



Next generation microbiological risk assessment: opportunities of whole genome sequencing (WGS) for foodborne pathogen surveillance, source tracking and risk assessment



Kalliopi Rantsiou^{a,*}, Sophia Kathariou^b, Annet Winkler^c, Panos Skandamis^d, Manuel Jimmy Saint-Cyr^g, Katia Rouzeau-Szynalski^e, Alejandro Amézquita^f

^a Department of Agriculture, Forest and Food Sciences, University of Turin, Largo Paolo Braccini 2, 10095 Grugliasco, Italy

^b North Carolina State University, Department of Food, Bioprocessing and Nutrition Sciences, Raleigh, NC, USA

^c Mondelez International, Kraft Foods R&D, Unterbibberger St. 15, D-81737 Munich, Germany

^d Laboratory of Food Quality Control & Hygiene, Department of Food Science & Technology, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece

^e Food Safety Microbiology, Nestec Ltd, Nestlé Research Center, Lausanne, Switzerland

^f Safety and Environmental Assurance Centre, Unilever, Colworth Science Park, Sharnbrook, Bedfordshire MK44 1LQ, UK

^g SECALIM Unit UMR1014, Oniris, INRA, Université Bretagne Loire, Nantes, France

ARTICLE INFO

Keywords:

Whole genome sequencing
Surveillance
Epidemiology
Foodborne pathogens
Source tracking

ABSTRACT

Whole genome sequencing (WGS) of important foodborne pathogens is a technology under development, but is already employed in routine surveillance by public health agencies and is being increasingly exploited in tracing transmission routes and identifying contamination events (source tracking) that take place in the farm-to-fork continuum. Furthermore, data generated from WGS, complemented by other -omics data, have the potential to be integrated into and strengthen microbiological risk assessment. In this paper, we discuss the contribution of WGS in diverse areas important to food safety and public health. Additionally, an outlook of future WGS applications, which should contribute to our understanding of the ecology and physiology of foodborne microorganisms, is presented.

1. Introduction

In the past few years, the intensive use of high-throughput Next Generation Sequencing (NGS) technologies has led to an unprecedented increase in speed and cost-effectiveness of the acquisition of very high volumes of nucleic acid sequence data (Deng et al., 2016). The technology has markedly increased the feasibility and routine implementation of a number of previously challenging applications. One such currently widely used application of NGS is whole genome sequencing (WGS), an analytical approach to determine the sequence of the whole genomic content of an organism (Mardis, 2008; Ståhl and Lundeberg, 2012).

In several industrialized countries, public health agencies and regulatory bodies [e.g. Food and Drug Administration (FDA), Centers for Disease Control and Prevention (CDC), Public Health England, European Center for Disease Prevention and Control (ECDC)] are now using WGS routinely to characterize clinical isolates of selected foodborne pathogens and to support epidemiological investigations. For instance, in the United States the CDC is performing WGS on all isolates

from human listeriosis cases and retrospectively on stored isolates from earlier cases of disease (Jackson et al., 2016). This would ensure that isolates can be compared after different time periods. The CDC has announced the intention to phase out current typing techniques and fully move to WGS by 2018. WGS is also included as a routine tool in the National Antimicrobial Resistance Surveillance System (NARMS). The FDA is routinely sequencing all *L. monocytogenes* isolates and has established the GenomeTrakr Network to facilitate sharing of genomic and other isolate-related information between laboratories. WGS is now used in parallel with other typing techniques in many countries of the European Union, with Denmark having already transitioned completely to WGS for certain foodborne pathogens (ECDC, 2016a). To further promote application of WGS, the ECDC will oversee and coordinate efforts to overcome current challenges and is aiming to establish WGS as the method of choice for foodborne pathogen typing and surveillance within the next five years (ECDC, 2016a).

WGS typically includes several steps: obtaining a pure culture of the organism of interest, DNA extraction, library preparation and amplification, and DNA sequencing, followed by possible alignment, data

* Corresponding author.

E-mail address: kalliopi.rantsiou@unito.it (K. Rantsiou).

analysis, and finally biological interpretation. Today, WGS is not necessarily followed by annotation, data analysis or functional (biological) interpretation. It is generally limited to obtaining and ideally depositing a file with the entire genome sequence in a database such as the National Center for Biotechnology Information (NCBI). Furthermore, in most instances WGS data are in the form of contigs, and in the absence of a closed genome sequence the available data are expected to contain sequence gaps. As WGS evolves from a research tool to a routine surveillance instrument, different pipelines, software and platforms are used to perform sequencing data analysis. Many of the pipelines are operational as research tools and under continuous development; for instance, FDA CFSAN provides periodic updates on the single nucleotide polymorphism Pipeline (<https://github.com/CFSAN-Biostatistics/snp-pipeline>; Davis et al., 2015). An overview of different bioinformatics approaches to WGS data analysis is provided by Franz et al. (2016). Standardization, i.e. validation and articulation of key quality requirements and criteria, is needed to be able to compare results produced by different pipelines, software and platforms.

In addition to epidemiological surveillance and clinical applications, WGS is slowly being explored by the food industry to improve food safety. In the context of food safety management systems (FSMS) and the implementation of measures to prevent foodborne pathogen contamination, food business operators (FBO) routinely monitor the production environment. Such monitoring programs rely on classical microbiological methods but may be complemented by molecular-based methods to obtain a more precise view of the hygienic state of the environment and a much more extended view of the ecological richness. The information collected from such approach may provide insight into contamination routes. Currently, the development of WGS is not at the same level in food industry compared to public health agencies, which are using it routinely. Food industry started to explore WGS for tracking the source of microbial contamination and for other purposes, such as determination of virulence and antibiotic resistance genes for pathogenic strains. Delays in adoption of WGS are mainly due to high costs, time and expertise/resources/capacities requirements. Extensive application by the food industry will reveal the potential of WGS to differentiate between strains persisting in a facility and those repeatedly introduced from an outside source.

The aim of this paper is to present the current status and the opportunities of WGS applications in foodborne pathogen surveillance, in understanding reservoirs, delineating transmission routes and integrating genomic data into risk assessment.

2. Foodborne pathogen surveillance

For the purposes of this paper, foodborne pathogen surveillance is defined as the systematic collection, analysis and interpretation of data essential to the planning, implementation and evaluation of public health practice, and the timely dissemination of this information for public health action (WHO, 2008). The main goals of surveillance programs are to identify, control and prevent foodborne outbreaks, determine the causes of foodborne disease, and monitor trends in occurrence of foodborne disease (Potter et al., 2000). Information gathered during surveillance programs generally contributes to the development of prevention strategies and to the assessment of their effectiveness (Braden and Tauxe, 2006). Moreover, information on the epidemiology of foodborne pathogens is highly valuable for risk assessment and HACCP plan development and implementation (Braden and Tauxe, 2006; Potter et al., 2000).

Although WGS can be used for all foodborne biological agents, it is currently most intensely used for bacteria. WGS information has been used to characterize, subtype or compare isolates of several bacterial species (Frey and Bishop-Lilly, 2015; Stasiewicz et al., 2015). To date, agents most extensively investigated by WGS are the bacteria *Listeria monocytogenes*, *Salmonella enterica* and *Escherichia coli*, partly because of their importance as human foodborne pathogens and the relatively

small size of their genomes (*L. monocytogenes*, approx. 3 Mb; *Salmonella* and *E. coli*, approx. 4.6 Mb each). Numerous WGS investigations have addressed other pathogens from food animals or food production environment. For instance, WGS of *Campylobacter jejuni* and *C. coli* from food animals and farm ecosystem have already yielded insights on potential host-adapted attributes and novel antimicrobial resistance determinants (Chen et al., 2013; Dutta et al., 2016; Miller et al., 2016; Qin et al., 2014; Zhao et al., 2015).

Routine use or complete replacement of other typing techniques by WGS will need adequate attention to overcome existing challenges and limitations such as costs, manpower, laboratory and bioinformatics infrastructure, safeguarding of findings, standardization and validation of the methodology, data interpretation requiring expert knowledge (both from the bioinformatics and the microbiology perspective). The lack of standards and harmonization is one of the challenges to overcome to enable full exploitation of WGS for routine epidemiological uses. On a more global level, WGS will face a different set of constraints including legal aspects of data sharing, international trade, integration of data into food safety decision making, and capabilities to perform WGS (EFSA, 2014; FAO, 2016).

Foodborne pathogen surveillance programs have long been implemented through networks bringing together several laboratories, such as FoodNet, PulseNet, the European Surveillance System (TESSy) (FDA, CDC, van Walle, 2013, ECDC). Until recently, the standard molecular subtyping and fingerprinting methods were pulsed-field gel electrophoresis (PFGE) and multiple-locus variable-number tandem repeat analysis (MLVA). PFGE lacks the resolution to effectively pinpoint the source of an outbreak while MLVA generally demonstrates a higher discriminatory power than PFGE (Best et al., 2007). However, MLVA has a notable limitation in relation to source attribution studies because many MLVA schemes are specific for certain clones only (Nadon et al., 2013). For instance, the MLVA scheme effective for one *Salmonella* serovar is frequently not equally discriminative for others (Ross et al., 2011). WGS has the potential to overcome challenges related to discriminatory power.

The integration of WGS in surveillance programs has already started to significantly contribute to the investigation of foodborne outbreaks involving bacterial agents. In the United States, the FDA and the CDC are intensively using this technology in outbreak investigation by comparing isolates from patients to those from food or food production environments. WGS data are uploaded to the open-access database GenomeTrakr. The GenomeTrakr network is currently comprising of numerous laboratories in the US and elsewhere, leading to a dramatic increase in the number of sequences in the database (Allard et al., 2016). In the European Union, the ECDC is also putting in place a network of public health laboratories employing WGS, with at least 19 countries enrolled in surveillance programs and outbreak investigations. A common European database for WGS data and metadata is under development (ECDC, 2016b).

In the context of outbreak investigations, GenomeTrakr and equivalent databases will make it increasingly possible to accurately establish links between sequences from outbreaks and from food or environmental sources. Furthermore, the high discriminatory power of WGS allows for detection of diffuse outbreaks by linking infrequent cases, which would otherwise have been considered sporadic cases without a common source. Such capacities are expected to help to contain outbreaks due to earlier outbreak detection and intervention at the source(s). Recent examples are given in “Applications of WGS in Food Safety management” (FAO, 2016) and are increasingly encountered in the literature (Butcher et al., 2016; Dallman et al., 2016; Holmes et al., 2015).

The development of such databases combined with epidemiological investigations can enhance the possibility to share sequences among countries and thus identify and monitor global outbreaks (Allard et al., 2016). The expansion of the databases will allow more links between present and past outbreaks or contamination events, thus potentially

identifying strains persistent through time in certain production facilities or other ecosystems. Thus, WGS data may be employed to retrospectively investigate unsolved outbreaks or contamination events. However, this type of analysis can only be supported by complete metadata providing the relevant context. The metadata should include information, such as where strains were isolated, from which kind of sample, country of origin and date of isolation.

It is assumed that all identical or nearly identical strains belonging to the same species have a common ancestry. Finding nearly identical strains at different processing facilities raises questions about the expected genetic variability within pathogens. Studies exploring overall genetic diversity among strains of a certain species are needed, especially for highly clonal microorganisms such as lineages within *L. monocytogenes*, *Salmonella* and others (Leekitcharoenphon et al., 2014). This evidence poses a challenge for the food industry; (nearly)-identical strains can be found in multiple food environments i.e. an isolate of one factory does not necessarily mean a causal link with a clinical isolate and thus an isolate of a human patient. Stasiewicz et al. (2015) showed that *L. monocytogenes* isolated in a retail deli-specific setting differed by very limited Single Nucleotide Polymorphisms (median 2 to 11 SNPs) and have concluded that identical or nearly identical strains of *L. monocytogenes* can occur at different deli environments at different locations, with no evident clear links in transmission. WGS data provided evidence for repeated introduction of *L. monocytogenes* from an external source to multiple retail establishments. Similarly, Fagerlund et al. (2016) reported nearly identical *L. monocytogenes* MLST Sequence Type 8 isolates in Danish and Norwegian salmon processing plants spanning a period of more than 10 years. In another case, nearly identical *L. monocytogenes* ST 8 isolates were found in two poultry processing plants in Norway (Fagerlund et al., 2016). These examples emphasized that linking disease back to isolates from processing environments can be challenging as almost identical isolates can be detected in more than one food processing plant. To resolve situations of uncertainty it is important that metadata are carefully collected and evaluated in combination with the WGS data.

Sharing of WGS data and metadata across countries will allow development of novel surveillance programs and facilitate global perspectives of pathogen trends and emergence of new pathogens or food safety issues. The need to globally monitor important foodborne pathogens was recently highlighted by Moura et al. (2016) who subjected 1696 strains of *L. monocytogenes* to WGS and demonstrated that circulation of isolates between continents/countries takes place. These approaches, can only be fully effective with harmonization of methods and data-sharing platforms (Deng et al., 2016). Specific pilot initiatives are already in place with other pathogens, e.g. molecular surveillance of multidrug-resistant *Mycobacterium tuberculosis* within the European Union (van Walle, 2013). International initiatives such as the Global Microbial Identifier (GMI) as well as tripartite expert groups that involve industry, academia and government (e.g. ILSI Europe Microbiological Food Safety Task Force), are making efforts in this direction, in particular for method standardization and data sharing among public health organizations, academia and industry, areas that currently present substantial challenges.

3. Understanding reservoirs and delineating transmission routes

The WGS data collected from surveillance programs can be used to perform phylogenetic analyses of the population for various foodborne pathogens. Compared to previously employed methods to infer phylogenetic relationships among pathogenic isolates (such as PFGE or MLST), WGS presents the unprecedented characteristics of providing sequence data that span the whole genome and a discriminatory power that can be adapted, based on the downstream data analysis performed, to the genome characteristics of the pathogen in question. Hence, WGS is a data generation tool that can be considered “universal” i.e. can be similarly applied to all pathogens, independently of the level of genome

diversity that their population exhibits. Based on the structure of the population (monophyletic vs. highly diverse), the discrimination level of the downstream data analysis can be attuned (for example SNP based vs. genome wide-MLST). WGS has therefore the potential to gradually replace all pre-existing typing methods, as long as robust standards of harmonization are put in place in all steps of the process, i.e. before, during and after the acquisition of the sequencing data.

The genetic relations inferred from phylogenetic studies can be used in source attribution. In source attribution the goal is to “quantify the relative importance of specific food sources and animal reservoirs for human cases of foodborne illness” (EFSA, 2008). The genetic relations may indicate associations with specific hosts (therefore potentially identifying important reservoirs of pathogens) or habitats (resident populations at specific locations) as well as point to the geographical distribution of a pathogen's subtypes at the local, regional or global level (therefore identifying putative routes of transmission). Several subtyping methods can provide information that may be used to calculate frequencies of isolation of certain subtypes from certain foods (EFSA, 2013). This information can be correlated to frequency of occurrence of the subtypes among clinical isolates to infer the most important food source leading to disease by a certain subtype of the pathogen. For example, multilocus sequence typing (MLST) has been used with success for source attribution and discrimination of *Campylobacter* (Sheppard et al., 2009; Guyard-Nicodème et al., 2015) whereas it appears to be less appropriate for *Salmonella* (Barco et al., 2013).

WGS is a powerful tool to identify transmission pathways (i.e. epidemiological links between reservoirs or sources and infections) and complement epidemiological data (i.e. spatial, temporal, and host characteristics) rather than replace them. This technology is highly suitable for *Salmonella* and several other foodborne pathogens. For *Salmonella* strains, WGS is replacing conventional tools such as PFGE, which often is inadequate in accurately tracking the source of contamination (Allard et al., 2012). This was already shown by Hoffmann et al. (2016) in the context of a *Salmonella* Bareilly outbreak with scraped tuna. WGS determined links between the outbreak isolates in the US and isolates from a fish processing facility in India, thus yielding a geo (spatio-)temporal transmission map for this outbreak. It was also demonstrated that WGS from “historical” archived isolates was useful in establishing links between isolates from 2003 and 2012 (Hoffmann et al., 2016). Furthermore, results were compared with PFGE, and WGS demonstrated a higher resolution power. WGS analysis of *S. Enteritidis* isolates from an egg outbreak in the UK revealed a clear link between human, egg and environmental *S. Enteritidis* isolates specific to the outbreak (Inns et al., 2015). Concerning *Campylobacter*, molecular epidemiology of *C. jejuni* and *C. coli* has remained challenging due to the extensive genomic and phenotypic diversity within these species (Bronnec et al., 2016; Haddad et al., 2010; Rivoal et al., 2005; Rodrigues et al., 2015). The rapid evolution of bacterial genomes has important consequences for interpretation of molecular typing information (Lomonaco et al., 2015). Outbreak isolates may be missed in cases where small genomic changes result in large changes of molecular profiles by conventional tools (Barton et al., 2007). Conversely, some strains appear to be clonal and may be linked by conventional typing methods despite significant differences in gene content among isolates and absence of epidemiological linkage (Taboada et al., 2008).

Apart from clearly defining genetic relationships among subtypes within pathogenic species and permitting associations to be made with specific reservoirs or habitats, WGS data can also be used to obtain basic biological insights that could explain such associations. Moreover, considering both genetic and functional (e.g. virulence, survival and antimicrobial resistance) relationships among isolates will provide added value to the source attribution models (EFSA, 2013). In this context it should be underlined that WGS data can be used to predict functional properties of isolates and it is foreseen that it will be possible to discover biomarker genes that indicate association with a specific reservoir or habitat (addressed also in Cocolin et al., den Besten et al.,

this issue).

4. Contribution to risk assessment: “upgrading” hazard identification

Microbiological Risk Assessment (MRA) is a structured, systematic, science-based approach that evaluates the likelihood of adverse human health effects occurring following exposure to a pathogenic microorganism or its products, e.g. a toxin. MRA is intended to support the understanding and management of microbiological risks, so that it informs the need for possible risk management options and enables evaluation of management practices in the farm-to-fork continuum. MRA consists of four steps (CAC, 2014; US EPA and USDA/FSIS, 2012): (Aanensen et al., 2016) *hazard identification* – knowledge about the microorganism and its association with adverse health effects; (Allard et al., 2012) *hazard characterization* – likelihood of infection given the level of exposure (i.e. dose-response relationship) and the consequences of infection; (Allard et al., 2016) *exposure assessment* – which considers the quantities of microorganisms in raw materials and the impact of preservation and processing systems to arrive at likely quantities in finished products, the product use and hence consumer exposure; and (Barco et al., 2013) *risk characterization* – which provides an estimate of the level of risk to the exposed individuals, allowing for risk management decisions to be made on the basis of risk instead of solely on the potential for the organism to be present.

Whilst the use of MRA in food safety has been established for almost two decades, the application of WGS in MRA is largely unexplored. Many challenges remain, particularly in reconciling the different scales of data necessary to inform a decision based on an estimate of risk to human health – i.e. from tens of thousands of SNPs, to thousands of genes, to tens of biologically relevant phenotypic traits, to a single measure of risk (Franz et al., 2016; Pielaat et al., 2015). Two ‘sister’ papers in this issue discuss the potential application of WGS and other omics technologies in ‘*hazard characterization*’ and ‘*exposure assessment*’. Therefore, this section focuses on its potential application in *hazard identification*.

Phenotypic diversity of foodborne pathogens concerns numerous traits, especially those related to (i) virulence potential based on epidemiological evidence; certain strains of a species may be strongly associated with human disease while others are encountered rarely, or not at all, in clinical cases; (ii) behavior within foods or in the food environment, for example different abilities to resist low pH, low a_w , heat, biocides or to form biofilm and persist in the processing environment; (iii) association with specific habitats or animal hosts (for agents with environmental or animal reservoirs). Such phenotypic traits may allow sorting of strains based on phenotypic differences, but may not readily correspond to population structure of the pathogens. Furthermore, phenotypic observations do not provide insights on the underlying molecular mechanisms responsible for a certain (phenotypic) behavior. On the other hand, understanding the differences among members of the same species, but also within clades or phylogenetic lineages of the same species, and discerning their unique genetic features has food safety relevance, for example in terms of defining pathogenic potential, in addition to providing ecological and evolutionary insights.

The advent of molecular biology has provided tools such as PFGE and multilocus sequence typing (MLST), which allowed phylogenetic studies, albeit of limited resolution (Klemm and Dougan, 2016). WGS, on the other hand, is proving to be a flexible instrument to map genetic relationships among isolates of both monophyletic, highly clonal species/lineages and polyphyletic, highly diverse species, based on post-sequencing analysis (Deng et al., 2010; Gordienko et al., 2013). Understanding of the population structure of a bacterial species or other taxonomic unit provides insights into mechanisms of evolution. Gene or gene function loss and horizontal gene transfer (HGT) are key mechanisms that drive bacterial evolution and can frequently underlie the emergence of either pathogenic or attenuated variants, adaptation to

specific environments or specific animal hosts, tolerance to specific stresses and tissue or organ tropism. WGS and comparative genomics have highlighted the role of mobile genetic elements and HGT in the evolution of major foodborne pathogens such as *S. enterica* and *Shigella* spp. (Foley et al., 2013; The et al., 2016). Phylogenetic analysis based on WGS data will also help in the understanding of the evolution and spread of multidrug resistance phenotypes or antibiotic-resistant clades within pathogenic species (Klemm and Dougan, 2016).

By comparing a large set of genomic data it is possible to associate genomic sequences with specific phenotypic traits and identify the molecular basis of a phenotype (Falush and Bowden, 2006). Such genome-wide association studies (GWAS), validated by other –omics (transcriptomics, proteomics) approaches, can eventually lead to the identification of genomic sequences that may serve as markers or indicators of a specific phenotype (e.g. virulence, tolerance to specific stress, host association, environmental distribution) and can also have a predictive value when a bacterial strain or species is unknown or “emerging”. In the future, a *hazard identification* approach can be assisted by the use of such indicator sequences. Whilst the idea of linking genes to phenotypic traits in bacteriology is not new, the technical feasibility of bacterial GWAS remains a challenge due to the clonal nature of bacterial reproduction and population structure; association tests that do not take this into account may result in non-causal relationships (Earle et al., 2016).

A clear understanding of diversity at the genome level and of the population structure of a given pathogen is critical when “representative” or “reference” strains are subjected to further analysis. This is of primary importance in *hazard identification* and an informed choice regarding the strains to be included or considered will impact the fulfillment of the criteria regarding MRA and its success. The study of Pielaat et al. (2015) presents a first attempt to use GWAS in the context of *hazard identification*, showing promising results. The authors coupled genomic (SNP) data with *in vitro* measurements of adherence of *E. coli* O157:H7 to Caco-2 cells (as a proxy for virulence), and highlighted the importance of considering population structure to establish causality between virulence phenotypes and thus, capacity of causing a disease, and genetic disposition, suggesting the need to use mixed-effects models. This is consistent with the findings of a more recent study (Earle et al., 2016), which reports good performance of GWAS in 26 association studies concerning the resistance of 4 bacterial species to 17 antimicrobials. Specifically, in 25 of the 26 associations the leading candidate was a validated antibiotic resistance gene. Similar efforts are underway in order to associate genotypes of foodborne pathogens with tolerance to common food stresses, such as cold, salt or acid (Hingston et al., 2017).

As the cost of sequencing continues to decrease and public databases for WGS data for foodborne pathogens continue to develop (e.g. GenomeTrakr), the amount of WGS data is dramatically increasing, and is expected to continue to do so in the coming years. The potential use of these ‘big data’ in *hazard identification* will enable risk assessors and risk managers to move from a ‘taxonomic’ understanding of the hazard to one based on genotype-phenotype associations and on a much deeper understanding of virulence at the strain level. An example of effective integration of bacterial population genomics with clinical and epidemiological data was recently presented with *L. monocytogenes*, with certain hypervirulent clones found to be over-represented among human disease cases, while hypovirulent ones were over-represented in foods (Maury et al., 2016). Interestingly, this analysis also revealed that commonly employed reference strains of *L. monocytogenes* actually belonged to hypovirulent clones (Maury et al., 2016). In addition to public WGS databases, further information about relevant factors for *hazard identification* needs to be developed. Especially needed are data to inform gene family profiling, as well as the establishment and agreement on relevant reference genomes for specific pathogens that reflect strains actually involved in human illness.

As a systematic approach to support decision-making, MRA requires

that all stakeholders in the food safety continuum have confidence in the robustness of the data, approaches and scientific principles underpinning the output of the risk assessment. Current paradigms for pathogens, such as *L. monocytogenes* in existing regulatory frameworks, has resulted in policies which are based at differentiation at the species level (i.e. ignoring differences in virulence among strains). Therefore, challenges to fully realize the potential for WGS data in *hazard identification* can be witnessed not only at the technical and scientific level but also at the decision-making level, if decisions are to be based on interpretation of high-resolution genomic data such as virulence variation at the strain level.

5. Other applications of WGS in food safety

WGS has high discriminatory power and will in the future have short turn-around time, and overall lower costs, compared to conventional methods. For example, the determination of multidrug resistance and virulence of a pathogen can be time consuming and costly, as it is generally linked with culture-based analysis of the isolates. WGS will allow detection of genes in a one-step analysis, and consequently determination of resistance and virulence (Allard et al., 2016; Köser et al., 2016; Kwong et al., 2015; Reuter et al., 2014). Chattaway et al. (2016) evaluated the use of WGS for the subtyping of Stx-encoding *E. coli* isolates. The results showed that WGS was able, under a routine surveillance set-up, to provide information of the pathogenic potential of each isolate, enabling the prediction of clinical outcomes and the monitoring of emergence of hypervirulent isolates. In another study (Aanensen et al., 2016), WGS data of invasive *Staphylococcus aureus* was combined with epidemiological and resistance data (based on specific genes content). It was then possible to identify high-risk isolates and map their spread through Europe. Data obtained will help to elucidate and follow evolutionary events leading to the emergence and dissemination of hypervirulent strains or antimicrobial resistance genes.

As previously discussed, genomic information gained from WGS should be interpreted in the context of epidemiological analyses. For other applications, e.g. antimicrobial resistance, virulence and other functional assessments, such as tolerance to environmental stress, WGS data should be validated by functional analyses and implementation of additional technologies, including other omics-based tools such as transcriptomics, proteomics and metabolomics.

In addition to its use in surveillance and outbreak investigation, WGS-based information can also provide other subtyping data of relevance to the target microorganism, such as *in silico* determination of serotype and sequence type (ST) (via *in silico* multi-locus sequence typing [MLST]). In fact, such determinations are not only more rapid and accurate but ultimately more cost-effective via WGS-based analysis (Kwong et al., 2016).

Besides the application of WGS on pure cultures of specific isolates, it is starting also to be used to study microbial populations. Whole genomes of microorganisms can be fully or partially reconstructed from metagenomics data (Cocolin et al. this issue). Metagenomics is a culture-independent application of NGS whereby the community DNA from a sample is sequenced, including DNA from all microbes in the sample. If a reference genome is available and the target microorganism being investigated is present at a certain detection level in the sample, its genome can be assembled, allowing culture-independent detection. Culture-independent detection of foodborne pathogens in food, clinical, veterinary or environmental samples circumvents issues related with the “culturability” of microorganisms, e.g. microorganisms that are slow, difficult or impossible to culture, or that cannot be recovered in conventional selective enrichments due to injury or the presence of competing microbiota. In addition it reduces the time necessary for detection since it bypasses the culture steps.

A further contribution of NGS to the process of understanding the association of pathogenic microorganisms with various environments is envisioned. Through a metagenomic approach, the microbial

communities of a sample can be described in a detail never before achieved. Metagenomic analysis of a sample that is associated (in space or time) with a specific WGS subtype of a pathogen, can shed light to events related with co-presence or co-exclusion, co-evolution, horizontal gene transfer and other microbial inter-relations that are, as of yet, poorly studied in nature. This has, among others, the potential to yield relevant indicators for pathogens in samples from specific habitats. Ultimately, it will be possible to understand (predict) the biotic factors that underlie the microbial propensity for a specific habitat.

6. Conclusions and future perspectives

The potential of WGS as a typing tool to be used in surveillance and outbreak investigation and to contribute in risk assessment is indisputable. Scientists in academia, the regulatory sector and the food industry increasingly recognize WGS as the method of choice for basic research applications and epidemiological investigations. Nevertheless, complete replacement of other typing techniques by WGS will need adequate attention to overcome existing challenges and limitations. These include expenses related to initial investment and running costs, human resources, laboratory and bioinformatics infrastructure but also the integration of WGS data into risk assessment and decision-making processes. A gradual phasing out of existing methodologies is foreseen as technical capacity is building up in, and becoming harmonized among, different countries. In this process, key issues that need to be addressed concern standardization, robustness and validation of the analytical methodology, from sample preparation to data interpretation. Specifically in data interpretation, a multidisciplinary approach is/will be required to match and integrate expertise in bioinformatics and biological/microbiological/epidemiological sciences. Initiatives to bring food industry, authorities from food safety and public health together to achieve common agreed understanding and approach of applications, in particular source tracking analysis, is urgently needed. Besides discussions on the technical aspects of the WGS workflow, the impact of WGS on the daily operation in the food industry needs to be addressed. Considering the worldwide food trade, source tracking needs to be performed at the global level, and authorities from different zones need to be involved in related initiatives.

Finally, the future may also open new doors to high-resolution whole-genome profiling of specific microbes from metagenomics datasets, markedly lowering the need for culture-based pathogen detection and characterization.

Glossary and abbreviations

The glossary and abbreviations list are presented in Supplementary Tables 1 and 2, respectively, and are reproduced in the four joint papers on “Next generation Microbiological Risk Assessment”.

Acknowledgements

The paper is the result of a working group within the ILSI workshop on “Next generation MRA (Microbiological Risk Assessment); integration of Omics data into assessment” that took place in Athens, Greece, May 13–14th 2016. This workshop was organised with the help of ILSI-Europe, IAEP and the ICFMH. This paper forms part of a series of four papers that address the contribution of NGS and omics approaches in MRA.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijfoodmicro.2017.11.007>.

References

- Aanensen, D.M., Feil, E.J., Holden, M.T., Dordel, J., Yeats, C.A., Fedosejev, A., Goater, R., Castillo-Ramírez, S., Corander, J., Colijn, C., Chlebowicz, M.A., Schouls, L., Heck, M., Pluister, G., Ruimy, R., Kahlemeter, G., Åhman, J., Matuschek, E., Friedrich, A.W., Parkhill, J., Bentley, S.D., Spratt, B.G., Grundmann, H., European SRL Working Group, 2016. Whole-genome sequencing for routine pathogen surveillance in public health: a population snapshot of invasive *Staphylococcus aureus* in Europe. *MBio* 7, e00444–16.
- Allard, M.W., Luo, Y., Strain, E., Li, C., Keys, C.E., Son, I., Brown, E.W., 2012. High resolution clustering of *Salmonella enterica* serovar Montevideo strains using a next-generation sequencing approach. *BMC Genomics* 13, 1.
- Allard, M.W., Strain, E., Melka, D., Bunning, K., Musser, S.M., Brown, E.W., Timme, R., 2016. The practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J. Clin. Microbiol.* 54, 1975–1983.
- Barco, L., Barrucci, F., Olsen, L.E., Ricci, A., 2013. *Salmonella* source attribution based on microbial subtyping. *Int. J. Food Microbiol.* 163, 193–203.
- Barton, C., Ng, L.K., Tyler, S.D., Clark, C.G., 2007. Temperate bacteriophages affect pulsed-field gel electrophoresis patterns of *Campylobacter jejuni*. *J. Clin. Microbiol.* 45, 386–391.
- Best, E.L., Lindstedt, B.A., Cook, A., Clifton Hadley, F.A., Threlfall, E.J., Liebana, E., 2007. Multiple-locus variable-number tandem repeat analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium: comparison of isolates from pigs, poultry and cases of human gastroenteritis. *J. Appl. Microbiol.* 103, 565–572.
- Braden, C.R., Tauxe, R.V., 2006. Surveillance for emerging pathogens in the United States. In: Motarjemi, Y., Adams, M. (Eds.), *Emerging Foodborne Pathogens*. CRC Press, Washington, DC.
- Bronnec, V., Haddad, N., Cruveiller, S., Hernould, M., Tresse, O., Zagorec, M., 2016. Draft genome sequence of *Campylobacter jejuni* Bf, an atypical strain able to grow under aerobiosis. *Genome Announc.* 4, e00120–16.
- Butcher, H., Elson, R., Chattaway, M.A., Featherstone, C.A., Willis, C., Jorgensen, F., Dallman, T.J., Jenkins, C., McLaughlin, J., Beck, C.R., Harrison, S., 2016. Whole genome sequencing improved case ascertainment in an outbreak of Shiga toxin-producing *Escherichia coli* O157 associated with raw drinking milk. *Epidemiol. Infect.* 144, 2812–2823.
- CAC (Codex Alimentarius Commission), 2014. Principles and Guidelines for the Conduct of Microbiological Risk Assessment. CAC/GL 30-1999. Food and Agriculture Organization of the United Nations/World Health Organization, Rome/Geneva Available at: http://www.codexalimentarius.org/download/standards/357/CXG_030e_2014.pdf, Accessed date: 11 January 2016.
- Chattaway, M.A., Dallman, T.J., Gentile, A., Wright, M.J., Long, S.E., Ashton, P.M., Perry, N.T., Jenkins, C., 2016. Whole genome sequencing for public health surveillance of shiga toxin-producing *Escherichia coli* other than serogroup O157. *Front. Microbiol.* 7, 258.
- Chen, Y., Mukherjee, S., Hoffmann, M., Kotewicz, M.L., Young, S., Abbott, J., Luo, Y., Davidson, M.K., Allard, M., McDermott, P., Zhao, S., 2013. Whole-genome sequencing of gentamicin-resistant *Campylobacter coli* isolated from U.S. retail meats reveals novel plasmid-mediated aminoglycoside resistance genes. *Antimicrob. Agents Chemother.* 57, 5398–5405.
- Dallman, T., Inns, T., Jombart, T., Ashton, P., Loman, N., Chatt, C., Messelhaeuser, U., Rabsch, W., Simon, T., Nikisins, S., Bernard, H., Hello, S.I., Jourdan da-Silva, N., Kornschöber, C., Mossong, J., Hawkey, P., de Pinna, E., Grant, K., Cleary, P., 2016. Phylogenetic structure of European *Salmonella* Enteritidis outbreak correlates with national and international egg distribution network. *Microb. Genom.* 2, e000070.
- Davis, S., Pettengill, J.B., Luo, Y., Payne, J., Shpuntoff, A., Rand, H., Strain, E., 2015. CFSAN pipeline: an automated method for constructing SNP matrices from next-generation sequence data. *Peer J. Comput. Sci.* 1, e20. <https://doi.org/10.7717/peerj.cs.20>.
- Deng, Z., Phillippy, A.M., Li, Z., Salzberg, S.L., Zhang, W., 2010. Probing the pan-genome of *Listeria monocytogenes*: new insights into intraspecific niche expansion and genomic diversification. *BMC Genomics* 11, 500.
- Deng, X., den Bakker, H.C., Hendriksen, R.S., 2016. Genomic epidemiology: whole-genome-sequencing-powered surveillance and outbreak investigation of foodborne bacterial pathogens. *Ann. Rev. Food. Sci. Technol.* 7, 353–374.
- Dutta, V., Altermann, E., Olson, J., Wray, G.A., Siletzky, R.M., Kathariou, S., 2016. Whole-genome sequences of agricultural, host-associated *Campylobacter coli* and *Campylobacter jejuni* strains. *Genome Announc.* 4 (4) (pii: e00833-16).
- Earle, S.G., Wo, C.-H., Charlesworth, J., Stoesser, N., Gordon, N.C., Walker, T.M., Spencer, C.C.A., Iqbal, Z., Clifton, D.A., Hopkins, K.L., Woodford, N., Smith, E.G., Ismail, N., Llewellyn, M.J., Peto, T.E., Crook, D.W., McVean, G., Walker, A.S., Wilson, D.J., 2016. Identifying lineage effects when controlling for population structure improves power in bacterial association studies. *Nat. Microbiol.* <http://dx.doi.org/10.1038/nmicrobiol.2016.41>.
- European Centre for Disease Prevention and Control, 2016a. ECDC roadmap for integration of molecular and genomic typing into European-level surveillance and epidemic preparedness – Version 2.1, 2016–19. ECDC, Stockholm <http://ecdc.europa.eu/en/publications/Publications/molecular-typing-EU-surveillance-epidemic-preparedness-2016-19-roadmap.pdf>.
- European Centre for Disease Prevention and Control, 2016b. Expert Opinion on Whole Genome Sequencing for Public Health Surveillance. ECDC, Stockholm. <http://ecdc.europa.eu/en/publications/Publications/whole-genome-sequencing-for-public-health-surveillance.pdf>.
- European Food Safety Authority BIOHAZ Panel, 2008. Scientific opinion of the panel on biological hazards on a request from EFSA on overview of methods for source attribution for human illness from food-borne microbiological hazards. *EFSA J.* 764, 1–43.
- European Food Safety Authority BIOHAZ Panel, 2013. Scientific opinion on the evaluation of molecular typing methods for major food-borne microbiological hazards and their use for attribution modeling, outbreak investigation and scanning surveillance: part 1 (evaluation of methods and applications). *EFSA J.* 11, 3502. <http://dx.doi.org/10.2903/j.efsa.2013.3502>.
- European Food Safety Authority Scientific Colloquium Summary Report, 2014 European Food Safety Authority Scientific Colloquium Summary. Use of Whole Genome Sequencing (WGS) of Foodborne Pathogens for Public Health Protection. <http://dx.doi.org/10.2805/66246>.
- Fagerlund, A., Langsrud, S., Schirmer, B.C.T., Møretro, T., Heir, E., 2016. Genome analysis of *Listeria monocytogenes* sequence type 8 strains persisting in salmon and poultry processing environments and comparison with related strains. *PLoS One.* <http://dx.doi.org/10.1371/journal.pone.0151117>.
- Falush, D., Bowden, R., 2006. Genome-wide association mapping in bacteria? *Trends Microbiol.* 14, 353–355.
- Foley, S.L., Johnson, T.J., Ricke, S.C., Nayak, R., Danzeisen, J., 2013. *Salmonella* pathogenicity and host adaptation in chicken-associated serovars. *Microbiol. Mol. Biol. Rev.* 77, 582–607.
- Food and Agriculture Organization of the United Nations (FAO), 2016 Food and Agriculture Organization of the United Nations (FA). Applications of WGS in food safety management, Rome. <http://www.fao.org/3/a-i5619e.pdf>.
- Franz, E., Mughini Gras, L., Dallman, T., 2016. Significance of whole genome sequencing for surveillance, source attribution and microbial risk assessment of foodborne pathogens. *Curr. Opin. Food Sci.* 8, 74–79.
- Frey, K.G., Bishop-Lilly, K.A., 2015. Next-generation sequencing for pathogen detection and identification. *Methods Microbiol.* 42, 525–554.
- Gordienko, E.N., Kazanov, M.D., Gelfand, M.S., 2013. Evolution of pan-genomes of *Escherichia coli*, *Shigella* spp. and *Salmonella enterica*. *J. Bacteriol.* 195, 2786–2792.
- Guyard-Nicodème, M., Rivoal, K., Houard, E., Rose, V., Quesne, S., Mourand, G., Rouxel, S., Kempf, I., Guillier, L., Gauschard, F., Chemaly, M., 2015. Prevalence and characterization of *Campylobacter jejuni* from chicken meat sold in French retail outlets. *Int. J. Food Microbiol.* 203, 8–14.
- Haddad, N., Maillart, G., Garénaux, A., Jugiau, F., Federighi, M., Cappelletti, J.M., 2010. Adhesion ability of *Campylobacter jejuni* to HT-29 cells increases with the augmentation of oxidant agent concentration. *Curr. Microbiol.* 61, 500–505.
- Hington, P., Chen, J., Dhillon, B.K., Laing, C., Bertelli, C., Gannon, V., Tasara, T., Allen, K., Brinkman, F.S.L., Hansen, T.L., Wang, S., 2017. Genotypes associated with *Listeria monocytogenes* isolates displaying impaired or enhanced tolerances to cold, salt, acid, or desiccation stress. *Front. Microbiol.* 8, 369. <http://dx.doi.org/10.3389/fmicb.2017.00369>.
- Hoffmann, M., Luo, Y., Monday, S.R., Gonzalez-Escalona, N., Ottesen, A.R., Muruvanda, T., Senturk, I.F., 2016. Tracing origins of the *Salmonella* Bareilly strain causing a food-borne outbreak in the United States. *J. Infect. Dis.* 213, 502–508.
- Holmes, A., Allison, L., Ward, M., Dallman, T.J., Clark, R., Fawkes, A., Murphy, L., Hanson, M., 2015. Utility of whole-genome sequencing of *Escherichia coli* O157 for outbreak detection and epidemiological surveillance. *J. Clin. Microbiol.* 53, 3565–3573.
- Inns, T., Lane, C., Peters, T., Dallman, T., Chatt, C., McFarland, N., Crook, P., Bishop, T., Edge, J., Hawker, J., Elson, R., Neal, K., Adak, G.K., Cleary, P., 2015. Outbreak control team, 2015. A multi-country *Salmonella* Enteritidis phage type 14b outbreak associated with eggs from a German producer: ‘near real-time’ application of whole genome sequencing and food chain investigations, United Kingdom, May to September 2014. *Euro Surveill.* 20, 21098.
- Jackson, B.R., Tarr, C., Strain, E., Jackson, K.A., Conrad, A., Carleton, H., Katz, L.S., Stroika, S., Gould, L.H., Mody, R.K., Silk, B.J., Beal, J., Chen, Y., Timme, R., Doyle, M., Fields, A., Wise, M., Tillman, G., Defibaugh-Chavez, S., Kucerova, Z., Sabol, A., Roache, K., Trees, E., Simmons, M., Wasilenko, J., Kubota, K., Pousele, H., Klimke, W., Besser, J., Brown, E., Allard, M., Gerner-Smidt, P., 2016. Implementation of nationwide real-time whole genome sequencing to enhance listeriosis outbreak detection and investigation. *Clin. Infect. Dis.* 63, 380–386.
- Klemm, E., Dougan, G., 2016. Advances in understanding bacterial pathogenesis gained from whole-genome sequencing and phylogenetics. *Cell Host Microbe* 19, 599–610.
- Köser, C.U., Ellington, M.J., Peacock, S.J., 2016. Whole-genome sequencing to control antimicrobial resistance. *Trends Genet.* 30, 401–407.
- Kwong, J.C., McCallum, N., Sintchenko, V., Howden, B.P., 2015. Whole genome sequencing in clinical and public health microbiology. *Pathology* 47, 199–210.
- Kwong, J.C., Mercoullia, K., Tomita, T., Easton, M., Li, H.Y., Bulach, D.M., Stinear, T.P., Seemann, T., Howden, B.P., 2016. Prospective whole-genome sequencing enhances national surveillance of *Listeria monocytogenes*. *J. Clin. Microbiol.* 54, 333–342. <http://dx.doi.org/10.1128/JCM.02344-15>. (Epub 2015 Nov 25).
- Leekitchaerophon, P., Nielsen, E.M., Kaas, R.S., Lund, O., Aarestrup, F.M., 2014. Evaluation of whole genome sequencing for outbreak detection of *Salmonella enterica*. *PLoS One* 9, e87991. <http://dx.doi.org/10.1371/journal.pone.0087991>.
- Lomonaco, S., Nucera, D., Filipello, V., 2015. The evolution and epidemiology of *Listeria monocytogenes* in Europe and the United States. *Infect. Genet. Evol.* 35, 172–183. <http://dx.doi.org/10.1016/j.meegid.2015.08.008>.
- Mardis, E.R., 2008. Next-generation DNA sequencing methods. *Ann. Rev. Genom. Human Genet.* 9, 387–402.
- Maury, M.M., Tsai, Y.-H., Charlier, C., Touchon, M., Chenal-Francois, V., Leclercq, A., Crisculo, A., Gaultier, C., Roussel, S., Brisabois, A., Disson, O., Rocha, E.P.C., Brisse, S., Lecuit, M., 2016. Uncovering *Listeria monocytogenes* hypervirulence by harnessing biodiversity. *Nat. Genet.* 48, 308–313.
- Miller, W.G., Huynh, S., Parker, C.T., Niedermeyer, J.A., Kathariou, S., 2016. Complete genome sequences of multidrug-resistant *Campylobacter jejuni* strain 14980A (turkey feces) and *Campylobacter coli* strain 14983A (housefly from a turkey farm), harboring

- a novel gentamicin resistance mobile element. *Genome Announc.* 4 (pii: e01175-16).
- Moura, A., Criscuolo, A., Pouseele, H., Maury, M.M., Leclercq, A., Tarr, C., Björkman, J.T., Dallman, T., Reimer, A., Enouf, V., Larssonneur, E., Carleton, H., Bracq-Dieye, H., Katz, L.S., Jones, L., Touchon, M., Tourdjman, M., Walker, M., Stroika, S., Cantinelli, T., Chenal-Francisque, V., Kucerova, Z., Rocha, E.P.C., Nadon, C., Grant, K., Nielsen, E.M., Pot, B., Gerner-Smidt, P., Lecuit, M., Sylvain Brisse, S., 2016. Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat Microbiol* 2, 16185. <http://dx.doi.org/10.1038/nmicrobiol.2016.185>.
- Nadon, C.A., Trees, E., Ng, L.K., Møller Nielsen, E., Reimer, A., Maxwell, N., Kubota, K.A., Gerner-Smidt, P., the MLVA Harmonization Working Group, 2013. Development and application of MLVA methods as a tool for inter-laboratory surveillance. *Euro Surveill.* 18 (pii = 20565).
- NARMS available at: <http://www.fda.gov/AnimalVeterinary/NewsEvents/CVMUpdates/ucm498038.htm>, Accessed date: 28 November 2016.
- Pielaat, A., Boer, M.P., Wijnands, L.M., van Hoek, A.H.A.M., Bouw, E., Barker, G.C., Teunis, P.F.M., Aarts, H.J.M., Franz, E., 2015. First step in using molecular data for microbial food safety risk assessment; hazard identification for of *Escherichia coli* O157:H7 by coupling genomic data with in vitro adherence to human epithelial cells. *Int. J. Food Microbiol.* 213, 130–138.
- Potter, M., Archer, D., Benson, A., Busta, F., Dickson, J., Doyle, M., Farber, J., Finlay, B.B., Goldblatt, M., Hedberg, C., Hooper, D., Jahncke, M., Jaykus, L.-A., Kaspar, C., Liang, A.P., Lindsay, J., Pestka, J., Pierson, M., Slade, P., Tompkin, R.B., Tortorello, M.L., 2000. Emerging Microbiological Food Safety Issues: Implications for Control in the 21st Century. Institute of Food Technology, Chicago.
- Qin, S., Wang, Y., Zhang, Q., Zhang, M., Deng, F., Shen, Z., Wu, C., Wang, S., Zhang, J., Shen, J., 2014. Report of ribosomal RNA methylase gene erm (B) in multidrug-resistant *Campylobacter coli*. *J. Antimicrob. Chemother.* 69, 964–968.
- Reuter, S., Ellington, M.J., Cartwright, E.J.P., Köser, C.U., Török, M.E., Gouliouris, T., Harris, S.R., Brown, N.M., Holden, M.T.G., Quail, M., Parkhill, J., Smith, G.P., Bentley, S.D., Peacock, S.J., 2014. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern. Med.* 173, 1397–1404.
- Rivoal, K., Ragimbeau, C., Salvat, G., Colin, P., Ermel, G., 2005. Genomic diversity of *Campylobacter coli* and *Campylobacter jejuni* isolates recovered from free-range broiler farms and comparison with isolates of various origins. *Appl. Environ. Microbiol.* 71, 6216–6227.
- Rodrigues, R.C., Pocheron, A.L., Hernould, M., Haddad, N., Tresse, O., Cappelier, J.M., 2015. Description of *Campylobacter jejuni* Bf, an atypical aero-tolerant strain. *Gut. Pathog.* 7, 1.
- Ross, I.L., Davos, D.E., Mwanri, L., Raupach, J., Heuzenroeder, M.W., 2011. MLVA and phage typing as complementary tools in the epidemiological investigation of *Salmonella enterica* serovar Typhimurium clusters. *Curr. Microbiol.* 62, 1034–1038.
- Sheppard, S.K., Dallas, J.F., Strachan, N.J., MacRae, M., McCarthy, N.D., Wilson, D.J., Forbes, K.J., 2009. *Campylobacter* genotyping to determine the source of human infection. *Clin. Infect. Dis.* 48, 1072–1078.
- Ståhl, P.L., Lundeberg, J., 2012. Towards the single-hour high-quality genome. *Annu. Rev. Biochem.* 81, 359–378.
- Stasiewicz, M.J., Oliver, H.F., Wiedmann, M., den Bakker, H.C., 2015. Whole-genome sequencing allows for improved identification of persistent *Listeria monocytogenes* in food-associated environments. *Appl. Environ. Microbiol.* 81, 6024–6037.
- Taboada, E.N., MacKinnon, J.M., Luebbert, C.C., Gannon, V.P., Nash, J.H., Rahn, K., 2008. Comparative genomic assessment of multi-locus sequence typing: rapid accumulation of genomic heterogeneity among clonal isolates of *Campylobacter jejuni*. *BMC Evol. Biol.* 8, 229.
- The, H.C., Thanh, D.P., Holt, K.E., Thomson, N.R., Baker, S., 2016. The genomic signatures of *Shigella* evolution, adaptation and geographical spread. *Nat. Rev.* 14, 235–250.
- US EPA and USDA/FSIS (Environmental Protection Agency and Department of Agriculture - Food Safety and Inspection Service), 2012 US EPA and USDA/FSIS (Environmental Protection Agency and D. Microbial risk assessment guideline: pathogenic microorganisms with focus on food and water. Available at: <http://www2.epa.gov/osa/microbial-risk-assessment-guideline-pathogenic-microorganisms-focus-food-and-water>, Accessed date: 1 November 2016.
- van Walle, I., 2013. ECDC starts pilot phase for collection of molecular typing data. *Euro Surveill.* 18 (pii 20357).
- World Health Organisation (WHO), 2008. Foodborne disease outbreaks: guidelines for investigation and control. http://www.who.int/foodsafety/publications/foodborne_disease/outbreak_guidelines.pdf.
- Zhao, S., Tyson, G.H., Chen, Y., Li, C., Murkerjee, S., Young, S., Lam, C., Folster, J.P., Whichard, J.M., McDermott, P.F., 2015. Whole-genome sequencing analysis accurately predicts antimicrobial resistance phenotypes in *Campylobacter* spp. *Appl. Environ. Microbiol.* 82, 459–466.