11th Global Microbial Identifier Meeting Presentation Abstracts

Day 2 Thursday 17th May 2018

Active Systems and Overcoming NGS Barriers in the Developing World

M. C. Enrique J. Delgado Suárez

Faculty of Veterinary Medicine, National Autonomous University of Mexico

WGS has proven to be a robust tool to improve the control of infectious diseases. Despite its cost has dropped significantly, the use of NGS in developing countries is still discrete. In these nations, active epidemiological surveillance is seldom practice, even with traditional methods of microbial identification. Generally, only the most severe cases that require patients' hospitalization involve a more thorough investigation of the involved pathogen, usually in regard to its antimicrobial resistance. However, these isolates often remained in the hospital's collection and the associated information is not further used for epidemiological purposes. There are multiple factors associated with this situation. In Mexico, for instance, it would require a profound amendment of current laws and regulations. Although some governmental institutions have a growing collection of pathogen isolates, the public sharing of this information is restricted by law. Moreover, research capacity in this area is limited, as well as government policies stimulating the application of funds for these purposes. Still, there has been some important investments in government laboratories, mostly due to the pressure of the intense commercial exchange of foods. Therefore, integrating governments into the GMI initiative may be the fastest way to overcome NGS barriers in the developing world.

Genomic Data Sharing under Nagoya Protocol – Future International Initiatives

George Haringhuizen

RIVM, NL

GMI is committed to a world where high quality microbiological genomic information from human, animal and plant domains is freely shared among all nations [……..]. The GMI mission is to build a global platform linked to an open and interactive worldwide network of databases for standardized identification, characterization and comparison of microorganisms through the storing of whole genome sequences of microorganisms, the connected metadata, and the provision of analytical facilities and shared standards.

In November of this year, at the 14th meeting of the Conference of Parties (CoP) of the CBD/Nagoya Protocol (NP), the status of NGS-data is on the agenda: do NGS-data fall by nature under the scope of NP or not? The CoP issued a scoping paper for discussion on which many research institutes, funding agencies, national institutes, and international organizations gave their opinion and expressed concern, among which WHO and FAO. If NGS-data must be considered as equivalent to tangible genomic resources, the status of open access NGS-databases will be comparable to existing Culture Collections and BioBanks that need already be compliant to the Protocol. So, how do these Biobanks cope with the Nagoya conditions and who bears the burden? What can we learn? We looked for systematic approaches and identified 4 models (American, Asian, Japanese, European) depending on who is responsible for the access and benefit sharing negotiations with the country of origin. All investigated models however end up being burdensome and compromise somehow the open and timely sharing of genetic resources, especially the fast sharing in situations of public health emergencies. Will it be possible under the NP to create

exemptions or otherwise mechanisms for free, fair and fast sharing of essential information, and thus uphold the original vision and mission of GMI?

The Vision of Sharing

Eric Stevens

US FDA, USA

The importance of sharing WGS data cannot be overstated. Making the data available to all sectors within the One Health framework in real time enhances the impact of WGS for food safety and public health. This presentation will demonstrate the power of sharing WGS data and relate its impact on our understand of human genomic variation and molecular evolution to what it can do for microbial populations. Examples of how WGS has been used and could be used will be discussed as we work towards the vision of making WGS data sharing a global reality.

Global surveillance of antimicrobial resistance

Frank Moeller Aarestrup

DTU, DK

Antimicrobial resistance (AMR) is one of the most serious global public health threats, however, obtaining representative data on AMR for healthy human and animal populations are difficult. We have developed online bioinformatics tools and abilities for analyzing and sharing whole genome sequencing data in an standardized way facilitating open and global sharing of data. There can be several barriers for sharing data from clinical infections and isolates from clinical infectious provides limited information on healthy human and animal populations. We have developed metagenomics approaches allowing determination of the abundance and diversity of all AMR genes in metagenomics samples and are now using this to study the occurrence and transmission of AMR globally. Sampling is currently ongoing in +100 countries and initial results from a sub-set of the data will be presented. Part of the raw sequencing data are shared in real-time for everybody to use.

Biology and Epidemiology of Shiga Toxin-Producing E. coli – NGS Investigations

Eelco Franz

National Institute for Public Health and the Environment (RIVM), The Netherlands

Shiga toxin-producing *Escherichia coli* (STEC) are globally dispersed pathogens associated with a broad spectrum of clinical manifestations in infected humans, including diarrhoea, haemorrhagic colitis and (occasionally fatal) haemolytic uremic syndrome (HUS). STEC are generally considered zoonotic with ruminants, and particular cattle and sheep, as the main reservoirs. *E. coli* O157:H7 is the most commonly reported STEC serotype. Although evolutionary models have been developed the geographical spread of the pathogen and the extent of inter- and intra-continental transmission are still to be analysed comprehensively and quantitatively. We provide the first comprehensive global phylo-geographical analysis of STEC O157, reconstructing the phylogenetic history and global spread of the contemporary clones. Next to STEC O157 some STEC are truly emerging pathogens, including stx2f-producing STEC which represents a significant fraction of human isolates and so far only is

isolated from pigeons. We conducted an in-depth genomic comparison of *stx*$_{2f}$-carrying *E. coli* from pigeons and humans in order to contribute to the understanding of the ecology and epidemiology of this emerging group of STEC. We hereby identified a new hybrid *E. coli* pathotype.

Practical Issues in Implementing Next-Generation- Sequencing in Routine Diagnostic Microbiology

John WA Rossen

University of Groningen, The Netherlands

Next generation sequencing (NGS) is increasingly being used in clinical microbiology. Like every new technology adopted in microbiology, the integration of NGS into clinical and routine workflows must be carefully managed. As the microbiology laboratories have to adhere to various national and international regulations and criteria for their accreditation, quality control issues for using WGS in microbiology, including the importance of proficiency testing will be discussed. Furthermore, the current and future place of this technology in the diagnostic hierarchy of microbiology will be presented as well as the necessity of maintaining backwards compatibility with already established methods. Finally, the question of whether WGS can entirely replace routine microbiology in the future and the tension between the fact that most sequencers are designed to process multiple samples in parallel whereas for optimal diagnosis a one-by-one processing of the samples is preferred will be addressed.

Prospective Genomic Surveillance in a Clinical Environment: Tracking Resistance and its Mobilization

Lynn Bry

Harvard Medical School, USA

Diverse vectors mobilize drug resistance in pathogens, including via conjugative plasmids and transposons. While some forms of resistance are strongly associated with specific vectors, such as transmission of *Klebsiella pneumonia* carbapenemases (KPC) by the Tn4401 transposon, others, such as the New Delhi metallo-beta-lactamases (NDMs), have a more diverse vectors that mediate spread. In addition to specifying capacity for mobilization of resistance within and among species, the complement of mobile vectors also provides a robust set of information to assist with outbreak and strain cluster analyses. However, the complex and repetitive nature of these vectors introduces complexities in genomic analyses, and often requires specific technical, experimental and computational approaches to resolve carrying vectors and those involving nested structures. We present clinical cases that used mobile vector information to support gene-, vector-, strain- and cluster-level analyses to provide actionable data to support infection control and clinical microbiology operations.

The Impact of Pathogen Genomics in U.S. Public Health

Gregory Armstrong

Centers for Disease Control and Prevention, U.S.A

Since 2014, the US CDC, together with federal, state and local partners, has been adapting next-generation sequencing technologies in dozens of infectious-disease public health domains, such as food safety, tuberculosis control, influenza monitoring, and monitoring of antimicrobial resistance in viral, bacterial, and eukaryotic pathogens. This transition has been facilitated by the Advanced Molecular Detection (AMD) program, a $30M-per-year initiative to bring genomic sequencing, bioinformatics and related technologies bear against public health threats. The program is now in its fifth year and at a point where state and local health departments in the US are rapidly acquiring the technology and implementing it, in most cases, first for bacterial foodborne pathogen characterization, but increasingly in other areas. This talk will focus on the AMD program, how it has been fostering the adoption of pathogen genomics into the US public health system, and where the program is currently having the greatest impact.

NGS Provides Functional Insight into the Survival and Persistence of Bacterial Pathogens: The Case of Salmonella

Jie Zheng, Yu Wang, Elizabeth Reed, and Eric W Brown

U.S. Food and Drug Administration, U.S.A

Genetic adaptations to food and the food production environment observed in *Salmonella* and other pathogens associated with foodborne illnesses have become a public health concern. This highlights the probability of evolutionary changes in these pathogens, in which a selective advantage was conferred for survival, persistence and even growth within food matrices and in the environment, increasing their propensity for morbidity and mortality.

For example, sprouts contamination is one of the recurrent problems both in the United States and around the world. Between 2000 and 2016, at least 17 salmonellosis outbreaks linked to the consumption of raw sprouts were documented internationally, nine of them in the US. Most involved alfalfa sprouts, but cress, mung bean, and clover sprouts were also implicated. A metatranscriptomics approach was applied to examine *Salmonella* functions in sprout spent irrigation water (SIW). Interestingly, genes with different functions in *Salmonella* were observed to be highly transcribed at different time points during the sprouting process. This study sheds light on the active interaction of *Salmonella* with the sprout microbial community.

The application of NGS in food safety will give rise to a deeper understanding of *Salmonella* adaptations to certain environmental conditions and help identify preventive control measures to inhibit pathogen growth.

Next Generation Sequencing technologies for plant pest diagnostics
Baldissera Giovani and Françoise Petter

European and Mediterranean Plant Protection Organization, France

Reliable and rapid diagnostic processes are essential to support inspection activities conducted by National Plant Protection Organisations (NPPOs) in the framework of their official mandate, and to evaluate the efficacy of measures taken. Official controls aim to prevent or reduce the risk of introducing new pests through the agri-food trade and to protect consumer interests (Giovani *et al*., 2018).
During the last decades the incidence of plant diseases has increased exponentially in terms of both numbers and severity, as a result of increased trade of plants and plant products, advances in transport technologies and the development of a complex network of global commerce (Santini *et al*., 2017). The European and Mediterranean Plant Protection Organization recommend more than 300 pests (any species, strain or biotype plant, animal or pathogenic agent injurious to plants or plant products) for regulation. Regulated pests are officially controlled and countries organise surveillance pre-border, at border and at places of productions on commodities that are the hosts of a given pest.
By allowing the sequencing of the whole genome of (micro)organisms present on a commodity without *a priori* knowledge, NGS technologies can greatly support the work of official diagnostic laboratories and empower inspectors *in situ*. The presentation will provide an overview of the plant health sector and of the various international initiatives organised to promote and adopt this new diagnostic paradigm while taking into account specific challenges such as the detection of unknown organisms and the international sharing of data.

NGS in the Detection of Genetic Exchange in Streptococci and Staphylococci from Food, Human and Animal Sources

Christoph Jans

ETH Zurich, Department of Health Sciences and Technology, Laboratory of Food Biotechnology, Switzerland

Streptococci and staphylococci are among the most important genera in relation to public and animal health. Furthermore, certain species of these genera share a strong association with food, either in the role of foodborne pathogens, but also as organisms to produce fermented food. The ability to proliferate in these various niches suggest genetic adaptation as a contributor to handle the different environments. Food-derived streptococci (*Streptococcus bovis* group) and staphylococci (*Staphylococcus aureus*) obtained in sub-Saharan Africa represent exotic examples to study genetic adaptation and gene exchange in comparison to a global pool of strains of animal, human and food sources. Among the *S. bovis* group, a recombinant rather than clonal history also has important implications on the reliability of classical phenotypic and single gene genotypic identification approaches. This is exemplified by genetic exchange and genome decay in the *S. bovis* group that affect basic carbohydrate metabolism such as that of lactose but also virulence features such as adhesion abilities or capsule properties. Next generation sequencing (NGS) provides the tool to investigate these adaptations in combination with gaining insights into evolution,

population structure and genome plasticity presented on the examples of *S. bovis* and *S. aureus*.

Proficiency Tests in genomics – for UNSGM and GMI

Rene S. Hendriksen

Research Group of Genomic Epidemiology, DTU-Food.

WHO Collaborating Centre for Antimicrobial Resistance in Foodborne Pathogens and Genomics (WHO CC). European Union Reference Laboratory for Antimicrobial Resistance (EURL-AMR).

Technical University of Denmark, National Food Institute, Kgs. Lyngby, Denmark

Proficiency testing (PT) is an important tool to assess data leading to ensure reliable. The Technical University of Denmark, the National Food Institute (DTU Food) has a long track record for providing PTs to WHO and EU. Recently, DTU Food has launch with partners the UNSGM PT for the Secretary-General's Mechanism for Investigation of Alleged Use of Chemical and Biological Weapons as well as maintaining the GMI PT. The purpose of the UNSGM PT is to assess laboratories ability to correctly detect and characterize a biological threat and associated genetic markers such as virulence factors and antimicrobial resistance using modern high-throughput genomic technologies, whole genome sequencing (WGS) to increase the global and national preparedness. Similarly, the GMI PT aim at facilitating the production of reliable laboratory results of consistently good quality within the area of WGS. A total of 56 laboratories participated in the UNSGM PT with 90.1% of all laboratories correctly identified the genomes to species level. Thirteen laboratories obtain an overall score above 80%. In the GMI PT, 66 laboratories participated with the majority performing well. A few laboratories were identified as being outliers due to either conmtamination o fthe reference material or poor sequencing ability. Overall in both PTs, the programmed successfully identified the laboratories show a high level of proficiency despite the different scopes.

Establishment of Quality Control in PulseNet/GenomeTrakr

Eija Trees[1] and Ruth Timme[2]

[1]Centers for Disease Control and Prevention, U.S.A.
[2]Food and Drug Administration, U.S.A.

As whole genome sequencing (WGS) is established as the primary subtyping tool for foodborne pathogens it is imperative to implement a rigorous quality assurance / quality control (QA/QC) program to ensure the integrity and comparability of the data deposited to public or limited access repositories. PulseNet and GenomeTrakr, the two US networks sequencing clinical, food and environmental isolates for foodborne disease surveillance, outbreak investigations and attribution, are working together towards harmonized QA/QC program. The QA/QC program being established is based on a quality manual, a collection of electronic standard operating procedures for both wet and dry lab procedures that include multiple quality control points for each procedure. Minimum quality thresholds at each quality

control point must be met in order for the procedure to move to the next step. New staff undergoes training that has been harmonized between the networks and for PulseNet are also required to pass a competency assessment (certification) before submitting data to the network databases. A laboratory maintains its certification status by participating in an annual proficiency test that is harmonized across both networks.

CDC experience on using WGS for patient management – implications on QA/QC

Eija Trees

Centers for Disease Control and Prevention, U.S.A.

The implementation of whole genome sequencing (WGS) as a primary subtyping tool in public health laboratories will introduce an important paradigm shift also for the reference and diagnostic laboratories. Traditional workflows that have been used up until now to identify the species, serotype, virulence and antimicrobial resistance profiles and the strain subtype are now being replaced by one workflow based on genomic sequencing. In the United States, laboratory tests that generate data to be used to make decisions on patient management are regulated by the federal Clinical Laboratory Improvement Amendments (CLIA) established in 1988. CLIA requires laboratories to be certified by their state before they can accept samples of human origin for diagnostic testing. Before a certification can be obtained, documentation needs to be in place on assay validation, staff training, equipment maintenance, and quality control measures for assay performance. Any modifications to the workflow require full re-validation or re-verification depending on workflow change significance. Participation in two proficiency tests annually is required and certified labs are inspected every other year. At the US CDC, Enteric Diseases Laboratory Branch (EDLB) is the first, and so far the only, program to obtain CLIA certification for WGS-based diagnostics.

Tracking the Resistome in One Health Surveillance

Patrick McDermott* and Heather Tate

Food and Drug Administration, U.S.A

In the US, the National Antimicrobial Resistance Monitoring System has been using WGS of *Salmonella* as a tool of routine surveillance since 2013. To date, NARMS has generated MIC and WGS data on over 10,000 *Salmonella* isolates from food animals, 9,000 from human clinical cases, 6,000 from retail meats, and 1,4OO from imported foods. Analysis of this dataset showed that the presence of known resistance determinants is very highly correlated with clinical resistance, indicating that WGS data can be used to predict resistance in *Salmonella* strains lacking traditional antimicrobial susceptibility data. To make these data accessible, the NARMS program has launched Resistome Tracker and other online tools that provide visually informative displays of antibiotic resistance genes. Resistome Tracker harvests resistance gene information on a weekly basis from genomic data deposited at NCBI. It presents the resistome data using interactive dashboards that allows users to customize visualizations by antibiotic drug class, compare resistance genes across different sources, identify new resistance genes, and map selected resistance genes to geographic region. The tool also provides alerts about new resistance traits as they emerge in a region or source to provide early warning on emergent trends.

The NCBI Pathogen Detection Browser: Integrating Antimicrobial Resistance Genotypes and Phenotypes

William Klimke

National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, U.S.A.

The NCBI Pathogen Detection pipeline takes raw sequencing data from a surveillance network of public health labs sequencing food and clinical pathogen isolates and generates publicly available reports. The pipeline takes the raw sequencing data, assembles, annotates, and clusters the isolates into phylogenetic trees to aid outbreak and traceback investigations. The annotated assemblies are checked for the presence of antimicrobial resistance genes/proteins using a process built by NCBI called AMRFinder. AMRFinder uses a set of curated proteins and hidden markov models (HMMs) to assign functional names to proteins. AMRFinder has recently been tested on 6,242 isolates from the National Antibiotic Resistance Monitoring System (NARMS. These isolates were a mix of different species for which antibiotic susceptibility test (AST) were done. Comparison of genotype predictions from AMRFinder and the 89,318 susceptibility tests performed across the set of isolates showed that AMRFinder made a call consistent with the phenotype 98.1% of the time. The genotype calls and phenotype information are added to the Pathogen Isolates Browser: https://www.ncbi.nlm.nih.gov/pathogens/isolates#/search/

The EUCAST consultations on WGS for predicting antimicrobial susceptibilities

Matthew J Ellington

Antimicrobial Resistance and Healthcare Associated Infections (AMRHAI) Reference Unit, National Infection Service, Public Health England, London, UK.

WGS offers the potential to predict antimicrobial susceptibility from a single assay. The European Committee on Antimicrobial Susceptibility Testing established a subcommittee to review the current development status of WGS for bacterial antimicrobial susceptibility testing (AST). Barriers and issues to the wider adoption of the technology were considered, including whether epidemiological cut-off values (ECOFFs) or clinical breakpoints are the most appropriate comparators for genotypic and phenotypic data. Analysis showed that the use of WGS-inferred AST for guiding clinical decision was not supported by the evidence base, but this area should be a funding priority if it is to become a rival to phenotypic AST.

11th Global Microbial Identifier Meeting Presentation Abstracts

Day 3 Friday 18th May 2018

Application of WGS in food establishments: One Health context

Ivan Nastasijevic*, Brankica Lakicevic and Branko Velebit

Institute of Meat Hygiene and Technology, Serbia

Food safety is a global concern, and consumers have the right to safe and nutritious food. However, the estimated burden of foodborne diseases at global level, e.g. 600 million food borne illnesses and 420,000 deaths from 31 major food safety hazards in 2010 (WHO, 2015), as well as at EU level, with around 356,000 confirmed human illnesses and 452 deaths (EFSA, 2017) from 5 major food safety hazards (Campylobacteriosis, Salmonellosis, Yersiniosis, STEC infections, Listeriosis) and the related social and economic costs (hospitalization, loss of income, employment and market access) remain unacceptably high. Tracking the foodborne pathogens in the food processing – distribution – retail – consumer continuum is of utmost importance for facilitation of outbreak investigation and rapid action in controlling/preventing foodborne outbreaks. Whole Genome Sequencing (WGS) has recently emerged as a new tool and offers great potential in the way we investigate, assess and manage microbiological food safety issues, food borne illnesses, including associated antimicrobial resistance (FAO, 2016). The use of WGS can facilitate the understanding of contamination/colonization routes of foodborne pathogens within the food production environment and can also enable efficient tracking of entering routes of pathogens and their distribution from farm – to – consumer. WGS is a powerful tool for obtaining genomic data, which gives a higher level of resolution discrimination, i.e. better information about genetic similarity between isolates than conventional molecular typing such as Fluorescent Amplified Fragment Length Polymorphism (fAFLP) or Pulsed Field Gel Electrophoresis (PFGE); such molecular methods only determine if isolates are the same or different but not how closely related they are genetically. Implementation of WGS is likely to be beneficial for many countries in the foreseeable future, in support of food safety management systems and foodborne outbreak investigations. In recent years, development and availability of user-friendly portable real-time devices for DNA and RNA sequencing enabling ultra-long read lengths are possible (hundreds of kb). The cost of WGS has also lowered significantly allowing its use in more routine applications since the price for bacterial genome sequencing fall to less than $50/isolate, in case that a considerable number of isolates is sequenced at the same time on a given instrument to achieve maximum economy of scale. For instance, it may encompass in-house sequencing when sequencing equipment is used for multiple applications (WGS of pathogens, starter cultures and spoilage organisms). Further, WGS is currently becoming the method of choice for characterizing isolates of the major food borne pathogens (*Salmonella*, *Campylobacter*, *Listeria monocytogenes,* STEC) in national reference laboratories across USA and EU. Inter-sectoral cooperation in `One Health` context between environmental, veterinary and health authorities through establishment of national public database of isolates (non-food, food and human origin) will allow the real-time exchange of information and comparison of genetic similarity between isolates of foodborne pathogens. Such approach will significantly facilitate understanding of distribution of

pathogens and antimicrobial resistance along the food chain continuum and improve the effectiveness of outbreak investigations.

The Impact of NGS at the Intersection of Good Agricultural Practices and Human Food Consumption

Rebecca Bell

Food and Drug Administration, U.S.A

An estimated 14% of all foodborne outbreaks reported in North America are attributed to the consumption of raw or minimally-processed fruits and vegetables. In particular, consumption of fresh tomatoes has been linked to numerous foodborne outbreaks involving various serovars of *Salmonella enterica*. The association of *Salmonella* with tomatoes is ecologically complex and not yet fully understood. Successful mitigation strategies will depend on extensive environmental monitoring and understanding of the unique adaptive changes acquired by these human pathogens in the plant environment. An extensive environmental monitoring program for *Salmonella* has been ongoing along the VES since 2009. Next generation sequencing (NGS) has aided in the understanding of the ecology and biology within this non-host environment in several ways from traceback during outbreaks to functional genomics that may lead to new mitigation strategies to reduce or prevent contamination of fresh produce.

EFSA moving on: WGS activities for Food Safety in a European context

Beatriz Guerra, Valentina Rizzi, Ernesto Liebana

European Food Safety Authority, EFSA, Parma, Italy

Given the growing importance of WGS analysis for food safety (multinational foodborne outbreak investigations, surveillance/monitoring fields including AMR, use for approval of regulatory products including GMOs), the European Food Safety Authority (EFSA) has recognized the need to move on. In EFSA, several WGS activities, grouped under a WGS Umbrella project, are currently taking place. These include, among others, i) “In house” capacity building; ii) EC Mandates: ECDC-EFSA are exploring the possibility to extend the current collection and analysis of molecular typing data from foodborne pathogens in the joint ECDC-EFSA molecular typing database to WGS data; and EFSA is exploring the possibility to use molecular typing methods including WGS for the harmonized monitoring of AMR in bacteria transmitted through food. Iii) Funding research projects (ENGAGE, INNUENDO, GENCAMP, Liseq, others). iv) Using WGS analyses (performed by the EURLs) to support EFSA routine activities: e.g. for the European Summary Reports on AMR data (with EURL-AR and MSs to support the data quality and detection of emerging resistance mechanisms/resistant clones), as well as for rapid outbreak assessments (with ECDC and EURLs). v) Collecting information on the availability of WGS for the main foodborne

pathogens in animals, food, feed and their related environment in European food safety and veterinary laboratories.

Metagenomic Approaches for Complete Characterization of Human Enteric Diseases

Heather Carleton

US CDC, USA

PulseNet, the U.S. national molecular foodborne surveillance program, is preparing for a future without isolate culture. Current molecular surveillance relies on isolates which are less frequently submitted to public health laboratories due to increased use of culture independent diagnostic methods (CIDTs) by the clinical laboratories. PulseNet is currently developing pathogen detection, subtyping and strain characterization methods that by themselves are independent of cultures. PulseNet is pursuing two approaches: (1) near term - amplicon sequencing targeting specific pathogens including enough markers to provide sufficient resolution for outbreak detection and investigation similar to the cgMLST approach to subtype cultures of specific pathogens; and (2) long term- non-specific metagenomics by shotgun sequencing which may identify and subtype known pathogens as well as novel hitherto unrecognized pathogens. The latter approach will leapfrog pathogen discovery and likely the identification of known and novel pathogens causing outbreaks of unknown etiology.

Genomes from Metagenomes

Stephan C. Schuster

Nanyang Technological University, 60 Nanyang Drive, SBS-B1N-27, Singapore 637551

Microbial communities constitute the largest domain of life in terms of biomass, however, less than 1% of all microbial organisms are believed to be cultivatable. With the onset of next-generation sequencing surveys of non-cultivatable microbial assemblages have become accessible by metagenomics analysis. This, however, only results in fractions of the total genomic information of a given community and does not allow for the identification of cellular genome information. In this presentation, I will demonstrate several methods that allow dissecting microbial communities by single-cell techniques in combination with single-cell sequencing. In addition, the generation of genomes from metagenomes by deep sequencing and assembly of individual genomes from metagenomics data is presented.

WGS – the One Health linkage

Eric Brown

US FDA, USA

Whole genome sequencing now regularly underpins public health surveillance and food safety traceability efforts in the US and other countries where disease outbreak monitoring and response programs are commonly implemented. One expected and welcome shift in the application of WGS to food and feedborne safety has been a much more detailed understanding of the natural reservoirs and avenues of dissemination of human enteric pathogens throughout the farm to fork continuum, oft times including environmental niches, livestock and other agricultural settings, as well as water and food sources directly linked to human consumption and exposure. Thus, the agnostic nature and widespread availability of WGS data has made genomic sequencing an excellent tool for integration and surveillance in a One Health approach. The GenomeTrakr global foodborne pathogen database further advances this effort by providing a consolidated open-source and curated WGS database for tens of thousands of food and environmental bacterial pathogen sequences. Culture-dependent examples of this approach are well established and include WGS-based pathogen links between agricultural waters and the human food supply as well as linkage of drug resistance and other adaptive changes in Salmonella that ensure its survival throughout this continuum. Moreover, culture-independent or "quasi" independent WGS-based metagenomic approaches are starting to provide such One Health linkages directly from food and environmental samples. The emerging role of direct WGS and a novel "metaGenomeTrakr" database in this area will be discussed.

Virome Profiling of Sewage for Human Disease Surveillance

My Phan

Erasmus MC, The Netherlands

Detecting viruses in sewage might be a big game changer for viral surveillance. If the detection methods are sensitive enough, the proponents argue, we could detect viruses in sewage at early times before a large number of people are infected. What is missing from these discussions are some concrete numbers on the quantity of viral sequences that might be found in sewage, the dilution and the degradation of the viruses that that may occur between the patient and the sewage collection point, the stability of disease-causing viruses in sewage and the challenges of assembling useful viral sequence contigs from essentially pooled samples. I will discuss some of our experiences applying next-generating sequencing (NGS) to detect and quantify the viral sequence content of sewage, using the global sewage samples collected from 63 countries.

One Health Surveillance and Risk Prediction in Influenza

Ron Fouchier

Erasmus MC, The Netherlands

Today, genome sequence data are available from public databases for well over 200,000 influenza virus strains from humans and animals around the world. Increasingly, this dataset is used to inform policy, for instance for epidemiological tracing during surveillance studies. More and more, the data is also used to infer changes in virus traits that are important for public and animal health. Immune escape of human influenza viruses (“antigenic drift”) is increasingly monitored primarily by NGS during epidemics, followed by phenotypic confirmation of newly emerging variants to identify new potential vaccine candidates. Likewise, during influenza outbreaks in animals, viral genotypes are used to predict viral phenotypes associated with antiviral resistance or pathogenic, zoonotic or pandemic risks. I will discuss the current state-of-the art with respect to influenza virus genome sequencing and prediction of viral phenotypes from the genomic data during surveillance studies.

Bringing NGS to Diagnostic Virology

Sander van Boheeman

Erasmus MC, The Netherlands

Next-generation sequencing (NGS) has substantially improved the possibilities in clinical virology. Within one sequencing run NGS can perform a wide variety of diagnostic assays, e.g. detection and discovery, genotyping, resistance mutations, virulence markers, andtransmission markers. However, many hurdles need to be overcome to implement NGS in routine testing in the clinical laboritory. Turnaround times of many sequencing platforms are not competitive with the turnaround time needed to inform clinical practice and hospital epidemiology. Large datasets of reads are difficult to analyse and store, often needing specialisted expertise to manage. In metagenomic sequencing a significant amount of human sequences are created, leading to privacy issues. Other examples are contamination, database usage and internal control. In this presentation, some of the key challenges and possible solutions will be discussed.

Fast and cost-effective sequencing of RNA virus genomes in clinical samples

Ramette A*, Grädel C, Terrazos Miani MA, Barbani MT, Steinlin-Schopfer J, Bittel P, Suter FM, Leib SL

Human enteroviruses are small RNA viruses that affect millions of people each year and lead to a variety of symptoms ranging from mild illness to severe neurological disorders. Those viruses are routinely diagnosed by PCR assays, often combined with partial sequencing for genotyping. Due to high genomic variability, PCR-based approaches can lead to false negative results. Whole-genome sequencing (WGS) provides complete genetic information, but the approach is more expensive and time consuming, and not routinely used in diagnostic laboratories. Nanopore sequencing offers now the possibility for fast WGS and also enables direct RNA sequencing. We developed a wet lab protocol and bioinformatics pipelines for fast WGS of enteroviruses in clinical samples.

WGS of enterovirus RNA genomes were performed from cell cultures and directly from clinical stool samples using a MinION nanopore sequencer, a small portable device allowing long reads and real-time acquisition of sequences. We compared the accuracy of both amplified cDNA and direct RNA sequencing with that of Sanger sequencing, and determined the duration of the nanopore sequencing and the bioinformatic steps needed to obtain the highest consensus accuracy.

Nanopore sequencing of cDNA molecules from enteroviruses grown in cell cultures readily provided >95% coverage of enterovirus genomes (e.g. CV-A6, CV-A16, E18), and >99% consensus accuracy as compared to Sanger sequencing. Direct RNA sequencing of total RNA extracted from enterovirus-positive stool sample obtained in June 2017 provided very long reads, often covering the near-complete RNA genome of a *Coxsackievirus* A6 strain. Comparison of manual extraction (Trizol) vs. automated extraction (Easymag, Biomerieux) indicated that direct RNA sequencing is compatible with automated extraction procedure. Yet, high RNA loading quantities are needed to successfully sequence RNA natively, affecting the sensitivity of the approach. Extra polishing steps using nanopolish software were required to obtain higher consensus accuracies.

Enterovirus genomes extracted directly from clinical samples were successfully sequenced using the nanopore sequencing technology. cDNA sequencing provides long, high quality reads and RNA direct sequencing enables the fastest turnaround time for identifying enteroviruses in clinical samples. This study demonstrates the usefulness of fast sequencing technology in the diagnostic laboratories.

Viromes As Genetic Reservoir for the Microbial Communities in Food-Associated Environments: A Focus on Antimicrobial- Resistance Genes

Diego Mora

Department of Food Environmental and Nutritional Sciences, University of Milan, Italy

The description of viral populations was carried out in three different food-related environments with a focus on the identification and mobilization of antibiotic resistance genes (ARGs). Here reported three different cases of study that involve a comprehensive characterization of viral and microbial communities and the identification of ARGs in virome and microbiomes. In particular we have characterized: i) water sample from an experimental aquaculture plant, ii) viral communities present in the air surrounding cheese production area in two different dairy plants, and iii) water samples from Lambro River characterized by different levels of urbanization. Shotgun metagenomic sequencing was used to study both microbes and viruses, while *16S rRNA* profiling analysis completed the characterization of the microbial community. The three different cases of study showed how antibiotic resistance genes are wide spread in different environments irrespective of the presence of associated anthropic activities. In this context, we hypothesize that the release of antibiotics molecules in the environment by the microbiota is a driving force able to maintain ARGs in the microbiome and the consequent mobilization in the virome. Identification of ARGs and, more in general, all the microbial genes identified in the viromes of the different cases of study did not reflect the microbiome taxonomy, thus suggesting that microbial genes mobilized in the genome of viruses should be considered as a reminiscence of the past recombination events rather than a picture of the current microbial diversity. Beyond that, the analysis of the distribution of the different metagenomic profiles underlined that the variations both in terms of presence of genes and in terms of abundances are more visible in viromes than in microbiome that are less susceptible to gene fluctuation.

Proficiency testing for viral high-throughput sequencing

Annika Brinkmann[1], Claudia Wylezich[2], Andreas Andrusch[1], Ariane Belka[2], Thomas Nordahl Petersen[4], Pierrick Lucas[3], Yannick Blanchard[3], Anna Papa[5], Angeliki Melidou[5], Bas B. Oude Munnink[6], Jelle Matthijnssens[7], Richard J. Ellis[8], Florian Hansmann[9], Wolfgang Baumgärtner[9], Maurilia Marcacci[10], João Paulo Gomes[11], Dennis Schmitz[12], Victor Corman[13], Isabella Eckerle[13], Soizick F. Le Guyader[14], Julien Schaeffer[14], Rene S Hendriksen[15], Martin Beer[4], Andreas Nitsche[1#]

[1] Robert Koch Institute; Centre for Biological Threats and Special Pathogens 1, Berlin, GERMANY
[2] Friedrich-Loeffler-Institut; Institute for Virus Diagnostics, Greifswald–Insel Riems, GERMANY
[3] French Agency for Food, Environmental and Occupational Health & Safety, Laboratory of Ploufragan, Unit of Viral Genetics and Biosafety, Ploufragan, FRANCE
[4] Technical University of Denmark; Center for Biological Sequence Analysis, Department of Systems Biology, Lyngby, DENMARK
[5] Aristotle University of Thessaloniki; School of Medicine, Microbiology Department, Thessaloniki, GREECE
[6] Erasmus Medical Centre, Rotterdam, THE NETHERLANDS
[7] REGA Institute KU Leuven, Leuven, Belgium
[8] Animal and Plant Health Agency, Addlestone, UNITED KINGDOM
[9] University of Veterinary Medicine; Department of Pathology, Hannover, GERMANY
[10] Istituto Zooprofilattico Sperimentale dell'Abruzzo e Molise "G. Caporale", Teramo, ITALY
[11] Instituto Nacional de Saúde Doutor Ricardo Jorge; Department of Infectious Diseases, Lisbon, PORTUGAL
[12] National Insitute for Public Health and the Environment; Bilthoven, THE NETHERLANDS
[13] University of Bonn Medical Centre; Institute of Virology, Bonn, GERMANY
[14] Institut français de recherche pour l'exploitation de la mer; Laboratoire de Microbiologie, Nantes, FRANCES
[15] Technical University of Denmark; National Food Institute, Lyngby, DENMARK

High-throughput sequencing (HTS) has enabled the identification of several viruses and is becoming increasingly important in clinical settings and disease outbreaks. The presented virus proficiency test (PT) was designed to contribute to the standardization of HTS "wet lab" and "dry lab" procedures, with the main goal to improve the detection, identification and analysis of viral pathogens in diagnostic samples. The participants' proficiency in identification and analysis of viral pathogens in complex samples and HTS datasets was evaluated in two general parts of the virus PT: (1) the laboratories' proficiency to analyze HTS data and identify viral pathogens, (2a) the laboratories' sample preparation and sequencing procedures and (2b) the laboratories' HTS performance and output. Therefore, participants of the PT were invited to (1) analyze a complex metagenomics fastq dataset to identify viral sequences of 4 viruses in varying numbers, (2a) perform purification, DNA/RNA extraction, library preparation and sequencing of a virus strain within a complex sample matrix and (2b) perform library preparation and sequencing of viral RNA/DNA. The results summarized in this talk indicate that numerous protocols, tools and different workflows are used for virus HTS and results of such workflows differ in sensitivity and specificity. So far, there are no standard procedures for sample preparation and virome analyses, and sharing and comparing reliable results of such analyses remain difficult.

11th Global Microbial Identifier Meeting

Abstracts for the Poster Presentations

Evergreen: a platform for surveillance of bacterial outbreaks

JUDIT SZARVAS*[1] & JOHANNE AHRENFELDT*[1], JOSE LUIS BELLOD CISNEROS[1], MARTIN CHRISTEN FRØLUND THOMSEN[1], FRANK MØLLER AARESTRUP[1], OLE LUND[2]

[1] Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark

[2] Genomic Epidemiology, Department of Bio and Health Informatics, Technical University of Denmark, Kongens Lyngby, Denmark

Whole-genome sequencing has an expanding role in public health with the technological advancements in high-throughput sequencing. Public health authorities sequence thousands of pathogenic isolates each month for microbial diagnostics and surveillance of pathogenic bacteria. Whole genome sequencing data could be used to infer the phylogeny of the isolates, thus spot outbreaks and connect cases with potential sources. Moreover, uploading the data to public repositories facilitates the surveillance of pathogenic bacteria on an international scale. We have built a platform for the surveillance of pathogenic bacteria, which incorporates publicly available sequencing data. The Evergreen pipeline downloads the newly available whole genome sequencing data daily from the public repositories. To decrease the computational burden, the data is divided into sets by matching the isolates to a closely related reference genome. The reads are mapped to the reference to gain a consensus sequence and the SNP based genetic distance is calculated between all the sequences that were mapped to the same reference. Sequences are clustered together with a threshold of 10 SNPs to reduce the redundancy in each set. Finally, phylogenetic trees are inferred from the non-redundant sequences and the clustered sequences are placed on a clade with the cluster representative sequence. By observing the trees made during the pilot phase, where we are only looking at *E. coli*, *Listeria*, *Shigella*, *Campylobacter* and *Salmonella*, we have identified several known outbreaks. If the data had been uploaded real time, we could have discovered these outbreaks ongoing and helped to prevent further outbreaks.

GenomeTrakr database and network: WGS network for real-time characterization and source tracking of foodborne pathogens

M. Allard*, R. Timme, M. Sanchez, E. Stevens, M. Hoffman, K. Yao, G. Kastanis, D. Miller, Tim Muruvanda, S. Lomonoco, E. Strain, R, J. Payne, A. Pightling, H. Rand, J. Pettengill, Y. Luo, Narjol Gonzalez-Escalona, D. Melka, Sabina Lindley, Y. Chen, S. Tallent and E. Brown

Food and Drug Administration, College Park MD USA

A national database of federal, state, academic and international laboratories has been using WGS data to rapidly characterize pathogens. This mature GenomeTrakr network is part of NCBI Pathogen Detection web site. Public health agencies (FDA, CDC and USDA FSIS) collect and share data in real time. This high-resolution, rapidly growing database is actively being used in outbreak investigations at state, national, and international levels. GenomeTrakr database has demonstrated how distributed network of desktop WGS sequencers can be used in concert with traditional epidemiology and investigation for source tracking of foodborne pathogens. This new “open data” model allows greater transparency between federal/state agencies, industry partners, academia, and international collaborators. This database has continued to grow and diversify the foodborne pathogen database doubling in the last year to ~170,000 draft genomes. Two new international surveillance efforts were added to collect food, animal and environmental isolates including *Campylobacter*. NCBI has release new data analysis tools that improve rapid interpretation and visualization. NCBI, currently is producing daily clustering results for 22 pathogens including: *Salmonella*, *Listeria*, *E. coli Campylobacter* and *Vibrio*. The high-resolution WGS data in concert with epidemiological or inspection evidence has drastically enhanced our ability to identify the food sources of current outbreaks for foodborne pathogens with ~200 regulatory clusters examined in 2017. Results demonstrate global benefits of having an open data model. Understanding root causes of foodborne contamination assists our academic, public health and industry partners to develop preventative controls to make food safer globally.

A validation approach of an end-to-end whole genome sequencing workflow for source tracking of *Listeria monocytogenes* and *Salmonella enterica*

Anne-Catherine PORTMANN[1], Coralie FOURNIER[2], Johan GIMONET[1], Catherine NGOM-BRU[1], Caroline BARRETTO*[1], Leen BAERT[1]

[1]Institute of Food Safety and Analytical Science, Nestle Research Center, Lausanne, Switzerland
[2]Multi-Omics Profiling, Nestlé Institute of Health Sciences SA, Lausanne, Switzerland

Whole genome sequencing (WGS), due to its high discriminatory power, is routinely being used for source tracking, pathogen surveillance and outbreak investigation. In the food industry, WGS used for source tracking is beneficial to support contamination investigations. Despite its increased use worldwide, no standards or guidelines are available today for its use in outbreak and/or trace-back investigations. The differences between genomes identified by WGS need to be trusted and a validation of all steps of the WGS workflow is therefore recommended. Here we present a validation of an end-to-end WGS workflow for *Listeria monocytogenes* and *Salmonella enterica*, including isolates sub-culturing, DNA extraction, sequencing and bioinformatics analysis. The following performance criteria were assessed: stability, repeatability, reproducibility, discriminatory power and epidemiological concordance. Few SNPs were observed for *L. monocytogenes* and *S. enterica* when comparing isolate sequences derived from the same subculture and between isolates after 10 subcultures. Consequently, the stability of the WGS workflow for *L. monocytogenes* and *S. enterica* was demonstrated despite the few genomic variations that can occur during sub-culturing steps. Repeatability and reproducibility were confirmed. The WGS workflow has a high discriminatory power and confirms genetic relatedness. Additionally, the WGS workflow was able to reproduce published outbreak results, illustrating the epidemiological concordance. The current study proposes a validation approach comprising all steps of a WGS workflow and demonstrates that the workflow can be applied to *L. monocytogenes* or *S. enterica*. This work is one of the first steps to the harmonization of WGS methodologies for source tracking.

Microbial Stock Culture Bank a Panacea to Mitigate the Outbreak of Infectious Diseases in the Developing World

Ayuba Sunday Buru

Genomic Research Laboratory, College of Medicine, Kaduna State University, Kaduna, Barau Dikko Teaching Hospital (BDTH) Kaduna

One of the most important, yet often neglected tasks in any routine microbiology laboratory is to preserve a collection of viral, bacterial and fungal as stock cultures. The microbial stock culture bank seek to have a more comprehensive and all inclusive data of medically important microorganism isolated daily from patient samples attending Barau Dikko Teaching Hospital. The isolated microorganisms will be identified, characterized and the genomic sequence will be established (phylogenetic tree), the data generated will be stored, access and research on, cutting across different thematic research areas. Which include but not limited to the Genomic Research group, Microbiome Research Group, Microbial phyto-therapy Research group, Antimicrobial surveillance program, epigenetics, transcriptomics, Epidemiology and Disease Surveillance Program etc. The emergence and re-emergence of resistant strains of microorganism is a wake up call for all to come together and develop a more robust public health template to mitigate the infection and to equally prevent cases of outbreak of dangerous and difficult to treat microorganisms.

This project will require all to be onboard to ensure we have a fully functional Genomic Research Laboratory that can embark on novel research that will contribute greatly to the scientific community.

Breaking WGS: Impact of sequence quality on bacterial whole-genome sequence (WGS) analysis for public-health microbiology

CATHERINE CARRILLO*, ADAM KOZIOL, ANDREW LOW, PAUL MANNINGER, BURTON BLAIS.

***Ottawa Laboratory (Carling), Canadian Food Inspection Agency, Ottawa, Ontario, Canada**

With the increasing adoption of whole-genome sequencing (WGS) for public-health microbiology worldwide, development of standards for WGS data is a pressing need. In the context of public-health microbiology, sequence data are commonly used to determine if strains are related (e.g. to a cluster of illnesses), or to infer strain phenotypes (e.g. virulence, antimicrobial resistance, serotype) based on identification of genetic markers. The purpose of this study was to determine how variations in sequencing methodology impacted interpretation of WGS of foodborne pathogens. Variations in growth of bacterial cultures, and DNA and sequencing library preparation methods were evaluated to characterize the influence of these parameters on downstream analyses, using the Illumina Nextera XT/ MiSeq sequencing platform as the model system. Optimization of bacterial growth conditions and DNA extraction methods were associated with more efficient sequence generation; however, if minimum coverage requirements were met, no impacts on downstream analyses were observed. In contrast, gaps in sequence coverage were identified when sequencing library size selection methods were used. In cases of sequence contamination with another strain, clustering of WGS data was impacted. However, contamination could be detected using bioinformatics tools (ConFindr) developed for this purpose. While there may be an economic gain to optimizing wet-lab sequencing methods, in most cases this did not influence data interpretation. Given the rapid advances in high-throughput sequencing methods, development of international standards for WGS of foodborne pathogens applicable to all sequencing platforms will be challenging. Such standards should focus on parameters that significantly impact WGS data interpretation.

PlasmidID: a mapping based tool for plasmid characterization and visualization

Pedro J. Sola-Campoy[1],[2], Sara Monzón[2], María Pérez-Vázquez[1], Belén Aracil[1], José Campos[1], Jesús Oteo[1], Isabel Cuesta[2]

[1] Reference and Research Laboratory for Antibiotic Resistance, CNM, ISCIII, Majadahonda, Madrid.

[2] Bioinformatic Unit, ISCIII, Majadahonda, Madrid.

Plasmids are primarily responsible for dissemination of antibiotic resistance genes (ARG), virulence factors and other accessory genes that enable bacteria survival and spread in hospitals. The modular nature of plasmid structure hinders its identification and further comparison using the most usual high-throughput sequencing techniques.
Here we report a plasmid identification tool (PlasmidID) and its applications for the analysis of bacterial plasmidome. It is a mapping-based, assembly-assisted plasmid characterization tool that analyzes and gives graphic solution for plasmid identification. lllumina reads from reference *Escherichia coli*, *Klebsiella pneumoniae* and *Salmonella enterica* were simulated to evaluate PlasmidID and adjust default parameters. Finally, the utility and efficacy of PlasmidID was tested in three clinical isolates of *K. pneumoniae* that harboured different carbapenemases and were responsible for outbreaks in Spain (ST11/OXA-48-like, ST147/VIM-1 and ST512/KPC-3).
PlasmidID is a bash script that maps reads over NCBI plasmid database sequences. The most covered sequences are clustered by identity to avoid redundancy and the longest are used as scaffold for plasmid reconstruction. Reads are assembled into contigs and those are annotated, including ARG, replisome and any user defined genes or sequences. All information generated from mapping, assembly, annotation and local alignment analyses is gathered and accurately represented in a circular image which allow user to determine plasmidic composition in any bacterial sample.
Since it is database dependent tool, it works for every plasmid length and species, and can be adapted to other aims such metaplasmid analysis or SMRT sequencing technologies

WGS-Outbreaker: A deep analysis tool for diagnostic microbiology and outbreak investigation in a public health microbiology laboratory

S. MONZON[1], A. HERNANDEZ[1], PJ SOLA[1,2], L. CHAPADO[1], S. HERRERA-LEON[3], M.T. LLORENTE[3], P. JIMENEZ[4], A. ZABALLOS[4] and I. CUESTA*[1]

[1]Bioinformatics Unit, Core Scientific and Technical Units, ISCIII, Spain
[2]Reference and Research Laboratory for Antibiotics Resistance CNM-ISCIII, Spain
[3]Reference and Research Laboratory for Food and Waterborne Bacterial Diseases CNM-ISCIII, Spain
[4]Genomics Unit, Core Scientific and Technical Units, ISCIII, Spain

Investigation of foodborne bacterial pathogens outbreaks is going through regeneration due to the development of high-throughput sequencing techniques, which provide higher genomic resolution and full genetic information than the gold standards methods. WGS-Outbreaker aims to facilitate the incorporation of Whole Genome Sequencing (WGS) in a microbiology reference laboratories with a public health role, for outbreak investigations and diagnostic microbiology. This workflow implements a comprehensive analysis, including bacterial species identification, sequence-based typing, presence of plasmids and resistance genes, SNP calling and phylogenetic relation inference. WGS-Outbreaker is developed for Illumina reads analysis and uses reference-based mapping approaches. Species identification is carried out by kmerfinder. SRST2 analyses the presence of plasmids and resistance genes, and also identifies the sequence type. For SNP calling, users can choose among CFSAN snp-pipeline or in-house pipeline using BWA and GATK , or both. Finally, a phylogenetic tree is made in with RAxML. In addition, a module for testing quality of reads and statistics of mapping and variant calling has been implemented. WGS-Outbreaker is written in bash. A configuration file is defined by the user with the desired analysis. WGS-Outbreaker is able to run in High Performance Computing environment as well as in local. The WGS analysis of Salmonella enterica´s outbreaks performed with WGS-Outbreaker agreed with results obtained by routinary microbiological techniques, outperforming them in execution time and resolution, allowing in addition the substitution of multiple laboratory tests by a single one.

A human *Campylobacter fetus* outbreak identified through next generation sequencing

Birgitta Duim[1,7], Aldert Zomer[1,7], Hetty Koppenaal[2], Fleur Groenendijk[2], Marcel van den Berge[3], Erwin Verkade[3], Aloys Tijsma[4], Marcel Spierenburg[4], Kees Verduin[5], Margreet te Wierik[6] and Jaap A. Wagenaar[1,7].

Utrecht University, Faculty of Veterinary Medicine[1], Public Health Service Zeeland[2], Admiraal de Ruyter Hospital[3], Netherlands Food and Consumer Product Safety Authority[4], Amphia Hospital[5], National Intitute for Public Health[6], WHO-collaboration center for Campylobacter[7].

Campylobacter fetus subsp. *fetus* (*Cff*) is primarily a veterinary pathogen and occasionally associated with severe infections in humans. From May-October 2015 an unexpectedly high number (6) of invasive *Cff* infections in human patients were reported in the province of Zeeland (n=5) and the neighbouring province of Brabant (n=1) in the Netherlands. Based on patient questionnaires, it was concluded that the patient from Brabant most probably acquired the infection abroad, whereas all patients in Zeeland consumed unripened sheep cheese from unpasteurized milk. For 4 patients, the product could be traced to one sheep farm, while for the 5th patient a second farm was indicated. Microbiological investigation of sheep faeces revealed *Cff* in sheep from the first farm but not of the second. MiSeq sequencing of isolates from patients, and the suspected sheep flock, identified that all isolates belonged to MLST ST6. The genome SNP phylogeny irrevocably showed that 5 isolates belonged to a distinct clone while the 6th isolate from the unrelated patient belonged to a separate clone. The core genome of 5 isolates generally differed between 0-7 SNPs, while epidemiological unrelated human isolates differed by at least 26 SNPs. The outbreak isolates harboured determinants for serum resistance (*glf*) and the surface array protein (*sap*), both associated with invasive infections. The genomes from the sheep isolates were identical to the patient isolates, confirming the epidemiological source identification. In conclusion, NGS confirmed that consumption of unpasteurized sheep cheese lead to an outbreak of *Cff* infections.

Assembly and Comparison of Whole Genomes of Different Bacteria using the Nextera™ DNA Flex Library Prep Kit

R.Esson[1]*, C.teiling[2], S.Falque[1], E. Baud[2], N. Nougarede[1], E. Abachin[1]

[1] Analytical R&D Department, Sanofi-Pasteur, Marcy l'Etoile – France

[2] Illumina, Inc., San Diego, CA.

Analyses of bacterial genomes provide information on genomic composition and organization. Nucleic acid sequence is mandatory to study virulence factors, antibiotic resistance, epidemiology, and nosocomial outbreaks. Over the last 15 years, NGS (Next Generation Sequencing) has become the most common approach for bacterial genome sequencing. While the handling time to perform sequencing and analysis has been reduced, the number of samples to be tested has dramatically increased.

The objective of the study was to perform an evaluation of the Nextera DNA Flex Library Prep kit for rapid DNA library preparation, the first step of the entire sequencing workflow. This protocol utilizes bead-based “tagmentation” chemistry by combining DNA fragmentation and adapter tagging in a single reaction on a solid support, reducing the need for accurate quantification at inputs greater than 100ng. Here we tested bacterial cultures and colonies as input by selecting the following bacterial species: *Escherichia coli*, *Staphylococcus aureus*, *Bordetella pertussis* and *Clostridium tetani*. Three different amounts of DNA (100 ng, 300 ng, and 500 ng) were used as input. Sequencing data were analyzed through different pipelines (Spades for assembling the genomes, SRST2 for MLST comparison, Bacterial analysis Pipeline, and others), all available on the BaseSpace Hub.

We have combined key features of a new innovative workflow to analyze whole bacterial genomes. Nextera DNA Flex Library Prep kit provided a method for rapid extraction of DNA from colonies, to a normalized library, ready for sequencing. This study evaluated multiple species of bacteria and demonstrated the suitability of the kit on projects where bacterial genomic composition is required. This kit can be applied to other bacterial species without any additional development needed.

Genotypic and phenotypic characterization of antimicrobial resistant *E.coli* isolated from ready-to-eat food in Singapore based on disk diffusion, broth microdilution and whole genome sequencing

Siyao Guo[1,2], Moon Y.F. Tay[1,2], Kyaw Thu Aung[1,2,4], Kelyn Seow Lee Ghee [1,2], Man Ling Chau[4], Ramona Alikiiteaga Gutiérrez[4], Lee Ching Ng[3,4], Jorgen Schlundt[1,2]*

[1] Nanyang Technological University Food Technology Centre (NAFTEC), 62 Nanyang Drive Singapore 637459

[2]School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive Singapore 637459

[3] School of Biological Science, Nanyang Technological University, 60 Nanyang Drive Singapore 637551

[4] Environmental Health Institute, National Environment Agency, 11 Biopolis Way, Helios Block #06-05/08 Singapore 138667

This study aimed to determine the antimicrobial resistance profile of *Escherichia coli* isolated from ready-to-eat food sold in hawker centers in Singapore. A total of 99 strains were isolated between 2009 and 2014 from poultry-based dishes (n=77) and fish-based dishes (n=22). The results from disk diffusion assay showed that 24 out of 99 *Escherichia coli* strains were resistant at least one antimicrobial agent. Thereafter, the 24 isolates were further subjected to further phenotypic (broth microdilution) and genotypic (whole genome sequencing (WGS)) characterization. Thirty-three antimicrobial agents (including classes of β-lactem, aminoglycosides, tetracycline, fluoroquinolones, polymyxin and others) were used for the MIC testing, which showed that 15 isolates carried multi-resistance (defined as resistance to three or more classes of antimicrobials). Two isolates which were resistant to multiple β-lactem antimicrobial agents were confirmed as Extended-Spectrum β-lactem (ESBL)- producing bacteria by double disk synergy method. Based on WGS data, online analysis tool ResFinder detected 7 classes of antimicrobial resistance genes and resistance-related chromosomal point mutations. For most classes of antimicrobials, the AMR genotype showed high consistency with the phenotypes. Most of the AMR genes were likely to be located on plasmids, increasing the risk of AMR transfer. Four known gene mutations which may cause antimicrobial resistance were detected. This study shows significant level of AMR in contaminating *E.coli* isolates from ready-to-eat food, and suggest potential risk of AMR transfer from ready-to-eat food to human.

Genotypic and phenotypic characterization of *Salmonella enterica* isolated from chicken meat and human in Sri Lanka

Moon YF Tay, Sujatha Pathirage, Kelyn LG Seow, Uddami SGH Wickramasuriya, SAN Sadeepani, Lakshi R. Palihawadana, IUW Ekanayaka, Saranga I Fonseka, Gayani Thusharika, Masami T Takeuchi, Jørgen Schlundt

A total of seventy-nine *Salmonella enterica* isolates from chicken meat and human in Sri Lanka were characterized by serotyping, disc diffusion assay and whole genome sequencing. Agona (ST13) was the most common serovar in chicken meat whereas Enteritidis (ST11) was the most common serovar in human. The level of agreement between serotyping and serotype prediction results was 92.4%. Among the thirty-three chicken meat isolates, 87.9% of them had at least one resistance gene and the most frequent resistance genotype was fosA7, which was only present in all sixteen Agona strains. The maximum number of resistance genes is six and they were found in two Kentucky (ST314) isolates. On the contrary, among the forty-six human isolates, only 28.3% of them had at least one resistance gene and the most frequent resistance genotype was QnrS1. The maximum number of resistance genes is nine and they were found in one Typhi (ST1) isolate. fosA7 gene was also found in only one human isolate that is of serovar Mountpleasant. In summary, this study identified serovars that were dominating in chicken meat and human, and showed the difference in antimicrobial resistance genetic profile of chicken meat and human strains from Sri Lanka.

Comparative genomics of *Lactobacillus reuteri* strains isolated from chicken gut

ANNA GREPPI[1]*, PAUL TETTEH ASARE[1], NIKLAUS ZEMP[2], CLARISSA SCHWAB[1], CHRISTOPHE LACROIX[1]

[1] Institute of Food, Nutrition and Health. Laboratory of Food Biotechnology, ETH, Zurich.

[2] Genetic Diversity Centre, ETH, Zürich

Lactobacillus reuteri is a probiotic bacterium encountered in the gastrointestinal-tract of humans and also found in many animals including rodents, pigs, chicken. The poultry isolates of *L. reuteri* can convert glycerol into reuterin which is a potent antimicrobial system. We currently investigate *L. reuteri* producing-reuterin for *Campylobacter* prevention, and the genetic diversity of this important species in poultry. We sequenced whole genomes of 25 *L. reuteri* strains isolated from chicken gut for comparative genomic analysis. Illumina reads were quality filtered and trimmed by Flash and Trimmomatic, and trimmed paired-end reads assembled using Spades. The overall quality of assemblies was evaluated by Quast and BUSCO. Plasmids were identified using a customized database. Assembled genomes were then annotated using Prokka and Rast.

The average number of paired-end reads across all genomes was 452,155. Assembled genomes size ranged between 1.95 Mbp and 2. 7 Mbp with 68 to 657 contigs longer than 500 bp, and an N50 between 22978 and 69730. Of the 443 BUSCO genes we found between 431 and 434 single copy genes indicating a high quality of the assemblies. Prokka annotation predicted a total number of coding sequences (CDS) from 1909 to 2298; all genomes contained 1 transfer message RNA (tmRNA), 45 to 68 transfer RNAs (tRNA), and 3 to 6 ribosomal RNAs (rRNA). Seven genomes contained 1 repeat region. Annotations

comparison is used for deep comparative genome analysis and combined with phenotypic traits to identify best candidates for preventing *Campylobacter* in chicken and reducing antibiotic burdens in poultry farms.

GenEpiO and FoodOn: Enabling data interoperability for infectious disease surveillance, investigation and control

EMMA GRIFFITHS[1]*, DAMION DOOLEY[2], GURINDER GOSAL[2], NABIL-FAREED ALIKHAN[3], MARIA SANCHEZ[4], THOMAS MATTHEWS[5], AARON PETKAU[5], JOSH ADAM[5], RUTH TIMME[4], MORAG GRAHAM[5,6], GARY VAN DOMSELAAR[5,6], FIONA BRINKMAN[1], WILLIAM HSIAO[2,7]

[1]Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, BC, Canada

[2]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada

[3]Department of Microbiology and Infection, Warwick Medical School, University of Warwick, Coventry, UK

[4]Office of Regulatory Science, US Food and Drug Administration, College Park, Maryland, USA

[5]National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, MB, Canada

[6]Department of Medical Microbiology and Infectious Diseases, Max Rady College of Medicine, University of Manitoba, Winnipeg, MB, Canada

[7]British Columbia Centre for Disease Control Public Health Laboratory, Vancouver, BC, Canada

The ability to share data between organizations is crucial for global, real-time infectious disease surveillance, investigation and control. Reliable capture and harmonization of whole genome sequencing (WGS) contextual information (sample source, experimental and bioinformatics methods, lab, clinical and epidemiological data) is critical for the interpretation of WGS results used for decision making in health crises. This data is often recorded using free text and institution-specific data dictionaries, requiring time-consuming and error-prone transformation before it can be used in investigations. Ontologies provide hierarchies of well-defined, standardized vocabulary enabling comparisons at different levels of granularity; universal IDs for disambiguating terms; built-in logic enhancing querying power; and synonyms that enable institutions to use preferred terms while linking to a standard, improving interoperability. We have created two ontologies to better harmonize and integrate genomics data into food microbiology and public health workflows, called the Genomic Epidemiology Ontology (GenEpiO) and the Food Ontology (FoodOn). To better implement our ontologies, we have also created a data parsing/text matching tool, and an ontology-driven data specification platform called the Genomic Epidemiology Entity Mart (GEEM). These tools and resources are currently being tested and evaluated for use in an International Organization for Standards (ISO) standard for the implementation of WGS for food microbiology, as well as key databases and platforms for typing and tracking foodborne pathogens – Enterobase, GenomeTrakr and IRIDA. The improved inferencing and computability of harmonized data provided by our work can enhance communication and

analyses, resulting in faster hypothesis generation during investigations, and ultimately, better health outcomes.

Validating cgMLST of S*almonella enterica* by comparing cgMLSTFinder and EnteroBase

CAMILLA HUNDAHL JOHNSEN*1, PIMLAPAS LEEKITCHAROENPHON1, ROLF KAAS1, PHILIP T.L.C. CLAUSEN1, OLE LUND2, AND FRANK M. AARESTRUP1

1 Research Group for Genomic Epidemiology, National Food Institute, Technical University of Denmark, Kongens Lyngby, Denmark.

2 Genomic Epidemiology, Department of Bio and Health Informatics, Technical University of Denmark, Kongens Lyngby, Denmark.

At Center at Genomic Epidemiology we have for many years provided users with a service for
detecting Multi-Locus Sequence Types (MLST) based on whole-genome sequencing (WGS) data (https://cge.cbs.dtu.dk/services/MLST/). However, due to limited number of genes, MLST sometimes lacks discriminative power. WGS data can provide sequence information across an entire bacterial genome and MLST can thereby be extended to be assigned for a defined core genome (cgMLST). Accordingly, we have generated a cgMLST service for 5 species, including *Salmonella enterica* (https://cge.cbs.dtu.dk/services/cgMLSTFinder/). The core genome schemes are obtained from PubMLST, Institute Pasteur and EnteroBase. We have validated cgMLSTFinder on the salmonella cgMLST scheme using a dataset containing 262 *Salmonella* isolates. We compared the detected ST type with EnteroBase determined cgMLST and compared the respective allele-profiles. The agreement between cgMLSTFinder and EnteroBase on the ST level was 92,3%. When looking into the allele-profiles we found that, on average, 14 out of 3002 salmonella core genome loci were assigned with different allele numbers when comparing cgMLSTFinder and EnteroBase. Notably, the different allele calls were not randomly distributed over the core genome. A limited number of loci were very frequently seen to disagree. We also found that many of the loci that showed a high frequency of disagreement were duplicated in many isolates. This suggest that the defined core genome should be reevaluated. Additionally, this study illustrates the large effect different mapping methods have on allele calling.

Challenges for the implementation of WGS of pathogenic isolates for surveillance and outbreak investigation in the context of a public health institute working under ISO accreditation

S. C.J. DE KEERSMAECKER[1*], S. HOFFMAN[1], M. DELVOYE[1], E. VANDERMASSEN[1], W. MATTHEUS[2], P.J. CEYSSENS[2], V. MATHYS[2], S. BERTRAND[2], S. DENAYER[3], B. VERHAEGEN[3], C. GARCIA-GRAELLS[3], N. BOTTELDOORN[3], A.M. VANHERLE[4], P. CLIQUET[4], A. SALTYKOVA[1], J. VAN BRAEKEL[1], B. BOGAERTS[1], K. VANNESTE[1], N. ROOSENS[1]

[1] Platform Biotechnology and Bioinformatics, Sciensano, Brussels, Belgium;
[2] National Reference Centre for *Salmonella* and *Shigella*, NRC – *Neisseria meningitidis*, NRC-*Listeria*, NRC-Tuberculosis and Mycobacteria, Scientific Service Bacterial Diseases, Sciensano, Brussels, Belgium
[3]National Reference Laboratory of Foodborne Outbreaks, NRL *Salmonella*-Food, NRL-*Listeria*, *NRL-E. coli*, NRL-antimicrobial resistance, Scientific Service Food Pathogens, Sciensano, Brussels, Belgium
[4] Quality Assurance, Sciensano, Brussels, Belgium

Whole Genome Sequencing (WGS) has the potential to replace currently used classical (sub)typing and characterization methods in the National Reference Centers (NRCs) and National Reference Laboratories (NRLs) for pathogen surveillance and outbreak investigation. However, before this actually can occur in the context of a public health institute working under ISO accreditation, this new technology needs to be validated to demonstrate it is fit for purpose within a quality system, and within an acceptable cost setting. This applies to both the wet and dry lab (bioinformatics) aspects. However, although several initiatives are ongoing, currently no official harmonized guideline is available. Based on the existing literature, we selected the performance metrics to be validated (trueness, repeatability, intra- and inter-reproducibility, and robustness) and their corresponding acceptance criteria. The quality of the obtained data was evaluated against these criteria. Additionally, downstream analysis with specific bioinformatics tools of WGS data of strains previously characterized, allowed evaluating whether the quality of the data was sufficient for an accurate characterization by WGS. The workflow we elaborated for the validation of the WGS technology to be used within the context of a public health institute allowed us drawing up the required standard operating procedures and applying for ISO 17025 accreditation. It would be interesting that international WGS guidelines become available to facilitate standardization and harmonization of WGS procedures, including the accreditation process itself, assessed by different auditors. This will assure portability and inter-laboratory exchangeability of WGS data and interpretation, which is crucial for accurate surveillance and outbreak investigation.

Population and evolutionary dynamics of Shiga-toxin Producing *Escherichia coli* O157 in an Australian beef herd

Meghan Jones[1], Sophie Octavia[1], Geraldine Lammers[2], Jane Heller[2], and Ruiting Lan[1,*]

1. School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, New South Wales, Australia;

2. School of Animal and Veterinary Science, Charles Sturt University, Wagga Wagga, New South Wales, Australia.

Shiga toxin producing *Escherichia coli* O157:H7 (STEC O157) is naturally found in the gastrointestinal tract of cattle and can cause severe disease in humans. There is limited understanding of the population dynamics and microevolution of STEC O157 at herd level. In this study, isolates from a closed beef herd of 23 cows were used to examine the population turnover in the herd. Of the nine STEC O157 clades previously described, clade 7 was found in 162 of the 169 isolates typed. Multiple locus variable number tandem repeat analysis (MLVA) analysis differentiated 169 isolates into 33 unique MLVA types. Five predominant MLVA types were evident with most of the remaining types containing only a single isolate. MLVA data suggest that over time clonal replacement occurred within the herd. Genome sequencing of 18 selected isolates found that the isolates were divided into four lineages, representing four different 'clones' in the herd. Genome data confirmed clonal replacement over time and provided evidence of cross transmission of strains between cows. The findings enhanced our understanding of the population dynamics of STEC O157 in its natural host that will help developing effective control measures to prevent the spread of the pathogen to the human population.

Comparative genomic analysis and physiological characterization of the emerging pathogen *Elizabethkingia anophelis*

YINGYING LI1*, LIANG YANG1,2

1. Singapore Centre for Environmental Life Sciences Engineering (SCELSE), Nanyang Technological University, Singapore 637551

2. School of Biological Sciences, Nanyang Technological University, Singapore

Elizabethkingia anophelis is an emerging Gram-negative pathogen that causes infections with high mortality rate in neonates and in immunocompromised patients. Characterization of the pathogenesis mechanisms of *E. anophelis* is limited due to the lack of comprehensive genomic information. To this end, we sequenced the complete genome of a clinical isolate *E. anophelis* strain NUHP1 and revealed the potential genes responsible for pathogenesis mechanisms such as iron uptake and stress response systems. Meanwhile, we compared the genomes between NUHP1 and assemblies of *E. anophelis* strains isolated from the mosquito gut, supporting our hypothesis of well-developed systems for scavenging iron and oxidative stress response in *E. anophelis*. Thereafter, transcriptome profiling was conducted for NUHP1, comparing the expression similarity and differences between hydrogen peroxide-treatment with mouse blood treatment. The transcriptomic data further showed that the presence of protective mechanism associated with iron siderophore and heme utilization in NUHP1 when it was treated with mouse blood. Our genomic and transcriptomic analysis

indicated that oxidative stress might be one of the major stresses for NUHP1 upon infection in hosts. Moreover, we sequenced 16 clinical *E. anophelis* isolates by using the Pacbio platform and compared them with 3 available *Elizabethkingia spp.* genomes. The phylogenetic analysis indicates that *E. anophelis* was transmitted between different ICUs in local hospitals.

Microbiological Status of Ready to Eat Food (Pickles, Jhalmuri, Velpuri/Fuska) sold in front of the schools of Dhaka City, Bangladesh

RAHMAN MATIUR* JAMES JACOB RIBERO[1] FERDOUSI SHAHNILA[2]

***Presenting author, Deputy Technical Manager, National Food Safety Laboratory, Institute of Public Health, Mohakhali, Dhaka, Bangladesh** matiurmrahman@yahoo.com

[1] Senior Research Associate, IMRC, Dhaka, Bangladesh

[2] Professor, Dhaka Medical College, Dhaka, Bangladesh

In Bangladesh, street food vendors are available in front of school, university, office, footpaths etc. Of the food items chotpoti, fuska, velpuri, samucha, daalpuri, lassi, pakora, pizza and patties are common. Studies reported that about 50 percent of the street foods tested was contaminated with various diarrhoea causing bacteria, and over 40 percent contained traces of faecal pathogens. Several newspapers reported that increased risk of child health associated with these foods. In this study, three most popular street food items namely velpuri/fuska, jhalmuri & pickles sold in front of schools of Dhaka city (n = 134) were tested to determine microbiological status by: Total Aerobic Viable Colony Count, Detection of coliform, *Escherichia* coli, Enumeration and Serotyping of *Salmonella* and Total Mold Count tests. Analysis was done according to the ISO Method. Among the samples, 85-90% of velpuri, fuska, and jhalmuri were found heavily contaminated with faecal coliform bacteria and other pathogens. Unacceptable levels of mold were found in 75%, 99% and 10% samples of velpuri, fuska and pickles respectively. It was also observed that, the street foods mostly sold in the open, may be further contaminated with physical and chemicals due to traffic pollution caused by different kinds of vehicular exhausts. Strong awareness of personal hygiene and food safety among the vendors and consumers may contribute in reducing the risk of food safety among the school children.

Host adaptations of *Salmonella enterica* subsp. *enterica* serovars deciphered by the first Genome Wide Association Study implementing accessory genes and coregenome variants

VILA NOVA Meryl1,2,*; LA Kevin1; FELTEN Arnaud1; DURIMEL Kevin1; BESSIERES Philippe2; MISTOU Michel-Yves1; MARIADASSOU Mahendra2 and RADOMSKI Nicolas1
*** Presenter**

1 Université PARIS-EST, Anses, Laboratory for food safety, Maisons-Alfort, France.

2 Université AgroParisTech, INRA, Applied Mathematics and Computer Science, from Genomes to the Environment, Jouy-en-Josas, France.

Salmonella enterica subsp. *enterica* is a major public health concern as a foodborne pathogen, and its adaptation mechanisms to specific animal hosts remain poorly understood. Genome Wide Association Study (GWAS), routinely used in human genetics, have recently been applied successfully on bacteria [1] to reveal the genetic basis of bacterial host speciation [2], antibiotic resistance [3] and virulence [4].

We propose to decipher host adaptations of *Salmonella* serovars based on the first GWAS implementing accessory genes and coregenome variants. We selected host-adapted genomes from a curated large database (i), listed accessory genes and coregenome variants (ii), and associated those genetic mutations to the host adaptations (iii).

In order to create a genome dataset underpinning host speciation, the Enterobase database were downloaded (i.e. December 2016; 70 682 records), curated and synthesized (i.e. 13 532 records; 277 serovars) (i). A set of genomes including 20 strains of 12 multi- and preferential-host serovars (i.e. 440 genomes) was selected. Genes and variants were then detected by pangenome extraction and variants (i.e. SNPs and InDels) calling analysis, respectively (ii). These mutations were associated to host-adaptations by a GWAS that corrects for population structure using a linear mixed model as described in Earle *et al.* 2016 [5].

With this innovative approach, coupling selection of preferential host-adapted genomes underpinned by synthesis of the large Enterobase database, as well as GWAS integrating an advanced correction of structure population, we are able to decipher the host adaptations of *Salmonella* serovars.

The future status of NGS-databases under the Nagoya Protocol: What can we learn from Culture Collections and Biobanks?

CAROLINA RIBEIRO1*, MARION KOOPMANS2, GEORGE HARINGHUIZEN1

1 Center for infectious Disease Control, RIVM (Netherlands National Institute for Public Health and the Environment), Bilthoven, Netherlands.

2 Department of Viroscience, Erasmus MC, Rotterdam, Netherlands.
*** Contact: *carolina.dos.santos.ribeiro@rivm.nl***

In this year of 2018, the Conference of Member States (CoP) of the CBD/Nagoya Protocol discusses the status of Next Generation Sequencing data (NGS-data): do they fall by nature under the scope of the Nagoya Protocol or not? If NGS-data is to be considered as equivalent to tangible genetic resources (physical samples), the status of NGS-databases will be comparable to existing Culture Collections and Biobanks that already need to be compliant to the Protocol. Hence, how biobanks have coped with the Nagoya conditions and who bears the burden? What can we learn from it? We looked for systematic approaches for dealing with the Nagoya Protocol measures, and identified four models: the American, Japanese, Asian, and European. These models differ depending on how transfers of tangible genetic resources are performed for commercial and non-commercial use, if their redistribution is allowed, and who is responsible for the Access and Benefit Sharing negotiations with the country of origin of the resources. All investigated models end up being burdensome and compromising the timely sharing of genetic resources. The applicability of these models to NGS-databases is still questionable due to the intangible nature of NGS-data and the difficulties to track their access, transfer and use. Attention should be placed on the ability of biobanks to provide genetic resources promptly for public health emergencies by creating exemptions for pathogen genetic resources and/or facilitated systems of transfers. This would be possible only through the political engagement of national and international stakeholders daring to collaborate across their specific institutional mandates and domains.

Rapid Identification of Stable Clusters in Bacterial Populations Using the Adjusted Wallace Coefficient.

Dillon OR Barker[1,2], João A Carriço[3], Peter Kruczkiewicz[4], Federica Palma[5], Mirko Rossi[6], Eduardo N Taboada*[1]

[1] Surveillance, Outbreak Detection and Response Section, National Microbiology Laboratory at Lethbridge, Public Health Agency of Canada, Lethbridge, Alberta, Canada

[2] Department of Biological Sciences, University of Lethbridge, Lethbridge, Alberta, Canada

[3] Instituto de Microbiologia, Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal

[4] Bioinformatics Core, National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada

[5] Department of Food and Agricultural Sciences, Alma Mater Studiorum - University of Bologna, Bologna, Italy

[6] Department of Food Hygiene and Environmental Health, Faculty of Veterinary Medicine, University of Helsinki, Helsinki, Finland

Whole-genome sequencing (WGS) of microbial pathogens has become an essential part of modern epidemiological investigations. Although WGS data can be analyzed using a number of different approaches, such as traditional phylogenetic methods, a critical requirement for global systems for pathogen surveillance is the development of approaches for transforming sequence data into WGS-based subtypes, which creates a nomenclature that describes their higher-order relationships to one another. To this end, subtype similarity thresholds are needed to define clusters of subtypes representing lineages of interest. WGS-based subtyping presents a challenge since both the addition of novel genome sequences and small adjustments in similarity thresholds can have a dramatic impact on cluster composition and stability. We present the Neighbourhood Adjusted Wallace Coefficient (nAWC), a method for evaluating cluster stability based on computing cluster concordance between neighbouring similarity thresholds. The nAWC can be used to identify areas in in which distance thresholds produce robust clusters. Using datasets from *Salmonella enterica* and *Campylobacter jejuni*, representing strongly and weakly clonal bacterial species respectively, we show that clusters generated using such thresholds are both stable and reflect basic units in their overall population structure. Our results suggest that the nAWC could be useful for defining robust clusters compatible with nomenclatures for global WGS-based surveillance networks, which require stable clusters to be defined that both harness the discriminatory power of WGS data while allowing for long-term tracking of strains of interest.

Use of NGS to characterize Escherichia coli isolates from clinical cases of colibacillosis in poultry

GHYSLAINE VANIER*, GABRIEL DESMARAIS, JOHN MORRIS FAIRBROTHER

OIE Reference Laboratory for Escherichia coli, Service de Diagnostic, Faculté de médecine vétérinaire, Saint-Hyacinthe, Québec, Canada.

Avian pathogenic Escherichia coli (APEC) causes colibacillosis, which results in morbidity, mortality and economic losses to the poultry industry. Historically, characterization of clinical isolates was done by PCR, slide agglutination and disk diffusion assay. The objective of this work was to characterize APEC isolates using Next Generation-Sequencing (NGS) in order to implement its use in our routine diagnostic microbiology. Fifteen E. coli isolates from clinical cases of colibacillosis in poultry from 2014 to 2017 were characterized by NGS using an Illumina MiSeq following DNA library preparation with Nextera XT library kit. In silico analysis was performed for virulence genes, O:H serotypes, multilocus sequence types, antimicrobial resistance (AMR) genes, and plasmid replicon types. In parallel, these isolates were examined by PCR for the detection of 15 APEC virulence genes, by slide agglutination for the detection of O-serotypes and by disk diffusion assay to determine AMR profile. Our results show 98,0% similarity between NGS and classical PCR for detection of virulence genes. When our NGS results were used to predict AMR, they were in agreement with disk diffusion assay results at 95,2%. The characterization of APEC clinical isolates by NGS will allow rapid and accurate detection and surveillance of the different sero-virotype and AMR gene profiles present in the poultry farms and help to determine their zoonotic potential. Moreover, it will facilitate the tracking of pathogenic and/or multi-drug resistantE. coli isolates between production sites, and the identification of candidate strains for production of vaccines as an aid to control of this disease.