ENGAGE Workshop and Training courses

Establishing Next Generation sequencing Ability for Genomic analysis in Europe (ENGAGE)

ENGAGE Workshop, 10-11 October 2016

Venue: NIPH-NIH, Warsaw, Poland

Attendees:
Alessia Franco (IZSLT, Italy)
Anthony Underwood (PHE, UK)
Antonio Battisti (IZSLT, Italy)
Bavo Verhaegen (WIV-ISP, NRL Belgium)
Beatriz Guerra Román (EFSA)
Burkhard Malorny (BfR, Germany)
Dariusz Wasyl (PIWet, Poland)
Eleonora Mastrorilli (IZSVe, Italy)
Kinga Wieczorek (PIWet, Poland)
Liljana Petrovska-Holmes (APHA, UK)
Manuel Francisco Hernandez (MAGRAMA, NRL Spain)
Maria Borowiak (BfR, Germany)
Michel-Yves MISTOU (ANSES, France)
Mirko Rossi (VETMED, Finland)
Pimlapas Leekitcharoenphon (Shinny) (DTU, Denmark)
Raf Winand (Scientific Institute of Public Health, Belgium)
Rafał Gierczyński (NIPH-NIH, Poland)
Rene S. Hendriksen (DTU, Denmark)
Sara Petrin (IZSVe, Italy)
Susanne Karlsmose Pedersen (DTU, Denmark)
Tomasz Wołkowicz (NIPH-NIH, Poland)
Valeria Michelacci (EURL-VTEC, Italy)

Reference:
As indicated below the presentations are made available on 'sharepoint' which refers to https://share.dtu.dk/sites/Engage_149200/SitePages/Home.aspx [login information sent to partners on email 9th June 2016. If you experience problems accessing the site, please contact suska@food.dtu.dk]

Meeting notes:

- Welcome and introduction to the ENGAGE workshop 2016 (Tomasz Wolkowicz)

- Update, ENGAGE (Rene Hendriksen, DTU Food) [presentation available on sharepoint]
    - Status of ongoing tasks presented.
    - Importance of meeting deliverables.
    - Risks for the project not succeeding were indicated, e.g. minimal synergy with other ongoing activities.
    - The difficulties due to avoiding duplications with tasks performed within other EU funded projects were commented. EFSA underlined the need to follow the signed agreement. Any changes made to improve the performance of the project (duplications/solutions not previously identified) should be communicated to EFSA.

- EFSA, WGS for Food Safety (Beatriz Guerra, European Food Safety Authority) [presentation available on sharepoint]
    - Presentation of the current and planned EFSA activities within the field of WGS.
    - Requirements from EFSA for the participants of the ENGAGE project:

- To deliver results, respect deadlines, inform/ask for permission EFSA about changes on the granted project proposal: personal, budget, deliverables, etc.; to inform/ ask EFSA for permission for any publication of results (manuscript, posters, etc).
- INNUENDO-ENGAGE Coordination meetings are organised regularly.
- Also, the EC is collecting information from the MS in a questionnaire to explore the current ongoing work in relation to WGS in MS.

- 'Tour de partner' (Chair: Rene Hendriksen, DTU Food) – Each partner institution (Availability of strains related to the target list; Number of strains sequenced; Deliverables completed) [presentations available on sharepoint]
  - DTU Food, Denmark – Till now, sequenced 153 *Salmonella/E.coli*. Another 96 will be sequenced in 2016.
  - IZSVe, Italy – Till now, sequenced 78 *Salmonella* (Derby, Napoli, monophasic Typhimurium)
  - APHA, UK – All target *Salmonella* serovars are available. Till now, sequenced 94+125+259 strains.
  - PIWet, Poland – From the target list; many isolates available. Till now, sequenced 21 strains (during twinning). Planned another 2x96 in 2016.
  - IZSLT, Italy – Till now, sequenced 93 *Salmonella* (mostly Infantis) and *E. coli*
  - NIPH, Poland – Till now, sequenced 7 strains (during twinning). More planned during the next month.
  - PHE, UK – Sequence all *Salmonella* we receive. Till now, sequenced 11,000 *Salmonella* and 3,600 *E. coli*. Referred to the website https://omictools.com/ which might be a useful reference.
  - BfR, Germany – Till now, sequenced 131 *Salmonella* strains.
  - Additional participants:
    - France (ANSES): Sequences internally at ANSES and is outsourcing. Also at a hospital.
    - Spain (MAGRAMA): Not sequencing at the moment.
    - Belgium (WIW-ISP): Performs MiSeq-sequencing.

- Projects; data availability, status (Chair: Liljana Petrovska, APHA) [Presentation available on sharepoint as part of the 'Tour-de-partner'-presentations]
  - Benchmarking of AMR (APHA, PHE, DTU) – no project description drafted yet
  - Benchmarking of serotyping (SeqSero, MOST) (APHA, PHE, DTU, INNUENDO) – no project description drafted yet
  - Colistin resistance *mcr*-1 gene in *E. coli* (DTU, NVRI, IZSLT, PIWet, maybe also APHA?! (Liljana will get back after she has checked), also BfR?! (Burkhard will ask Anne Marie)) - plasmids in *E. coli* harbouring *mcr-1*
  - *S.* Napoli (IZSVe, BfR) – Epidemiology study; comparative genomics
  - *S.* Derby (IZSVe, BfR, DTU, APHA, PIWet, NIPH-NIH, IZSLT)
  - *S.* Typhimurium Monophasic strains (IZSVe) – clarification pending
  - *Salmonella* spp. from pets and rare serotypes (NVRI, DTU, APHA)
  - *S.* Paratyphi v. Java (BfR, DTU, PHE) – plasmid diversity, phylogeny
  - *S.* Infantis (IZSLT, PIWet, APHA, NIPH-NIH, BfR, DTU) – focus at co-resistance

Action item: each project lead takes action to start out and run the project

- Benchmarking; plan (Chair: Anthony Underwood) [presentation available on sharepoint]
  - Serotyping version 1:
    - Discussion of the necessary sensitivity of the test (SeqSero, MOST, SalmonellaTypeFinder)

- **Action item:** Mirko (INNUENDO coordinator), will test the strains using SISTR. Anthony and Liljana will inquire if any from PHE or APHA wants to lead the publication of the data.

  - Serotyping version 2:
    - Discussion on how many/which serotype we ***must*** be able to identify for the 2nd serotyping exercise.

- We need this list to be able to proceed with the benchmarking, e.g. top50 including regulated serovars
- The EFSA/ECDC zoonosis report to be published in December will show the most prevalent serotypes both in the human and animal side. This data could be useful for the top50 serovars list.

Action item: We will approach ECDC (Anthony/Kathie), EFSA (Liljana -> Rob/Kirsten) and the EURL-Salm (Liljana -> Kirsten) to obtain a top50 list. Liljana will write the exercise outline.

- AMR:
  - 414 strains already selected. More could be added from PHE.
  - Discussed whether it should be divided in the different antimicrobial classes.
  - Discussed whether we could run some of the testing command-line (the full set of selected strains) and then a subset web-based (e.g. 20-30 of each of *E. coli/Salmonella*)
  - Additional discussion items:
    - What samples should be analysed?
    - How many?
    - What resistances should be included?
    - Are there other tools that should be included?

Action item: Liljana writes up the outline of the AMR benchmarking

Future benchmarking exercises will be discussed during a future call.

- About the GMI PT 2015 incl. obtained results (Rene Hendriksen, DTU Food [presentation available on sharepoint]
  - Introduction to the GMI PT, wet-lab and dry-lab component
  - Report related to the wet-lab component will be compiled in a report (pending), the report related to the dry-lab component is available for download on. EFSA asked if the ENGAGE partners would get a separate report on their performance. DTU confirmed that this is the intention.
  - http://www.globalmicrobialidentifier.org/workgroups/gmi-proficiency-test-reports
  - Discussions about the results obtained from the dry-lab component whether it is possible to define expected results.

- About the PHE integrated pipeline (Anthony Underwood, PHE) [presentation available on sharepoint]
  - Lessons learned when introducing WGS routinely.
  - Currently PHE is sequencing all *Salmonella* they receive (two HiSeq-runs per week; 4 flow-cells)
  - K-mer based identification is used for identification of the organism.
  - The bioinformatics analysis is set up in a modular workflow allowing for tools to be included or excluded in a plug and play set-up.

- About INNUENDO; Integrating genomics into food-borne surveillance (Mirko Rossi, Finland) [presentation available on sharepoint]
-
  - Collaboration between ENGAGE and INNUENDO is ongoing, currently related to benchmarking of *Salmonella* serotyping tools. INNUENDO was offered participation on the GMI-PT and THL (the Finnish Public Health Authority) as member of INNUENDO signed up to participate.
  - For the wgMLST scheme, INNUENDO is working with basing a cutoff on the topology structure

- Whole genome sequencing: from library preparation to sequence analysis (Maria Borowiak, BfR) [presentation available on sharepoint]
  - Experiences when reducing the volume to half for the Nextera XT Library preparation; handling is more difficult but the quality of the raw reads is similar to the standard protocol.
  - Experiences when adding more strains to a flowcell (tested 48 strains of *Salmonella*; originally 24). With 48 strains, the number of contigs was considered too high and the N50 was considered too low, so 36 was the compromise.

- Experiences with run failures on the MiSeq
- Experience of the influence of assembler when detecting e.g. resistance genes (Velvet, SPAdes, PATRIC)

- Intro to the updated ENGAGE working space (Pimlapas Leekitcharoenphon (Shinny), DTU Food) [presentation available on sharepoint]
  - The original working space crashed and a new more stable one has been built. Data need to be re-submitted.
  - It was suggested it should be possible to edit metadata subsequent to submission.
  - It was suggested that when data in the working space is being downloaded by a colleague, an email is sent to the original provider for information.
  - Some firewall issues have been experienced.
  - It is the intention that data in the working space will be submitted to ENA by the termination of the ENGAGE project.

  Action item: Shinny follows up on 1) the possibility to edit metadata subsequent to submission, 2) setting up the database so that an email is sent to the original provider of sequences when they are being downloaded by a colleague, and 3) firewall issues.

- Plan for ENGAGE E-learning (Pimlapas Leekitcharoenphon (Shinny), DTU Food) [presentation available on sharepoint]
  - Tools are presented on E-learning and will be launched on Coursera (probably by mid-2017).

- Break-out groups focusing at technical challenges in bioinformatics analysis (Chairs: Rene Hendriksen, DTU Food, Anthony Underwood, PHE)
  - Bioinformatics Infrastructure
    - Federalized or centralized nationally appears as the best solutions for the structure with the analysis locally based
    - Discussion of how to store/share relevant metadata; possibility to store a summary of K-mers
  - Analytical pipelines
    - User limitations were discussed, depending on the bioinformatics knowledge and computational power available
    - Ideally, we would agree on 'winner' tools; tools that can function as reference tools for the purpose of comparison
    - Discussed how many SNPs can be expected on clonal isolates
  - Operational and Legal issues
    - Whole-genome-sequences may be exploited and it is relevant to consider the Intellectual Property of the sequences
    - Discussed that relatedness between phenotype and genotype is a key element

- Discussion: Needs for further twinning? (Chair: Rene Hendriksen, DTU Food)
  - Next twinning round will be in September 2017
  - Consortium is welcomed to send any comments related to twinning needs to rshe@food.dtu.dk

- Future opportunities and AOB (Chair: Rene Hendriksen, DTU Food)
  - If aiming at a common cgMLST, we need a large number of reference sequences
  - Suggested that the reference genomes are stored at the relevant EURLs
  - The intention is that we have a Galaxy pipeline provided by PHE we can use for the purpose of this project
  - Consortium is welcomed to send any further comments related to future opportunities to rshe@food.dtu.dk

- Summary of the ENGAGE workshop (Chair: Rene Hendriksen, DTU Food)
  - Project proposals that was discussed and will be ongoing

- Benchmarking, plan:
  - Benchmarking, serotyping #I
    - Mirko will run SISTR
    - We will look into capturing the findings in a scientific publication
  - Benchmarking, AMR
    - Liljana is in charge of this task.
  - Benchmarking, serotyping #II (1st quarter of 2017)
    - Liljana will draft a set-up based on the ISO 16140
  - Benchmarking, phylogeny (2nd quarter of 2017)
    - Mirko has a dataset which can be the basis

--- --- ---

Establishing Next Generation sequencing Ability for Genomic analysis in Europe (ENGAGE)

Venue: NIPH-NIH, Warsaw, Poland

Summary report, Workshop and Training course 2016

The first ENGAGE Workshop and Training Course was held from 10 to 14 October, 2016 at the National Institute of Public Health – National Institute of Hygiene (NIPH – NIH) in Warsaw.

ENGAGE Workshop, 10-11.10.2016

The objective of the workshop was to update information about project progress, deliverables, identifying any emerging issues with project activities. About 22 participants attended the workshop, including the participants of the ENGAGE projects, representative of EFSA, funding organization and four non-Engage participants (representatives of: EURL-VTEC (Italy), ANSES (France), MAGRAMA (Spain), WIW-ISP (Belgium)).

On the first day of the workshop each partner institutions presented the status of the ongoing task, including information about availability of strains related to target list, number of sequenced so far strains. Benchmarking plan concerning Salmonella serotyping, phylogeny and AMR exercises (Antimicrobial Resistance) was discussed.

Representative of EFSA, (Beatriz Guerra) presented the current and planned EFSA activities in the field of implementation WGS for food safety management. Moreover, requirements for the participants of ENGAGE project were mentioned. Each change on the granted proposal (budget, personal, etc.) or result publishing need previous permission of EFSA.

The second day started with presentation and discussion about the Global Microbial Identifier Proficiency Test 2015 (GMI PT 2015) and obtained results from the carried out proficiency tests related to the wet and dry-lab component. Two separate reports will be perform. (http://www.globalmicrobialidentifier.org/Workgroups/GMI-Proficiency-Test-Reports).

The following presentation related to the PHE integrated pipeline and INNUENDO project, which focus on WGS integration in routine surveillance and epidemiological investigations. Aspects of ongoing collaboration between ENGAGE and INNUENDO concerning benchmarking of Salmonella serotyping tools were presented. INNUENDO offered also participation on the GMI-PT and THL project.

The problems with ENGAGE working space which occurred have been repaired, new working space has been built. Data which are submitted need to be re-submitted. According to the plan for ENGAGE E-learning, an

online course on Whole Genome Sequencing tools and applications in surveillance of pathogens is now under preparation by DTU and will be launched on Coursera after summer 2017.

Next twinning round is planned in September 2017. Any suggestions from consortium partners are welcomed. For developing a common Core Genome MLST (cgMLST) scheme a large number of reference sequences are needed. The relevant EURLs are the potential resource of them.

Summary of the break-out group session

Subject: Technical Challenges of Implementing Bioinformatics for Public Health. Bioinformatics Infrastructure.

For Public Health purposes as surveillance, quick reports, outbreak investigations, large-scale epidemiology) a centralized model is needed with established rules for data access to accredited users by appropriate institution.

The main role of international projects (e.g COMPARE, GMI, EBI, NCBI) is to facilitate and disseminate data to scientific and decision making agencies. Participant intended to store their raw data, intermediate results and assembled genomes. For species identification, MLST, pMLST, AMR genes phylogeny CGE pipeline are used usually. For MLST and assembly other database, programs are employed, including publicly available like Enterobase. The advantages of CGE platform are: Robustness, accuracy, user-friendly, based on underlying and declared software and databases, server-client approach.

Training course on NGS data analysis, Warsaw, 12-14.10.2016

Trainees:

Maria Borowiak	BfR	Germany
Jens Andre Hammeri	BfR	Germany
Kerstin Stingl	BfR	Germany
Magdalena Zając	NVRI	Poland
Katarzyna Półtorak	NVRI	Poland
Arkadiusz Bomba	NVRI	Poland
Aleksandra Giza	NVRI	Poland
Edyta Świętoń	NVRI	Poland
Yue Tang	APHA	United Kingdom
Sara Petrin	IZSVe	Italy
Lelde Grantina-Ievina	BIOR	Latvia
Bojan Papić	VF-UL	Slovenia
Manuel Francisco Hernandez	MAGRAMA	Spain
Veldman Kees WUR	Netherlands	
Daniela Ceccarelli	WUR	Netherlands
Bavo Verhaegen	WIV-ISP	Belgium
Raf Winand	WIV-ISP	Belgium
Mistou Michel Yves	ANSES	France
Tomasz Wołkowicz	NIPH-NIH	Poland
Katarzyna Zacharczuk	NIPH-NIH	Poland
Natalia Wolaniuk	NIPH-NIH	Poland
Magdalena Rzeczkowska	NIPH-NIH	Poland
Aleksander Masny	NIPH-NIH	Poland
Łukasz Mąka	NIPH-NIH	Poland
Maciej Polak	NIPH-NIH	Poland

The first ENGAGE training course on “Next Generation Sequencing data analysis” was conducted on 12-14 October 2016. It was hosted and organised by NIPH-NIH. Training course was leaded by three bioinformatics expert, with a high degree of knowledge of this subject: Pimlapas Leekitcharoenphon (DTU, Denmark), Eleonora Mastrorilli (IZSVe, Italy) and Ali al-Shahib (PHE, UK).

All ENGAGE partners were invited to participate in the training course. Additionally external, non-ENGAGE representatives from European and National Reference Laboratories were also invited. Finally 25 peoples took part in the Training Course (complete list of participants is reported in Table 1). They represented a wide variety of institution from various European countries (including: France, UK, Poland, Germany, Italy, Latvia, Belgium, Spain, Slovenia, Netherlands). Majority of group was the person at the beginner’s level in this field. It was consistent with objectives of the project which assumed preparation in the first year of project duration, training based on fundamental introduction to NGS data analysis. It was designed for users who want to acquire the competence to analyse the next-generation sequencing (NGS) and other large-scale data sets independently and/or in the future a proficient manner.

Program consists of two parts: theoretical and practical part, respectively. The lecture part included an overview of NGS technologies, implementation of NGS for microbiology studies (clinical microbiology, genomic epidemiology) and introduction to training part about two open-source, web-based platform for sequence data analysis: Galaxy and CGE. The hand on training was covering issues of computer analysis on WGS data. All participants were using their own computers and were able to analyse previously prepared example sequences as well as their own WGS data.

During the Training Course participants have learned how to trim, analyse and assemble the raw reads with different assemblers (Velvet and SPAdes) and how to interpret the quality of the reads and the assembly. It was also trained how to find out the genus (using both k-mers and of 16S rDNA analysis), MLST pattern, plasmids occurrence, resistance genes and chromosomal resistance mutations, how to predict *Salmonella* and *E. coli* serotype and *E. coli* pathotype. Additionally participants have learned how to analyse genetic similarity using SNP analysis. Using Galaxy, participants have learned also how to create workflows using different analytical steps and different tools.

After the Training Course, every participant received a signed certificate of attendance.

Course material from training was freely available, including presentations and others materials provided by trainers.

Evaluation summary

After the training course all participants had an opportunity to share their opinions about the course. Overall, the participants reported that the training course was highly scientific and well explained. They highlighted knowledge and professionalism of the trainers. According to their opinion, the training gave them a lot of basic knowledge about NGS data analysis and information about bioinformatics tools for data analysis. Participants found the presentations clear and both part of training (theoretical and practical) good prepared. In their opinion the structure of the course was properly composed and content presented in a logical sequence: reliable introduction to Galaxy and CGE interface and then using them in practice. Opportunity for open discussion between teachers and participants was found valuable.

--- --- ---

Workshop and Training course 2017

Establishing Next Generation sequencing Ability for Genomic analysis in Europe (ENGAGE)

ENGAGE Workshop, 23-24 October 2017

Venue: LAZIOCREA (Lazio Region), Rome, Italy

Attendees:
Beatriz Guerra Román (EFSA)
Anaïs Painset (PHE, United Kingdom)
Liljana Petrovska-Holmes (APHA, United Kingdom)
Maria Borowiak, (BfR, Germany)
Burkhard Malorny (BfR, Germany)
Magdalena Zając (NVRI, Poland)
Katarzyna Półtorak (NVRI, Poland)
Rene S. Hendriksen (DTU, Denmark)
Susanne Karlsmose Pedersen (DTU, Denmark)
Pimlapas Leekitcharoenphon (DTU, Denmark)
Tomasz Wołkowicz (NIPH-NIH, Poland)
Katarzyna Zacharczuk (NIPH-NIH, Poland)
Antonio Battisti (IZSLT, Italy)
Virginia Carfora (IZSLT, Italy)
Angela Ianzano (IZSLT, Italy)
Fabiola Feltrin (IZSLT, Italy)
Alessia Franco (IZSLT, Italy)
Alba Patricia (IZSLT, Italy)
Lisa Barco (IZSVe, Italy)
Eleonora Mastrorilli (IZSVe, Italy)
Carmen Losasso (IZSVe, Italy)
Angela van Hoek (RIVM, Netherlands)
Valeria Michelacci (ISS – EURL-VTEC, Italy)
Mirko Rossi (VETMED, Finland)

Reference:
As indicated below the presentations are made available on 'sharepoint' which refers to https://share.dtu.dk/sites/Engage_149200/SitePages/Home.aspx [login information sent to partners on email 9th June 2016. If you experience problems accessing the site, please contact suska@food.dtu.dk]

Meeting notes:
- Welcome and introduction to the ENGAGE workshop 2017 (Antonio Battisti)

- Update, ENGAGE (Rene Hendriksen, DTU Food) [presentation available on sharepoint]
    - Objectives of ENGAGE presented.
    - Status of ongoing tasks presented.

- EFSA, WGS for Food Safety (Beatriz Guerra, European Food Safety Authority) [presentation available on sharepoint]
    - Presentation of the current and planned EFSA activities within the field of WGS.
    - WGS has already been in use for two multi-country outbreaks (New *Salmonella* serotype, S. 11:z41: e,n,z15 linked to sesame seeds and *S.* Enteritidis linked to eggs).
    - WGS is applied in the EUSR-AMR reference testing (to support the quality of the data reported to EFSA by the MSs for the EUSR-AMR report, to detect emerging resistance

mechanisms/resistant clones, to see the correlation between phenotypes and genotypes, and to demonstrate the value of WGS supporting for AMR surveillance).
    - Results from the EC Questionnaire supported by EFSA on the availability of WGS methods in the MS have now been analysed.
    - ECDC and EFSA have received a mandate from the EC for "Technical support to collect and analyse whole genome sequencing (WGS) data in the joint ECDC-EFSA molecular typing database" at least for *Salmonella*, *Listeria* and *E. coli*. A joint WG ECDC-EFSA has been set up. Deadline to deliver to EC is April 2019.
    - At EFSA, in-house capacity building is ongoing, currently mainly focusing at *Salmonella*, *E. coli* and *Listeria*.
    - Many advantages based on the outcomes of the ENGAGE project

  Discussions:
    - Is the aim one tool or 'any/many'?
        - EFSA: We will try several tools but for now there is no particular tool we suggest.

- Status from INNUENDO; Integrating genomics into food-borne surveillance (Mirko Rossi, Finland) [presentation available on sharepoint]
    - Collaboration between ENGAGE and INNUENDO is ongoing.
    - INNUENDO platform has been set up targeting *E. coli*, *Salmonella enterica*, *Campylobacter* spp., *Yersinia enterocolitica*.
        - Main modules: sequence check (INNUca), Allele call link (chewBBACA), automatic curation and performance. These modules will be available in Galaxy soon.

  Discussions:
    - The assembler used for the sequence analysis makes a difference for the analysis outcome.

- 'Tour de partner' (Chair: Rene Hendriksen, DTU Food) – Each partner institution (Availability of strains related to the target list; Number of strains sequenced; Deliverables completed) [presentations available on sharepoint]
    - IZSLT, Italy – Till now, 373 strains sequenced. Most were 93 *Salmonella* Infantis and also *E. coli*. Need to perform a little sequencing still. Chaired the *Salmonella* Infantis project. Organized the training course and workshop.
    - NIPH-NIH, Poland – Till now, 174 strains completely sequenced and uploaded to the working space. Under sequencing: 31 VTEC strains and 116 *Salmonella*. Participated in sequencing for ENGAGE projects. Most important thing that ENGAGE has contributed with for our institute is that now we are able to perform analysis of the sequences.
      Question: Is the outbreak strain a TEM-1B or a TEM-169?
    - BfR, Germany – Till now, sequenced 380 strains of *Salmonella* (mostly Paratyphi B Java and Typhimurium) and *E. coli*. Three publications (on VIM-1-producing *Salmonella* Infantis, first *mcr-1* Paratyphi B-isolate, identification of the mcr-5)
      Comment: Uploading the results to ENA via the working space currently means that the owner is DTU. DTU will follow-up to check if this can be changed and report back to the ENGAGE partners.
    - IZSVe, Italy – concluded the sequencing of 461 strains. All were *Salmonella* (mostly Derby, Napoli, monophasic Typhimurium). Main achievement is that we have this technique available now. Organizing of twinning has been concluded with one more twinning session in 2017.
    - PIWet, Poland – Since the last workshop, we bought an Illumina MiSeq and we do not anymore need to outsource. We were trained before and as soon as we received the equipment, we could start sequencing, and we sequenced 180 strains from April till now. Till now, for the ENGAGE project we sequenced 305 strains. Additional 68 strains expected by the end of November.
    - PHE, UK – Sequence all *Salmonella* we receive. Per year, we sequenced 11,000 *Salmonella* and 3,600 *E. coli*. ENGAGE has been important for us to compare and improve our methodology. Benchmarking exercises have been important.
    - APHA, UK – for the ENGAGE we sequenced 422 strains, all are uploaded to the ENGAGE working space. The benchmarking has been very important for us.

- DTU Food, Denmark – Till now, sequenced 223 *Salmonella* and 313 *E.coli*. Deliverables completed have been: WGS, benchmarking, PT, working space, administration, E-learning.

- Overall data availability with ENGAGE (Chair: Liljana Petrovska, APHA) [Presentation available on sharepoint]
  - In total 3220 sequenced (2763 sequences uploaded) (2623 *Salmonella* and 597 *E. coli*)
  - By the end of the ENGAGE project, all data must be available for download with an accession number for each sequence.

- ENGAGE projects; preliminary data of [Presentations available on sharepoint]
  - *S*. Infantis (Antonio Battisti) (IZSLT, PIWet, APHA, NIPH-NIH, BfR, DTU, PHE) In Italy the *S.* Infantis is predominant in broiler flocks and there is a statistical significant rise in the proportion detected. The project focuses at phylogeny among isolates in the EU Countries and abroad, for molecular epidemiology and source attribution purposes, the emerging ESBL-producing clone, co-resistance, and in looking into if the same clone is also detected in other countries. Till now, 299 isolates sequenced. Pending tasks are to complete the phylogeny and to analyse all the results with ResFinder 3.0.
  - *S*. Derby (Lisa Barco and Eleonora Mastrorilli) (IZSVe, BfR, DTU, APHA, PIWet, NIPH-NIH, IZSLT) One of the ten *Salmonella* serovars that is reported to give most infections in humans. *S.* Derby from turkey was only reported from the UK. Till now, 377 strains sequenced.
  - *S*. Napoli (Carmen Lossasso) (IZSVe, BfR) – This serovar causes severe infection but is not common in Europe. The study aims to clarify the epidemiology. Till now, sequenced 157 strains. Pending: Analysis of data.
  - WGS of rare and unrecognized *Salmonella* serovars (Magdalena Zajac) (NVRI, DTU, APHA) Isolates that caused difficulties in conventional serotyping were included. From a genetic point of view, Newport/Bardo and Senftenberg/Dessau is the same serotype. Further data analysis is in progress.

    Suggestion: How about including sequences available in ENA in the analysis?
  - *mcr-1* positive *E. coli* in Poland (Magdalena Zajac) (NVRI, DTU, APHA) Out of 5878 *E. coli* that were included in the analysis, 128 isolates were suspected to harbor *mcr-1*. 80 were found to harbor *mcr-1* and none were found to harbor *mcr-2*.

- Benchmarking session: update on benchmarking
  Status on planned publications – presentation of data and analysis performed and timeline for the publication [Presentations available on sharepoint]
  - Benchmarking on *Salmonella* serotyping tools (Liljana Petrovska-Holmes, APHA). 798 WGS data files included in the analysis. In some cases it was detected that the pipelines predicted the same serotypes but different from the phenotypes, sometimes due to incorrect phenotype, sometimes due to an updated name in the Kauffmann-White, sometimes because the serotype is not present in 'Enterobase'. Prediction of Enteritidis or Gallinarum needs improving (sdf-/sdf+ will be included to confidently say that it is an Enteritidis or not). Combination of 3 pipelines may give more accurate results.

    Discussion in plenum:
    - APHA currently calculates 72 hours from isolated bacteria to WGS serotype.

  - Benchmarking on antimicrobial profiling using WGS (Anaïs Painset, PHE) Benchmarking was performed on a well-characterized dataset. One of the challenges was that it was not possible to verify the gold standard.

    Discussion in plenum:
    - The trimming will also have an impact. The data analysed for this benchmarking was trimmed before the participants downloaded it
    - Individual feedback to participants will be given subsequent to this meeting.

- Potentially a comparison without gold-standard would be an option. Anaïs will explore this.

- Benchmarking on ISO for serotyping (Angela van Hoek, RIVM) 27 isolates were included for species identification and serotyping of the *Salmonella*. 13 sets of results were submitted. For the expected result, *S.* Eko, the tools indicated Hadar/Istanbul. Related to the ISO 16140-6, for the outcome of the inclusivity and exclusivity study indicated that related to the inclusivity, the WGS did not pass the acceptance level. Scientific publication is on the way.

  Discussion in plenum:
  - All steps in the analysis must be included for the analysis, e.g. the assembler is also relevant – not only the tools employed for the serotyping.
  - Additional relevant information could be depth, coverage and other characteristics about the analysed sequences.
  - Discussions ongoing related to how to calculate the inclusivity.

- Benchmarking on AMR/point-mutation – Ali
  - Comment from Anaïs that the analysis and writing of the manuscript is ongoing.

- About the GMI PT 2017 incl. obtained results (Rene Hendriksen, DTU Food [presentation available on sharepoint]
  - Introduction to the GMI PT, wet-lab component
  - Data related to the wet-lab component will be compiled in a report.
    Question: Is more than 300 contigs really considered a poor performance?

- Updates for the bioinformatics tools that have already been identified (Anaïs Painset, PHE) [presentation available on sharepoint]
  - Relevant tools to update, e.g. 'glue-together-scripts' and others (Anaïs will work with the ENGAGE group to update the list)
  - Anaïs will update the current list of tool including also comments/pros/cons on the mentioned tools e.g. that SPAdes does not work on HiSeq-data (and including a disclaimer).

- Discovery of *mcr-5* and distribution of *mcr* genes among colistin resistant German *Salmonella* isolates from animals, food and environment (Maria Borowiak, BfR) [presentation available on sharepoint]
  - Originally research was done on S. Paratyphi B to identify *mcr-1*. Of these strains, 32 were negative for *mcr-1*, *-2, -3, and 4*. Subsequently 14 of these were found to harbor *mcr-5* based on an assembly of the un-mapped genes and annotating using RastTK and confirmation by PCR and PFGE, southern blot and hybridisation.
  - The *mcr-5* was found in plasmids in 14 of the strains and in one isolate variant, the transposon including the *mcr-5* was found to have been incorporated in the chromosome. These were found to express different phenotypes where the isolates harbouring the *mcr-5* in plasmids exhibiting an MIC of 8 and the isolate(s) harbouring the *mcr-5* in the chromosome exhibiting an MIC of 4.
  - A global spread of this gene is likely.

- E-learning available on Coursera – what to add to the E- learning course and how do we proceed (Chair: Pimlapas Leekitcharoenphon (Shinny), DTU Food) [presentation available on sharepoint] and breakout group discussions and follow-up
  - Suggestions on topics which will be include in the E-learning:
    - cgMLSTFinder 1.0 tool (Responsible: Shinny)
    - SalmonellaTypeFinder 1.4 tool (Responsible: Shinny)
    - Sequencing quality (Responsible: Shinny)
  - Suggestions on topics which will be not be include in the E-learning as part of the ENGAGE project:
    - VirtualBox – how to install bioinformatics tools (to use virtualbox is straightforward. We will provide a guideline to install VirtualBox on our website, responsible: Shinny)

- Wet lab introduction (resource-wise it is currently not possible. Wet lab introduction as E-learning will be considered for another project such as EURL)
- Bioinformatics tools (pros/cons) (Anaïs will provide details of bioinformatics tools including pros and cons so there is no need a video for this)

- Discussion: Plan for uploading the sequences to ENA (Chair: Rene Hendriksen, DTU Food)
  - Important that accession numbers are available for all the 3,000 genomes sequences as part of the ENGAGE project. All must be available via ENA.
  - Currently 225 sequences are shared on the working space. Those that uploaded the sequence to the working space should share them with the ENGAGE group.
  - Shinny will follow up on how to upload from the working space to ENA and report back to the ENGAGE partners on how to proceed with the upload to ENA. There's also a command line tool how to batch-upload sequences to ENA. It would be preferable that in ENA, all sequences should be under the ENGAGE umbrella.
  - **Deadline for each partner to upload to ENA: December 1st 2017**.
  - Coordinator/contact: Anaïs Painset, PHE.

- Discussion: ENGAGE final report (Chair: Rene Hendriksen, DTU Food). The final report will be published and publically available. All partners should co-author the ENGAGE final report. We can base the final report on the Interim report and expand it with the relevant additional accomplishments.
  - WP 1 (Project Cooperation Framework) – Responsible: DTU (Rene Hendriksen)
  - WP 2 (Data Collection), including the logic behind the selection of strains – Responsible: APHA (Liljana Petrovska-Holmes)
  - WP 3 (Whole Genome Sequencing) – Responsible: PHE (Anaïs Painset)
    - For WP 3, all consortium partners should contribute with a description (about half a page) from their institute on how WGS has been applied during the ENGAGE project, and including description of the 'proof of concept' projects. Further description of the projects should be placed in appendices. Responsible: All consortium partners (please send your contribution directly to Anaïs Painset **at the latest on December 1st 2017** for her to include it in the WP 3 description.
  - WP 4 (Genome Database) – Responsible: DTU (Rene Hendriksen)
  - WP 5 (Analysis/Benchmarking) – Responsible: PHE (Anaïs Painset will confirm who will be the contact person)
  - WP 6 (Proficiency Testing) – Responsible: DTU (Rene Hendriksen)
  - WP 7 (Output/Outcome) – Responsible: APHA (Liljana Petrovska-Holmes)
  - WP 8 (Training/E-learning) – Responsible: NIPH-NIH and IZSLT (Tomasz Wołkowicz)
  - WP 9 (Workshops) – Responsible:
  - WP 10 (Dissemination) – Responsible: DTU (Rene Hendriksen)
  - WP 11 (Project Management) – Responsible: DTU (Rene Hendriksen)

  Also, we need to include the 'stories' from each institute.

  EFSA: What has been produced from ENGAGE has been very useful for others also. The outputs of this and other ongoing projects can be in fact be very useful to support some managerial measures (i.e. changes in the Legislation to include the use of WGS-based methods as alternatives/complementarily to current methods in use) . Thus the importance of presenting the facts, well-documented and providing enough background on the exercises performed in the final and benchmarking reports. An example could be what was the background for selecting the isolates that were included in the benchmarking. There are differences in the application of the tools, and this could be interesting to look into. Straight forward conclusions are important, also for each work package.
  The pilot projects and benchmarking projects are not seen in the Gantt scheme, but are in fact main subprojects of ENGAGE (diversity of activities performed, capacity built, and very valuable results) and should not be mentioned in an appendix only but should be presented in the body text of the final report (i.e., main findings/conclusions should be in the body text of the report and the full benchmarking/project reports should then be included as appendices. Also, many things were

discussed at this meeting but have not yet been included in the benchmarking reports.

It was agreed that to adjust the benchmarking reports, DTU Food will send round the benchmarking reports and ask all consortium partners (one by one) to add changes using tracked changes. Publications that are on the way should be included in the report in summary. As regards the final full publication, please share the draft publications with EFSA.

All contributions for the final report should be sent to rshe@food.dtu.dk and suska@food.dtu.dk at the latest on: **Friday December 8th 2017**

- Future opportunities and AOB (Chair: Rene Hendriksen, DTU Food)
  - Discussions were included during the meeting and no further issues were brought up.

- Summary of the ENGAGE workshop (Chair: Rene Hendriksen, DTU Food)

--- --- ---

Establishing Next Generation sequencing Ability for Genomic analysis in Europe (ENGAGE)

Venue: LAZIOCREA, Rome, Italy

Summary report, Workshop and Training course 2017

ENGAGE Workshop, Rome, 23-24.10.2017

The second ENGAGE Workshop was held from 23 to 24 October 2017 at LAZIOCREA (Lazio Region), Rome, Italy and organized by the Istituto Zooprofilattico Sperimentale del Lazio e Toscana "M. Aleandri" (IZSLT; Italy) D. O. Diagnostica Generale, NRL-AR, Italy.

The objective of the workshop was to update information about project progress, deliverables and identifying any emerging issue with project activities one year after the ENGAGE Workshop 2016, held in Warsaw (October 2016). A total 24 participants attended the workshop, including the participants of the ENGAGE projects, representative of EFSA and three non-Engage participants, representatives of: EURL-VTEC (Italy), RIVM-NL/EURL Salmonella (The Netherlands), and INNUENDO consortium (Finland).

In the 1° part of the first day of the workshop, Rene Hendriksen (DTU) presented an overview of the specific objectives of the project, the tasks already achieved and the status of the activities still ongoing. Briefly, from the timeline of the kick-off phase of the project emerged that the ENGAGE network has been extended, including other partnerships such as the EURL-Salmonella (RIVM, The Netherlands). The implementation of a new functional cloud-based protected working space and its regular use for data sharing/analysis among all the members of the consortium was underlined. As for the timeline of the analytic phase, an update of the analysis and the benchmarking exercises performed during 2017 on Salmonella serotyping tools/ISO Protocol, phylogeny and AMR (Antimicrobial Resistance) tools were introduced. The timeline of the proficiency testing phase showed that all the participants completed the Global Microbial Identifier Proficiency Test 2017 for the wet lab part. As part of the completed deliverables, the workshops, the training-courses, the E-learning modules and the twinning projects performed during 2017, together with the list of the submitted articles and dissemination of data through international conferences were presented.

-Representative of EFSA and coordinator at EFSA of the ENGAGE activities Beatriz Guerra presented the current approach of EFSA to WGS technologies in the field of zoonoses surveillance and foodborne pathogens detection

during outbreak investigations. The use of WGS analysis supporting phenotypical AMR data according to Decision 2013/652/EC on AMR monitoring in zoonotic and commensal bacteria was also discussed. The capacity of Member States (MSs) to detect foodborne pathogens by WGS analysis was analyzed based on questionnaires distributed by ECDC and EFSA on the behalf of European Commission (EU) and was discussed together with the current status of the "WGS umbrella project" for group together WGS EFSA current activities and projects. The use of command line tools/pipelines, web tools (e.g. CGE), commercial software (e.g. Bionumerics) and research project platforms (e.g. INNUENDO platform) was discussed.

- Representative of the INNUENDO consortium (Mirko Rossi), focused on update information on the use of analytical platforms and standard procedures in the frame of the INNUENDO project (www.innuendoweb.org), aiming at the integration of WGS to surveillance and the outbreak investigation of food-borne pathogens in the context of small countries with limited resources (more information on INNUENDO workshop held in July 2017 at https://ehutb.ehu.es/series/59b66d7bf82b2b150d8b468e). M. Rossi also presented the results of the INNUENDO consortium, describing the tools developed in the context of the project for the quality control and assembly of the raw reads and wgMLST. Among the planned activities, implementation of cgMLST /wgMLST schemes into the CGE tool was included.

After that, all the partner showed the work done so far and the accomplished deliverables (number of strains sequenced so far, number of sequences upload to the ENGAGE site, number of sequenced upload to ENA, and related projects to the sequences strains). At the ending, Liljana Petrovska (APHA) gave us a summary of the sequenced and upload strains from all the partners.

-Preliminary data of the following ENGAGE projects were presented: *Salmonella Infantis* (Antonio Battisti, IZSLT), *S. Derby* (Lisa Barco/Eleonora Mastrorilli, IZSVe), *S. Napoli* (Carmen Losasso, IZSVe), WGS of rare and unrecognized Salmonella serovars (Magdalena Zajac, NVRI).

In the Benchmarking Session, update information on benchmarking activities performed since the last ENGAGE Workshop (October 2016) was presented. Regarding the benchmarking exercises performed for Salmonella serotyping, a total of 27 genomes (23 Salmonella and 4 Non-Salmonella Enterobacteriaceae genomes) from APHA, DTU and RIVM were included in the data set to evaluate available bioinformatics tools for predicting Salmonella serotypes according to ISO 16140-6 -Part 6 "protocol for the validation of alternative methods for microbiological confirmation and typing procedures". All the ENGAGE partners participated in the benchmarking exercise. The obtained results demonstrated that serotyping using WGS is feasible although depends on the chosen tool. Future plans included possible publication of this benchmarking effort as a note at the beginning of 2018.

In order to evaluate the Salmonella serotyping tools available, a larger benchmarking exercise was performed by APHA, DTU, PHE and INNUENDO consortium (Univ. Helsinki). The raw reads of 798 Salmonellas, including 137 different serovars, were analyzed using MOST (PHE), SeqSero (Georgia University), Salmonella TypeFinder (DTU) and SISTR (PHA Canada).

A phylogeny benchmarking exercise was also performed. 30 *Salmonella enteriditis* outbreak-related genome sequences were send to the all ENGAGE participants in order to build the phylogeny using on-line tools or bioinformatics command-line software. Each obtained phylogeny was compared to the "true phylogeny" in terms of variants calling, SNPs alignments and tree topologies (PHE). From the obtained results two crucial points appeared to influence the final phylogeny: variants calling and the method used to build the tree.

Updates on AMR benchmarking were also presented: Point mutation analysis and ARIBA (Antimicrobial Resistance Identification by Assembly) testing are under progress. A draft on AMR benchmarking has been prepared and will be sent in the beginning of 2018.

The second day started with the presentation and discussion of the results obtained by the ENGAGE participants in the web lab component of the Global Microbial Identifier (GMI) Proficiency Test 2017 by Rene Hendriksen (DTU). A total 6 isolates (2 *Salmonella*, 2 *S. aureus* and 2 *E. coli*) along with their extracted DNA were send to each participant laboratory. The results were satisfactory for most of the participant labs.- Updates for the bioinformatics tools availables for the analysis of raw reads were discussed, particularly regarding SNPs detection (e.g. Snippy, PHEnix...). At this regard, the different use of the tools according the type of raw data were discussed and the inclusion of the pros and cons of the tools most used in the list of bioinformatic tools available was proposed.

- Maria Borowiak (BfR) presented the already published paper: "Discovery of mcr-5 and distribution of mcr genes among colistin resistant German Salmonella isolates from animals, food and environment", as part of the Engage projects.

Summary of the break-out

Subject: E-learning course

Break-out groups focused at E-learning course. An online course on Whole Genome Sequencing and applications of CGE tools in surveillance of pathogens is already available on Coursera. For the completion of this WP of the project, al least one more e-learning course should be done and published. The proposed topics for the new videos were:

- installation of VirtualBox and Biolinux

- cgMLSTFinder tool

- SalmonellaTypeFinder tool

- Quality control

The most discussed was the topic "Quality control" because it could be a useful tool for ensure the quality of the reads prior to the analysis. Volunteers for preparing the new video were requested.

In the last part of the second day, the plan for uploading sequences to European Nucleotide Archive (ENA) was discussed, since one of the aims of the project is made public available the sequences obtained at the end of the project (January 2018). The ENGAGE working space has the option of uploading the raw reads directly to ENA, but in this case reads result to be uploaded by the DTU. DTU would ask if this functionality of the working space can be changed. At the end, was concluded that each partner should upload all the sequences by the beginning of December (2017) and decided if upload it directly to ENA or use the ENGAGE workspace by the end of December.

A final discussion on drawing up the ENGAGE final report due in January 2018, was tackled. The main points covered decision on the format to be used, expectations and responsibilities from all the members of the consortium. Tasks were assigned to the different participants. The inclusion of a new subsection with the description of the meaning of WGS/ENGAGE for each institution and a brief description of the project and selected strains was proposed and accepted. EFSA reminded to the Consortium the importance of a good final report and good conclusions.

Training course on NGS analysis based on command line tools, Rome, 25-27.10.2017

List of Participants:

1) Carmen Garcia Pelayo (APHA.GSI.GOV-UK)
2) Yue Tang (APHA.GSI.GOV-UK)
3) Maria Borowiak (BFR-DE)
4) Josephine Gruetzke (BFR-DE)
5) Pimlapas Leekitcharoenphon (DTU-DK)
6) Patricia Alba (IZSLT-IT)
7) Antonio Battisti (IZSLT-IT)
8) Francesco Bottoni (IZSLT-IT)
9) Virginia Carfora (IZSLT-IT)
10) Fabiola Feltrin (IZSLT-IT)
11) Alessia Franco (IZSLT-IT)
12) Angela Ianzano (IZSLT-IT)
13) Eleonora Mastrorilli (IZSVE-IT)
14) Katarzyna Półtorak (NVRI-PL)
15) Tomasz Wołkowicz (PZH-PL)
16) Katarzyna Zacharczuk (PZH-PL)
17) Magdalena Zając (NVRI-PL)
18) Anais Painset (PHE.GOV-UK)
19) Angela van Hoek (RIVM-NL)

The Training Course on "NGS analysis based on command line tools" was held from 25 to 27 October 2017 at LAZIOCREA (Lazio Region), Rome, Italy and organized by the Istituto Zooprofilattico Sperimentale del Lazio e della Toscana "M. Aleandri" (IZSLT; Italy).

The Training course lasted 3 full days, from 25 to 27 October 2017. The event could have been attended by internal (ENGAGE Consortium) as well as external (non-ENGAGE Consortium) participants. A total 19 persons with different background and experience with NGS data analysis took part in this course. They represented a variety of institutions from different European countries (including: UK, Poland, Germany, Italy and the Netherlands).

The majority of the group had a quite long experience in using web-based tools for NGS data analysis but not all the participants were familiar with the Unix/Linux command line. The Training Course was designed for users who want to acquire the competence to analyze more efficiently NGS data by using command line tools. This course was structured as a continuation of the first ENGAGE training course held in 2016 in Poland.

The aim of this course was to give an overview of the LINUX systems and the use of command line in the NGS data analysis, and provide to the participants a basic knowledge of tools used for the analysis independently of the web tools.

The program consists of three parts, leaded respectively by three bioinformatics expert with a high degree of knowledge of this subject: Pimplas Leekitcharoenphon (DTU, Denmark), Eleonora Mastrorilli (IZSVe, Italy) and Anais Painset (PHE, UK). Every part consists of theoretical and practical sessions. For the practical part, participants were asked to have their own computer. Before attending the course, specific guidelines were provided by email by trainers to give the opportunity for those participants who were not used to Unix/Linux interface or not having a Unix/Linux machine, to set up on their own pc a virtual machine simulating a Linux environment, and also to install some tools from command line (Trimmomatic, SPAdes). In this way, the participants will be able to perform the data analysis by their own.

The theoretical part included an updated overview of the NGS, an introduction to the basic LINUX command line and the description of the tools used during the training as FastQC, Trimmomatic, SPAdes, BLAST, and the CGE tools installed in a Docker systems.

First Part - Eleonora Mastrorilli (IZSVe, Italy) - 25.10.17

Started with an updated overview of NGS Technologies used so far and then introduced the installation procedures of Virtual Box. An introduction on the use of the Unix bash along with the basic commands was presented, followed by practical exercises on Linux basic commands.

Second Part - Anais Painset (PHE, UK) - 25-26.10.17

Started with a lecture on using command line tool for clipping, trimming and assessing quality reads. Practical exercises on quality assessment were performed using FastQC for the evaluation of the raw reads (fastq files) quality and Trimmomatic for the trimming of those Fastq files. This part continued with an introduction on using SPAdes for “de novo” assembly, followed by practical exercises on assembling genomes by using SPAdes assembly toolkit. For evaluating the quality of the assembly, the Quast tool was introduced and used.

At the end, the BLAST command line tool was presented for the comparison of the assembled genomes with specific gene databases created by the user.

Third Part- Pimplas Leekitcharoenphon (DTU, Denmark). 26-27.10.17

The third part started with an introduction on using and installing CGE Docker tools (CGE pipelines running in a Docker container). Theoretical introduction and practical exercises on using the following CGE tools (command lines) have been addressed:

- Assembly
- Contig Analyzer
- KmerFinder for species identification
- MLST for ST typing
- ResFinder for identification of resistance genes
- PlasmidFinder to identify plasmid replicons
- pMLST for plasmid ST typing
- VirulenceFinder for determining virulence genes
- PlasmidFinder to identify plasmid replicons
- pMLST for plasmid ST typing
- VirulenceFinder for determining virulence genes
- Bacterial Analysis Pipeline (paired end reads)
- Bacterial Analysis Pipeline (contigs)

Course material from training was freely available, including presentations and others materials provided by trainers, one to one feedback moments were also forecasted during the course. All the trainers were constantly present in the three full days in order to help participants during practical sessions.

Training course. Evaluation summary.

The attendees rated the training course very high, and gave constructive and positive feedback about strong points and some ideas to improve it.

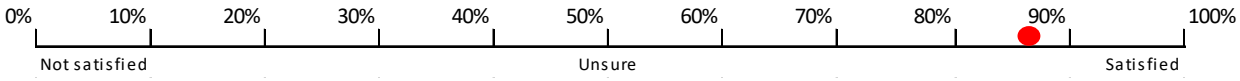
Participant Evaluation Form:

Please rate the contents of the training course	Not useful	Somewhat useful	Unsure	Useful	Very useful
a) Relevance of the training course in relation to my job				3	12
b) Scientific knowledge gained during the training course		1		1	13
c) VirtualBox			1	3	11
d) Introduction to UNIX command line				2	13
e) Command line tools for WC and trimming raw reads				1	14
f) De novo assembly using SPADES	1		1		13
g) BLAST				1	14
h) CGE docker tools				2	13
i) Discussion, questions and answers		1		1	13

Please rate the contents of the training course	Strongly disagree	Disagree	Unsure	Agree	Strongly agree
a) The organisation prior to the training course was efficient				5	10
b) The training course was overall well organised				5	10
c) The rime assigned to presentations was appropriate				5	10
d) The time assigned to pratical laboratory session was appropriate*		1	1	5	7
e) The time for discussion and partecipant interactions was appropriate				5	10
f) The training course facilities and location were satisfactory			2	7	6
g) The content of the material handed out at the workshop will be useful				6	9
*only for fourteen partecipants					

Please indicate with a mark on this scale your level of overall satisfaction with the training course:

Average (only for ten partecipants): 87%



DTU Food
National Food Institute