DNA SEQUENCING — DAY 2			
Time	**Activity Description**	**Intended Learning Outcomes** *After completion, trainees will (be able to):*	**Relevance** *Why this is important for you as:*
0930-1100	Sequencing Technologies (Chiara Crestani)	Presentation of the different sequencing technologies and of the latest advances in NGS Compare and understand the advantages and disadvantages of each technology Presentation of the basic vocabulary and concepts of genomic sequencing	**Bioinformaticians** must recognise the best NGS technology to fit their needs for bioinformatic analyses for public health surveillance. **Epidemiologists** will become familiar with the type of information that can be obtained from NGS data **Microbiologists** will be more familiar with the advantages and disadvantages of the different NGS methods, and will be able to select the most suitable one to answer their public health needs.
1110-1155	Galaxy Platform (Remi Planel)	Know how to navigate the Galaxy Platform and using its basic functionalities Being able to perform data analysis and construct workflows	**Bioinformaticians** must familiarize themselves with the Galaxy Platform and its basics to leverage its functionalities for data analysis and workflow construction, enabling streamlined genomic analysis. **Microbiologists** must understand the fundamentals of the Galaxy Platform to access its tools and workflows for processing and interpreting genomic data, facilitating accurate taxonomic classification and comparative genomics studies. **Epidemiologists** must leverage the capabilities of the Galaxy Platform to review the existing resources and pipelines , aiding in the investigation of pathogen transmission dynamics and disease outbreaks.

1400-1530	Raw data and Assembly (Nabil-Fareed Alikhan)	Learn to handle raw sequence data, perform quality assessment using fastQC Know about processing steps such as merging, error correction, and trimming to improve data quality How to read and interpret a QC report Gain insights into contamination detection Learn to recognize false SNPs and poor quality assemblies Understanding the impact of poor data quality on epidemiological inferences	**Bioinformaticians** should master raw data quality assessment and processing, including tools like fastQC, to ensure precise genomic data analysis. **Microbiologists** must understand the impact of DNA library quality and sequencing quality for accurate data generation. **Epidemiologists** must familiarize themselves with these concepts to ensure integration of high-quality NGS data with public health information.

Details

Sequencing technologies

Additionally, sequencing technologies will be explained in detail, comparing and discussing advantages and disadvantages of the different platforms such as Sanger, Illumina, and Nanopore (R9 vs R10). Basic concepts and vocabulary for NGS data will be covered, including the definitions of fastq and fasta files, single-end (SE) and paired-end (PE) sequencing, depth of coverage, long-reads and short-reads.

Through these sessions, participants will gain a thorough understanding of the principles, limitations and practical applications of microbial genomics and sequencing technologies in contemporary research and public health practices.

Galaxy Platform

This course will provide participants with an understanding of the Galaxy Platform and its fundamental functionalities. Participants will learn how to navigate the platform, access tools and workflows, and perform basic data analysis tasks. Topics covered include an introduction to the Galaxy interface, uploading and managing datasets, utilizing QC or assembly analysis tools and workflows, and interpreting analysis results. By the end of the course, participants will have gained the skills and confidence needed to effectively use the Galaxy Platform for their bioinformatics research and data analysis needs.

Raw data and Assembly

Participants will be exposed to the principles of handling and assessing raw sequence data, pre-processing steps like read trimming, and quality assessment using tools such as fastQC.

In the second part of the course, attendees will learn about the impact of genome data and assembly quality , including contamination, on epidemiological interpretation of genomic data.