Proposed common WGS-based genome analysis methods and standard protocols for national surveillance and integrated outbreak investigations of *Pseudomonas aeruginosa* and *Acinetobacter baumannii*"

Valeria Bortolaia

Funded by the European Union

Objectives

To obtain **feed-back** on the document recently sent out for a 2-week written consultation within the network:

Proposed **common WGS-based** genome analysis methods and standard protocols for national surveillance and integrated outbreak investigations of ***Pseudomonas aeruginosa*** and ***Acinetobacter baumannii***

Funded by the European Union

Outline

1. Walk-through the guidance document

2. Open discussion

Funded by the European Union

DNA extraction and QC

- Pure cultures
- Commercial kits (no boiling lysis)
- Considerations about plasmid DNA

DNA concentration and dilution

- Commercial instruments and kits (use of a specific concentration of DNA is crucial for genomic library preparation)

Funded by the European Union

Library preparation and DNA sequencing

- Commercial kits
- Protocols with preparation guidelines for specific library kits and guidelines for sequencing on the specific machinery are frequently updated and available on the Illumina website

Raw reads extraction, QC and filtering

- FASTQ files
- Minimum coverage of 50X would be ideal in public health settings
- Contamination checks (open source software)
- Trimming (open source software)

Funded by the European Union

Genome assembly and QC

- FASTA files (open source software; it is crucial that the assembly pipeline includes post-assembly corrections by mapping of the reads to the assembly and updating the consensus sequence)
- QC (open source software)
  - Number of contigs should be less than 500
  - Genome size should be within 10% of deviation of the expected genome size
  - Etc.

Bacterial species identification and QC

- Curated bioinformatics tools to perform species identification and ensure that QC thresholds specific for the selected tool are fulfilled

Funded by the European Union

Bacterial isolate typing

- MLST
- Detection of PM and genes mediating AMR
  - Curated tools
  - Species-specificity
  - Version #

Funded by the European Union

Cluster analysis and quality thresholds

- SNP (reference genome)
  - For *A. baumannii*, a genetic relatedness threshold of **≤ 2-3 SNPs** has been suggested to distinguish non-outbreak from outbreak strains
  - For *P. aeruginosa*, a genetic relatedness threshold of **≤ 5 SNPs** has been suggested to distinguish non-outbreak from outbreak strains
  - Higher thresholds (e.g. **up to 25 SNPs**) should be examined too
  - At least 90% of each query genome should have been included in the alignment to create the distance matrix

Funded by the European Union

Cluster analysis and quality thresholds

- cgMLST
  - For *A. baumannii*, laboratory-validated allelic thresholds of relatedness are:
    - **≤ 9** allelic differences: related
    - 10 to 200 allelic differences: possibly related
    - > 200 allelic differences: unrelated.
  - For *P. aeruginosa*, an allelic threshold of **≤ 12** allelic differences could identify epidemiologically linked isolates
  - At least 90% of the cgMLST loci present in the scheme must be assigned to each isolate being compared

Funded by the European Union

Data and metadata storage

- Raw reads/assemblies/bioinf. results
- Minimum metadata parameters. Examples of minimum fields are: bacterial species, sample collection date, type of clinical specimen, antimicrobial susceptibility test results, storage location. Patient data, epidemiological and clinical data and hospital data would ideally be linked to isolate data. An example of data useful from a European surveillance perspective can be found in the Protocol-genomic-surveillance-resistant-Enterobacteriaceae

Funded by the European Union

Open discussion