commercial





#### MLST

Several tools used, consolidated databases for 7-genes

#### Virulotype

Most common tools: VirulenceFinder, ABRicate, Virulotyper Most common databases used: VirulenceFinder, VFDB

#### Serotype

Most common tools: SerotypeFinder, SeqSero, SISTR, Serotyper Species-specific databases

#### **Resistance genes**

Tools used: ResFinder, RGI+CARD, ARIBA, ABRicate, ARG-ANNOT, GeneFinder Most common databases used: ResFinder, CARD DB





Reason	No of NRLs	
NGS is currently out of scope of the tasks for our NRL	5	
NGS is currently out of scope of the tasks for our NRL (But capacity is present in the institute)	1	
No plans due to lack of resources	1	
No plans due to lack of resources and/or relevant expertise	•	In rep
Seek cooperation with NVI	1	<b>-</b>
Yes just started	5	in the
Yes just started, but still lack of relevant expertise	1	Th
Yes, we have (unspecific) plans but no resources allocated	8	US
Yes, we have (unspecific) plans but lack of relevant expertise	2	pe ap
Yes, we have (unspecific) plans to setup NGS activities at our NRL	6	SC
Total	34	

**Follow-up Survey for** those replying not to use NGS (November 2018)

- some MSs **limited funding** stills presents a hindrance for accessing NGS
- e EURLs can effectively **cover the gap in** e capacity
- e awareness of the situation should be ed to install a policy aiming at the diffusion the technology overcoming the lack of rception of NGS as a must have proach to surveillance/monitoring and the arcity of resources

lotal

### **Courses on Bioinformatics for NGS data mining**

#### 17-18/10/2019

4th Joint Training Course on molecular typing EURL Lm - EURL VTEC - EURL Salmonella Basic course on Bioinformatics tools for WGS data analysis (BioNumerics, SeqSphere and ARIES)

#### @EURL Lm (ANSES)











### **Courses on Bioinformatics for NGS data mining**

- 22 participants in 2015 (7 NRLs, EFSA, OLs)
- 22 participants in 2016 (6 NRLs, EFSA, ECDC, OLs)
- 11 participants in 2017 (9 NRLs, OLs)
- 16 attendees in 2018 (6 NRLs, Ols)
- 14 participants in 2019 (EC, 10 NRLs, EFSA) OUTBREAK INVESTIGATION EXERCISE



The National Agency of Health is investigating a cluster of HUS cases associated with STEC O26:H11 ST29 *eae+ stx*2+, identified through the analysis of whole genome sequencing (WGS). The whole genome sequences of two STEC strains isolated are annexed to this notice (Strain1 and Strain2).

Scenario 1

There have been three laboratory confirmed cases sharing less than 10 allelic differences with cgMLST analysis.

Onset dates range from 4 June 2019 to 12 June 2019. The cases were dispersed across two regions (South East and East of the country). All confirmed cases are female (100%) with an age of 2, 3 and 4 years.

The cluster identified displays an allelic profile that had never been detected in the country previously.

The epidemiologic questionnaire highlighted the consumption of a mixed salad with cheese and ground beef as the possible source of the infections. Your NRL was asked to support the national investigation by testing food and providing evidence of possible correlations of the STEC isolated from food with the outbreak strains.

#### WGS of the STEC strains isolated during the investigation:

÷

Strain ID	Source
Strain3	Mixed salad with cheese
Strain4	Raw milk cheese of the type present in the
	salad (batch A)









OLALLIS SUPERIORS

DI SANT