Deliverable

13.4 E-learning Materials

Version: 01

Due: Month 12 (delayed until M60)

Completed: Month 60

Author: Jeffrey Skiby, DTU

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 643476.

Contents

Deliverable Description

This Deliverable is a report on the e-learning materials developed, updated, or compiled under the COMPARE Consortium and in tandem with other supported projects to further the vision of COMPARE. These materials may include SOPs, LOPs, guidance documents, manuals, instructional videos and massive open online courses (MOOCs).

Though development and work on these many types of materials took place with participants from several workpackages, the reporting of this work comes under workpackage 13 (Dissemination and Training).

In this report, we will describe the e-learning courses available via the COMPARE website, as well as the protocols and SOPs compiled/developed during the project.

We wanted to ensure relevant stakeholders of COMPARE were adequately informed about COMPARE's progress and results and have access to the training they need in order to apply the harmonized workflows, analytical tools and data resources developed and implemented by COMPARE in their pathogen detection and outbreak response activities.

As part of COMPARE's resources being made available to the stakeholders, various e-learning resources were made available as they were developed.

Initially, this Deliverable was planned to be delivered in Month 12 of the project. It was overly ambitious to set the Deliverable deadline for D 13.4 as Month 12. It had taken the first 12 months of the project to get the COMPARE partners aligned on priorities for training and e-learning. The first videos for how to sequence and analyze were under development within the first year of the project. Additional courses were added over the life of the project, and thus it would have been premature to report so early in the project. The delay in reporting did not affect any other work in the project.

Massive open online courses (MOOCs)

MOOCs offer an opportunity to train not only stakeholders, but also interested parties and students. Using an established platform, such as Coursera, allows the researchers to focus on the materials and give feedback to the students; Coursera manages the enrollment and maintenance of the site/platform.

Providing online courses relates to overall goals and objectives of the project of open source tools, open data and open science.

The courses are described below and available via the COMPARE website, https://www.compare-europe.eu/library/e-learning.

Antimicrobial resistance - theory and methods

Launched in June 2016, this MOOC has had almost 3000 students complete the course.

The course covers the topics related to antimicrobial resistance with basic definitions and overview on antimicrobials, their use and the emergence and spread of resistance. The course guides you through the concepts and the importance of resistance spread and dissemination and how that happens. It shows you how bacteria become resistant and which mechanisms they might use for this. The course also provides training in methods for antimicrobial susceptibility testing (AST) and detection of specific resistance in the microbiological laboratories with the basic methods available and with focus on obtaining good quality results that can be interpreted and used for different purposes.

Additionally, it will show you how to use genomic analysis tools to analyze whole genome sequencing data to detect resistance genes (and or other genes of interest) in a simple and easy way using online tools freely available.

Course modules:

- Antimicrobial and antimicrobial action
- Antimicrobials and resistance
- Antimicrobial susceptibility testing
- Interpretation
- Quality assurance
- Alternatives to AST- Genome analysis tools
- Specific resistances

Whole genome sequencing of bacterial genomes - tools and applications

Launched in July 2017, this MOOC has had almost 1000 students complete the course.

The field of monitoring and surveillance of pathogens, their characterization and subtyping as well as phylogeny and screening for antimicrobial resistance are highly relevant for Public Health and for professionals in the medical sectors. This also includes the veterinary field, the environment, pharmaceutical companies and private laboratories, especially those working with institutions or companies performing diagnostics, research or monitoring activities.

This online course includes video lectures about bacterial subtyping and applications of next generation technology to perform research and monitoring, and also demonstrations and guided exercises on the use of online tools.

The lectures provide a background for WGS analyses and present the Center for Genomic Epidemiology (CGE) tools, which are available online and can be used widely and by non-bioinformaticians.

The lecturers in this course include both microbiologists and bioinformaticians. The lecturers themselves use the WGS technologies for both research and monitoring purposes and the bioinformaticians have been involved in building these tools and pipelines in addition to teaching common users in a very practical way.

The course includes the following Modules: has the following videos:

Module 1
- General Principles of typing of bacteria
- Surveillance of Antimicrobial Resistance using whole genome sequencing

Module 2
- Application of Genomic tools- One technology takes it all
- Introduction to Next-Generation sequencing (NGS)
- De novo assembly- from raw reads to contigs: assembler tool description and application

Module 3
- Species identification: KmerFinder tool description and applications
- MLST Typing: MLST tool description and applications
- Resistance gene detection: Resfinder tool description and applications

Module 4
- Salmonella Serotype identification: SeqSero tool description and applications
- E. coli serotype identification: serotype finder tool description and applications
- Plasmid replicon identification and plasmid typing

Module 5
- Bacterial analysis pipeline- batch upload
- Phylogenetic relatedness: CSI Phylogeny tool description and applications
- Multipurpose detection of genetic markers- My Db finder tool description and applications

Metagenomics applied to surveillance of pathogens and antimicrobial resistance

Launched in May 2018, this MOOC has had more than 150 students complete the course.

The field of metagenomics and whole community sequencing is a promising area to unravel the content of microbial communities and their relationship to disease and antimicrobial resistance in the human population. Bioinformatic tools are extremely important for making sense out of metagenomics data, by estimating the presence of pathogens and antimicrobial resistance determinants in complex samples. Combined with relevant explanatory data, metagenomics is a powerful tool for surveillance.

In this course, metagenomics applied to surveillance of pathogens and antimicrobial resistance, we teach about the potential of metagenomics for surveillance and give the learners an overview of the steps and considerations in a metagenomics study.

The course has the following modules:

From sampling to sequencing

- In this module, metagenomics are introduced, some of the considerations and controls that need to be in place in a metagenomics study, and to the topic of antimicrobial resistance. You will also learn about: 1) Sampling and sample handling - the considerations behind a sampling plan, how to perform sampling in practice, and how sample storage can affect metagenomics results; 2) DNA and RNA extraction methods - both for bacterial and viral microorganisms; and 3) Sequencing - from library preparation to the basics of different sequencing technologies.

From reads to results

- In this module, the basics of bioinformatics analysis of metagenomics data are introduced, including the different types of analysis possible and the different algorithms available. You will then learn about quality control, MGmapper and KRAKEN (two freely available bioinformatics pipelines), and ResFinder (a database of antimicrobial resistance genes). An example from a real study will demonstrate some of the challenges of a bioinformatics analysis.

Interpretation of results and potential of metagenomics for surveillance

- In this module, two different approaches to analyse and interpret sequence reads - classification and assembly are reviewed. You will see examples of methods to visualize read counts and to analyse metagenomics together with explanatory data. Last, you will learn the potential of metagenomics for the development of a future global and integrated surveillance, and the challenges you may encounter during that process.

Next Generation Sequencing of Viral Genomes

We are currently developing the course “Next Generation Sequencing of Viral Genomes” with support of the Erasmus University community learning innovation (CLI). When finalized, it will made available via the COMPARE website as well as through Coursera.

In the field of next generation sequencing (NGS) there are many new and exciting developments, but it can be difficult to determine which technique is best suited for your pathogen, budget and research question. During the course, the student will learn about the different sequencing options that are currently available, their opportunities and challenges. The course includes both theory and methods and is developed for a broad audience, from master students who want in-depth knowledge on the latest deep sequencing technologies to researchers that want to set up/perform NGS in their laboratory.

The course modules are divided in the following topics:
Sample preparation

- In this module, we discuss the challenges in sample preparation for NGS of viral genomes. Which method is optimal depends on the target virus (RNA/DNA/Enveloped/Non-enveloped), sample matrix, host NA, abundancy of the target virus and many other variables. We will discuss methods for the clinical workflow, food, tissues and environmental samples for several types of viruses.

Sequencing options

- This modules gives an overview of current NGS techniques used in virology. We will discuss the opportunities and challenges (including the costs) of the different approaches, and how the method selection (amplicon, other ways of targeting sequencing, agnostic sequencing) is dependent on the questions at hand. We will further discuss implementation of NGS in a diagnostic setting and the progress in field sequencing.

Analytical workflows

- Data analyses of NGS runs can be complicated. We will provide a clear overview of the different steps in the analysis, available software and common pitfalls. We will further discuss the importance of reference databases for analyses of NGS data.

In-depth analyses of NGS data

- Finally we will provide an overview of techniques used for in-depth analyses of NGS data. We will discuss the application of NGS in genomic epidemiology, cluster analysis methods used in outbreaks and phylogenetic reconstruction.

Protocols and Standard Operating Procedures

In addition to the MOOCs made available via the COMPARE website, protocols and standard operating procedures (SOPs) have been reviewed and compiled for the COMPARE vision.

Protocols

An inventory of existing, and where possible harmonised, protocols has been developed in order to map the types of samples that are currently recommended at the EU or international level for known diseases of public and veterinary health importance. This approach provides a foundation for the development of novel risk-based sampling algorithms and protocols described in Workpackage 1, Risk assessment and risk-based strategies for sample and data collection. It represents a first step in achieving the COMPARE objective to improve the rapid identification of emerging infectious diseases, and allows for the exploration of novel integrated forms of surveillance for emerging infectious disease (EID). Understanding the extent of current sampling of food, human and animal populations through this inventory of known, and in some cases, notifiable pathogens helps to predict the characteristics of samples that are likely to be supplied or made available through existing surveillance systems. If known/notifiable pathogens are not identified upon conventional analysis of these samples, the samples may become the substrate for NGS laboratory protocols that are the subject of COMPARE Workpackage 2, Harmonised standards for sample processing and sequencing.

The inventory of protocols is available via the COMPARE website, https://www.compare-europe.eu/library/epidemiological-datasets

Standard Operating Procedures

Workpackage 2 in COMPARE was tasked to develop harmonized analytical workflows for generation of high-quality NGS data in combination with relevant metadata for pathogen detection and typing across sample types, pathogens and domains.

A collection of protocols based on those collected from the partners of COMPARE is available via the COMPARE website, https://www.compare-europe.eu/library/protocols-and-sops .

The protocols provided can be either Laboratory Operating Procedures (LOPs) or Standard Operating Procedures (SOPs).

LOPs are those protocols that have been used at least by the provider, but as yet have not been validated in detail.

SOPs are protocols that have been validated in detail and are already published. In some cases, different matrices and different pathogens (such as virus, bacteria, and/or parasite) have been used for validation

The SOPs/LOPs are divided between Metagenomics and Single isolates.