DISPATCHES

Go COMPARE!

"One serves all" next-generation sequencing framework

Amie Adkin[1] , Claudia Wylezich[2], Martin Beer[2], Jeffrey Skiby[3], Marion Koopmans[4,5], Frank Aarestrup[3].

COMPARE (Collaborative Management Platform for detection and Analyses of [Re-] emerging and foodborne outbreaks in Europe) is a multidisciplinary research network that is set up with the common vision of becoming the enabling analytical framework and globally linked data and information sharing platform system for the rapid identification, containment and mitigation of emerging infectious diseases and foodborne outbreaks.

The system sets out to integrate state-of-the-art strategies, tools, technologies and methods for collecting, processing and analysing sequence-based pathogen data in combination with associated (clinical, epidemiological and other) data, for the generation of actionable information to relevant authorities and other users in the human health, animal health and food safety areas.

Next-generation sequencing (NGS) used for whole genome sequencing (WGS) or whole community sequencing (WCS or metagenomics) enables generation of complete genomic information from the isolate or sample, independent of both the sector (public health, veterinary health or food safety), and the type of pathogen (viruses, bacteria or parasites). The outputs (sequence data) provide one common language that can be exchanged and compared between laboratories and over time, in combination with other associated data defined here as "metadata" including contextual data (*e.g.* data on sample type and process, clinical, microbiological, epidemiological and other data), primary data (raw sequence reads) and derived data (*e.g.* genomic alignments of reads, assemblies and functional annotation data

Keywords

- Data sharing platform
- Emerging infectious diseases
- Foodborne outbreak
- Laboratory proficiency testing
- Research network
- Sequence analysis

[1] Animal and Plant Health Agency (APHA), Department of Epidemiological Sciences, Weybridge, KT15 3NB, UK.

[2] Friedrich Loeffler Institut (FLI), Institute of Diagnostic Virology, 17493 Greifswald-Insel Riems, Germany

[3] National Food Institute, Technical University of Denmark (DTU Food), Research Group for Genomic Epidemiology, 2800 Kongens, Lyngby, Denmark

[4] National Institute for Public Health and the Environment (RIVM), Centre for Infectious Diseases Research, Diagnostics and Screening (IDS), 3720 BA Bilthoven, The Netherlands

[5] Erasmus MC, Department of Viroscience, 3015 CN Rotterdam, The Netherlands

* Corresponding author : amie.adkin@apha.gsi.gov.uk

sets). COMPARE aims to harness the rapid advances in these technologies to improve identification and mitigation of emerging infectious diseases and foodborne outbreaks.

To this end, COMPARE is establishing a “One serves all” analytical framework – as depicted in Figure 1 on the following page – presenting the different components of the COMPARE Analytical Framework.

COMPARE runs from 1 December 2014 to 30 November 2019. The consortium (Table 1) has been awarded EUR 20 million funding under the European Union’s Horizon 2020 research and innovation programme under grant agreement No 643476.

Table 1 / The COMPARE consortium consists of the following members:

1	Technical University of Denmark (DTU), Denmark	16	University of Cambridge, UK
2	Erasmus University Medical Center (Erasmus MC), The Netherlands	17	Tierärztliche Hochschule Hannover (TIHO) / The University of Veterinary Medicine Hannover (TiHo), Germany
3	Statens Serum Institute (SSI), Denmark	18	Universidad de Castilla- la Mancha (UCLM), Spain
4	Friedrich-Loeffler-Institut (FLI), Germany	19	Fondation Mérieux, France
5	Agence nationale de sécurité sanitaire de l’alimentation, de l’environnement et du travail / French Agency for Food, Environmental and Occupational Health & Safety (ANSES), France	20	Aristotelio Panepistimio Thessalonikis / Aristotle University of Thessaloniki (AUTH), Greece
6	Robert Koch-Institute (RKI), Germany	21	Institut Français de Recherche pour l’Exploitation de la Mer / French Public Institute for Marine Research (IFREMER), France
7	European Molecular Biology Laboratory (EMBL), UK	22	Erasmus University Rotterdam, The Netherlands
8	Istituto Superiore di Sanità (ISS), Italy	23	The Australian National University (ANU), Australia
9	Rijksinstituut voor Volksgezondheid en Milieu / National Institute for Public Health and the Environment (RIVM), The Netherlands	24	Magyar Tudomanyos Akademia Wigner Fizikai Kutatokozpont / Wigner Research Centre for Physics, Hungarian Academy of Sciences, Hungary
10	Animal and Plant Health Agency, UK	25	Civic Consulting GmbH (CIVIC), Germany
11	University of Edinburgh (UEDIN), UK	26	Responsible Technology (RT), France
12	University of Bonn Medical Centre (UK-Bonn), Germany	27	University of Bologna (UNIBO), Italy
13	Academisch Medisch Centrum Universiteit van Amsterdam / Academic Medical Center (AMC), University of Amsterdam, The Netherlands	28	Leibniz Institut Deutsche Sammlung von Mikroorganismen und Zellkulturen / German Collection of Microorganisms and Cell Cultures (DSMZ), Germany
14	Universiteit Antwerpen / University of Antwerp (UA), Belgium	29	Wellcome Trust Sanger Institute (WTSI), UK
15	Artemis One Health Research Institute (Artemis), The Netherlands		

DISPATCHES

Within the project, the first two work packages are researching approaches to support how more efficient risk-based sampling can be carried out, and the identification and harmonisation of laboratory protocols for samples intended for next-generation sequencing. All work packages are listed in Table 2.

Risk-based sampling inventory

WP1 leader: Dr Amie Adkin (amie.adkin@apha.gsi.gov.uk)

Understanding the extent of existing sampling protocols for generating clinical and diagnostic data arising from food, human, livestock and wildlife populations helps to predict the characteristics of samples that are likely to be supplied or made available through existing surveillance systems. We have been developing an inventory of existing, and where possible harmonised, protocols 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. Work is underway to make this list openly available to other researchers. The work required was across the different disciplines of human clinical data and the equivalent animal information. This was accessed through the various websites of multinational organisations and EU-FP7 projects, and gave rise to some surprising comparisons. Firstly, accessing livestock information was, overall, easier and these datasets were structured more logically than those held for human health. However, for all areas, information was dispersed and sometimes incomplete, leading to a feeling of a treasure hunt! The accessibility and coverage of datasets made available from several organisations was commended including the World Organisation from Animal Health (OIE) and US Centres for Disease Control and Prevention (CDC). Whilst the inventories are still being completed, they have already proven useful for one of the contributors currently working at Food and Agriculture Organization of the United Nations (FAO).

Optimising and harmonising handling protocols

WP2 leader: Prof Dr Martin Beer (martin.beer@fli.de)

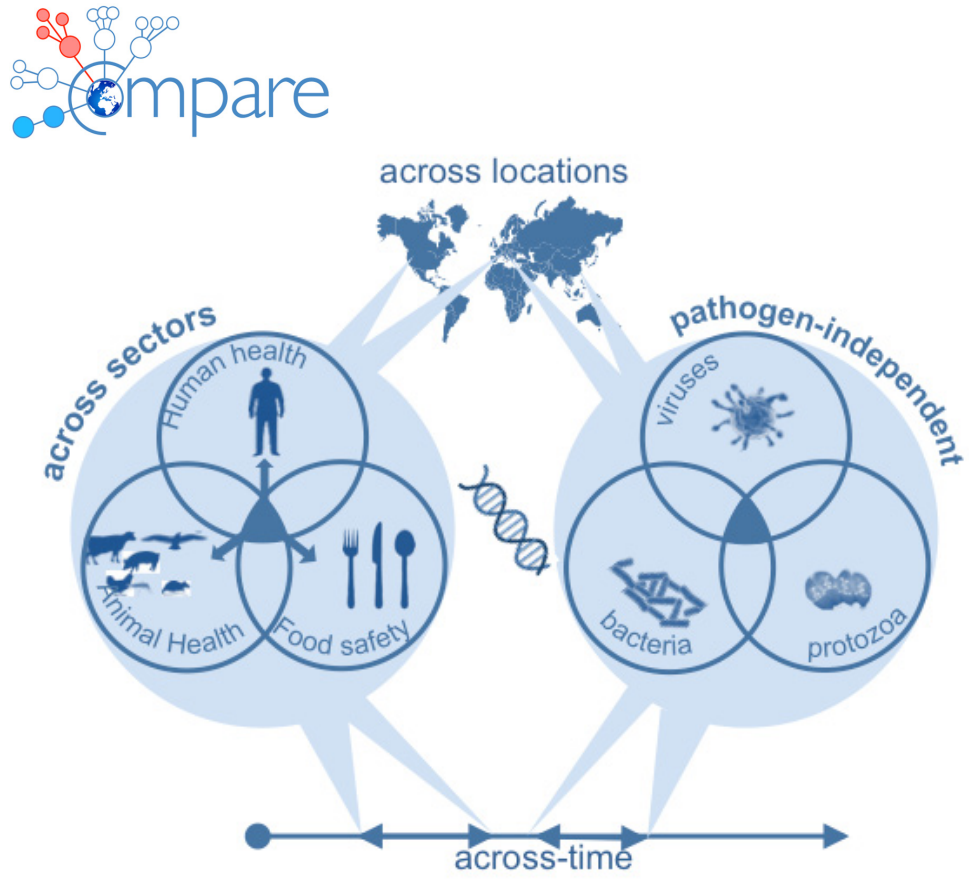
Careful sample handling is a crucial step in gaining high-quality information from next-generation sequencing to ensure maximum benefit for clinical and public health. Samples have to be treated with great care to minimise significant shifts in the microbial community composition of samples during transportation. This is an important prerequisite in order to display the initial sample situation within the sequencing outcome, and to successfully detect causative agents via sequencing. Therefore, work package 2 (WP2) is addressing the harmonisation of standards for sample handling as a basis for other tasks in the COMPARE project. During the first year of the project, an inventory of commonly used protocols with respect to collection, handling, transport and storage of samples was conducted via a survey. Based on survey results, experiments were designed to investigate the influence of various treatments and handling procedures such as fixation, storage temperature and duration on different sample matrices such as tissue, body fluids, faeces, sewage samples, as well as ticks and insects containing pathogens.

In parallel, sample-processing pipelines including pathogen inactivation, nucleic acid extraction, and subsequent processing until sequencing were developed and are being intensively tested for different matrices (*e.g.* tissue, ticks, bacterial suspensions, and food samples). Protocols providing best results regarding quality and quantity of extracted nucleic acids as well as sequence reads have already been disseminated in the form of Laboratory Operating Procedures (LOPs) for their further review and application in the laboratories of COMPARE members.

DISPATCHES

These sample-handling experiments will be tested and sample-processing pipelines further refined during the coming months. The ambitious aim of these optimisation steps is to develop and to provide one protocol for all samples for metagenomics.

Figure 1 / The COMPARE analytical framework



For more information, contact:
Prof. Dr. Frank Aarestrup of DTU, fmaa@food.dtu.dk

Prof. Dr. Marion Koopmans of Erasmus MC: m.koopmans@erasmusmc.nl

DISPATCHES

Table 2/ COMPARE Work Packages, WP leaders and overall tasks

Work Package	Title	WP Leader and Co-leader	Overall task of WP
WP1	Risk assessment and risk-based strategies for sample and data collection	WP leader: Amie Adkin, DEFRA/APHA (10) amie.adkin@apha.gsi.gov.uk Co-leader: Christian Gortazar, UCLM (18) christian.gortazar@uclm.es	To develop risk assessment models and risk-based sampling and data collection strategies for NGS-based analyses of foodborne and (re-) emerging infections.
WP2	Harmonised standards for sample processing and sequencing	WP leader: Martin Beer, FLI (4) martin.beer@fli.bund.de Co-leaders: Simone Caccio, ISS (8) simone.caccio@iss.it	To develop harmonised 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
WP3/6	From comparable data to actionable information: Analytical workflows for frontline diagnostics	WP leader: Surbhi Malhotra, UA (14) surbhi.malhotra@uantwerpen.be Co-leaders: Constance Schultsz, AMC (13) c.schultsz@gmail.com, Anne Pohlmann, FLI (4) anne.pohlmann@fli.de	To develop an analytical workflow for the use of single isolate and metagenomic NGS in human and veterinary clinical microbiology. To assess the feasibility of NGS/WGS/WCS for clinical diagnostic use and hospital epidemiology
WP4/7	From comparable data to actionable information: Analytical workflows for foodborne pathogen surveillance, outbreak detection and epidemiological analysis	WP leader: Eva Møller Nielsen, SSI (3) emn@ssi.dk Co-leaders: Tine Hald, DTU (1) tiha@food.dtu.dk, Michel-Yves Mistou, ANSES (5) michel-yves.mistou@anses.fr	To develop a general analytical workflow for population-based disease surveillance, outbreak detection and epidemiological modelling of foodborne infections.
WP5/8	From comparable data to actionable information: Additional tools for detection of and response to (re-) emerging infections	WP leader: Ron Fouchier, EMC (2) r.fouchier@erasmusmc.nl Co-leader: Mark Woolhouse, UEDIN (11) mark.woolhouse@ed.ac.uk	To develop cross-sector and cross-pathogen methods for support of emerging pathogen identification and characterisation in support of outbreak investigations and epidemiological analysis.
WP9	COMPARE data and information platform	WP leader: Guy Cochrane, EMBL (7) cocharne@ebi.ac.uk Co-leaders: Ole Lund, DTU (1) lund@cbs.dtu.dk, Istvan Csabai, WIGNER (24) csabai.istvan@wigner.mta.hu	Support rapid sharing, integration and analysis of sequence-based pathogen data in combination with other contextual metadata; the system will be linked to existing and future complementary systems, networks and databases such as those used by ECDC, NCBI and EFSA.
WP10	COMPARE risk communication tools	WP leader: Emilio Mordini, RT (26) emilio.mordini@responsibletechnology.eu	To design and develop appropriate risk communication tools and strategies for stakeholders.
WP11	User consultations	WP leader: Marion Koopmans, EMC (2) m.koopmans@erasmusmc.nl Co-leader: Frank Aarestrup, DTU (1) fmaa@food.dtu.dk	To design the COMPARE systems' analytical workflow and its main components based on the expert inputs and associated information needs of its intended future users and other stakeholders.
WP12	Barriers to open data sharing	WP leader: George Haringhuizen, RIVM (9) george.haringhuizen@rivm.nl Co-leader: Jørgen Schlundt, DTU (1) jschlundt@ntu.edu.sg	To identify, clarify and, as far as feasible, develop practical solutions for Political, Ethical, Administrative, Regulatory and Legal (PEARL) barriers that hamper the timely and open sharing of data through COMPARE.
WP13	Dissemination and training	WP leader: Frank Aarestrup, DTU (1) fmaa@food.dtu.dk Co-leader: Marion Koopmans, EMC (2) m.koopmans@erasmusmc.nl	To ensure that relevant stakeholders of COMPARE are adequately informed about COMPARE's progress and results and have access to the training they need in order to apply the harmonised workflows, analytical tools and data resources developed and implemented by COMPARE in their pathogen detection and outbreak response activities.
WP14	Cost-effectiveness framework	WP leader: Pieter van Baal, EUR (22) vanbaal@bmg.eur.nl Co-leader: Frank Allewedlt, CIVIC (25) alleweldt@civic-consulting.de	To develop a standardised framework for estimating the cost-effectiveness of the COMPARE system and related methods and tools, including the value of safety.
WP15	Management	WP leader: Frank Aarestrup, DTU (1) fmaa@food.dtu.dk Co-leader: Marion Koopmans, EMC (2) m.koopmans@erasmusmc.nl	To implement the appropriate organisational structures and processes to ensure COMPARE's compliance with the EU Grant Agreement and the COMPARE Consortium Agreement (CA).