Advances in foodborne outbreak investigation and source tracking using whole genome sequencing

W Ruppitsch¹, A Pietzka¹, A Cabal¹, A Chakeri¹, D Schmid¹, B Lakicevic², S Lepuschitz¹ and F Allerberger¹

¹ Public Health Division, Austrian Agency for Health and Food Safety, Währingerstrasse 25a, 1090 Vienna, Austria
² Institute of Meat Hygiene and Technology, Kacanskog 13, Belgrade, Serbia

E-mail: werner.ruppitsch@ages.at

Abstract. The progress in sequencing technology has revolutionized the fields of public health and food microbiology. Today, whole genome sequencing allows high-throughput analysis of entire bacterial genomes at affordable costs. Whole genome sequencing has become a daily routine process for surveillance of foodborne infectious diseases, outbreak investigation, and pathogen source tracking. Several studies on a variety of bacterial species have shown that whole genome sequence-based typing approaches are currently the most powerful typing tools. Whole genome sequencing allows the extraction of information on phylogenetic relatedness, antibiotic resistance, virulence-traits, serotype and multilocus sequence type of an isolate from a single analysis. The optimal typing resolution achievable by whole genome sequencing makes it possible to monitor even small genetic variations occurring in an outbreak strain during the course of an outbreak, making transmission events traceable. Whole genome sequencing allowed the creation of global databases based on standardized nomenclatures like the current multilocus sequence type databases. The benefit of global databases is the international exchange of data as a prerequisite for cross border outbreak investigation, strain tracking, and source identification in the global food chain. With further technological advancement, metagenomic approaches may provide future solutions, allowing complete pathogen detection and characterization directly from specimens.

1. Introduction

With increasing industrialization of food production and the international trading of fresh and frozen food, food safety is today a global challenge [1]. The most important aim of food analysis is to ensure the safety and quality of food and to protect consumer from food adulteration [2]. Food associated health problems caused by food contamination or malnutrition have a worldwide impact on public health and economics. Adaptation of microorganisms, climate change, modifications of the human life style and demographics, changes in economic development, excessive land use and increasing environmental pollution can cause the emergence of new microbial threats [3]. Food contamination with pathogenic microorganisms and toxins through soil, water and air, and livestock that was fed with contaminated feed, is a major global health threat [4]. Surveillance is an indispensable requirement for early detection of microbial threats, which allows the timely implementation of appropriate measures to terminate outbreaks, and prevent further transmission and morbidity. In our today’s global world, where pathogens easily cross national borders, disease monitoring requires efficient local, national and international surveillance systems [3]. The rapid progress in new high-throughput technologies like genomics, transcriptomics, proteomics, and metabolomics has defined the term “foodomics” as a new discipline and increased the standards for food safety [5, 6] (Figure 1). These foodomics tools represent new standards used in foodborne outbreak investigation, in food analysis and in food monitoring from harvesting, processing, transport and storage to final consumption [2].
2. Whole genome sequencing applications in food microbiology

For detection of pathogenic microorganisms, and for foodborne outbreak investigation and identification of the source of infection, the rapid progress in sequencing technology from Sanger sequencing to whole genome sequencing (WGS) or next generation sequencing (NGS) is revolutionizing the fields of public health and food microbiology [7]. The superiority of WGS-based strain characterization has led to the replacement of former gold standard typing tools like for example fluorescent amplified fragment length polymorphism (fAFLP), pulsed-field gel electrophoresis (PFGE), multiple-locus variable number tandem repeat analysis (MLVA) and serotyping. WGS-based typing based on either single nucleotide variants (SNVs), on gene-by-gene allelic profiling using core genome multilocus sequence typing (cgMLST) or whole genome multilocus sequence typing (wgMLST) is currently the most powerful diagnostic typing tool. The general benefits of WGS-based strain characterization approaches compared to traditional methods are robustness and superior discriminatory power, and the possibility to infer the geographic origin and to obtain evolutionary information for outbreak isolates [8]. The significant decrease of WGS costs allows, nowadays, the broad use of these technologies in daily routine applications in public health and food agencies. In addition, for public health and food laboratories, the high data quality, the reproducibility and accuracy of WGS technology has been demonstrated [9]. For backward compatibility to datasets obtained with traditional methods, information on serotype, classical multilocus sequence type (MLST) or MLVA data can be extracted from WGS data [10]. For public health and food agencies, it is important that both SNP- (single nucleotide polymorphism) and MLST-based approaches yield concordant results for phylogenetic clustering and complement each other [11-13], which allow, in the case of foodborne outbreak investigation, the responsible outbreak source to be identified with a high level of confidence independent of the used analysis pipeline [11]. The resulting information is the basis for correct decisions required in outbreak situations to stop further transmission and to terminate the outbreak [14]. The setup of open accessible databases allows data sharing between public health and food laboratories worldwide and facilitates international source tracking and multinational outbreak investigation [15].

Therefore, the application of WGS in combination with epidemiological analysis provided a new level on the investigation of foodborne pathogens involving *Cronobacter sakazakii* [16], *Listeria monocytogenes* [15, 17-22], *Salmonella* [23-24], Shiga-toxin producing *Escherichia coli* [8, 25-27],
and *Yersinia enterocolitica* [28]. WGS allows the efficient tracking of pathogens entry routes and distribution from farm-to-consumer. Subpopulations of bacterial pathogens can be transmitted from environmental sources outside processing facilities (animals, incoming raw materials, soil, dust and water) into food processing environments. Bacteria can persist in biofilms on stainless steel surfaces, equipment, floors and cold storage areas over long periods. From production facilities they can spread *via* aerosols, personnel, food workflows, and contaminated contact materials to food, and finally, to the consumer [19, 29]. WGS allows the characterization of these subpopulations at each step – from the product, the environment and clinical samples. Since farms and suppliers can have more than one customer, genetically identical bacteria can spread to multiple consumers, distributors or food processing facilities [20, 29].

The high resolution achieved by WGS-based typing applications allows faster identification and more successful investigation of outbreaks when epidemiological information is unavailable [21-22]. Expanding the use of WGS-based typing analysis globally will ensure the rapid implementation of interventions to protect public health, inform risk assessment and facilitate the management of national and international foodborne outbreaks.

WGS-based strain characterization is an already established process used in microbiological laboratories in daily routine diagnostics for strain characterization, surveillance, outbreak investigation, and source tracking [14, 16, 19-20, 30]. However, the use of metagenomics, an umbrella term that is generally used for 16S rRNA amplicon sequencing and shotgun metagenome or whole metagenome sequencing, is significantly more challenging and still needs improvements before it can be used as an accurate routine diagnostic tool [31-32]. A critical step in the analysis of metagenomics data are sequence assembly algorithms that must be able to reconstruct genes and organisms from complex mixtures. Therefore, due to the complexity of the entire gene mixture present in a sample, the assembly of metagenomics data is a major bioinformatic challenge [33-34]. However, despite this drawback, the benefit of metagenome sequencing, compared to current microbiological methods is the culture-independent analysis, providing nearly unbiased information on microbial communities from either food materials or the production environment [35-36]. In addition, whole metagenome sequencing allows the extraction of further information relevant to food safety like the occurrence of antimicrobial resistance genes, virulence genes and toxin genes and, when linked to transcriptomics or proteomics data, functional capacities and biochemical activities of microbial populations can be identified (Figure 1). The future improvement of WGS techniques together with the development of new and more accurate bioinformatics tools and pipelines will facilitate the application of WGS technologies on a daily routine basis.

In conclusion, the use of WGS provides several advantages of superior discriminatory power for strain characterization, robustness and stability, which is decisive in cluster detection, backtracking the source and reservoir of the causative strain, when epidemiological information is scare, and for gaining knowledge on the evolution of emerging pathogenic strains. WGS technologies provide not only benefits for public health and food agencies but also for the food industry throughout the farm-to-fork principle and upcoming improvements in technology and bioinformatics with the perspective of metagenomic sequencing applied directly to the sample specimen.

References
[1] Allerberger F, Zoltan B, Huhulescu S and Pietzka A A 2015 Listeriosis: The dark side of refrigeration and ensiling In: Sing, A. (Ed.): Zoonoses - Infections affecting humans and animals – Focus on public health aspects Springer Verlag Heidelberg pp 249–86
[2] Andjelković U, Šrajer Gajdošik M, Gašo-Sokač D, Martinović T and Josić D 2017 Foodomics and Food Safety: Where We Are Food Technol. Biotechnol. 55 290–307
[3] Ruppitsch W 2016 Molecular typing of bacteria for epidemiological surveillance and outbreak investigation *Die Bodenkultur: Journal of Land Management, Food and Environment* 67 199–224 DOI: 10.1515/boku-2016-0017
[4] Giacometti J and Josić D 2013 Foodomics in microbial safety *Trends Analys. Chem.* 52 16–22
Herrero M, Simó C, García-Cañas V, Ibáñez E and Cifuentes A 2012 Foodomics: MS-based strategies in modern food science and nutrition Mass. Spectrom. Rev. 31 49–69 doi:10.1002/mas.20335

Jagadeesan B, Gerner-Smidt P, Allard M W, Leuillet S, Winkler A, Xiao Y, Chaffron S, Van Der Vossen J, Tang S, Katase M, McClure P, Kimura B, Ching Chai L, Chapman J and Grant K 2018 The use of next generation sequencing for improving food safety: Translation into practice Food Microbiol. 79 96–115

Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, Gilpin B, Smith AM, Man Kam K, Perez E, Trees E, Kubota K, Takken J, Nielsen EM, Carleton H and FWD-NEXT Expert Panel 2017 PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Euro Surveill. 22 30544 doi: 10.2807/1560-7917.ES.2017.22.23.30544

Jenkins C, Dallman T J and Grant K A 2019 Impact of whole genome sequencing on the investigation of food-borne outbreaks of Shiga toxin-producing Escherichia coli serogroup O157:H7, England, 2013 to 2017 Euro Surveill. 24 1800346 doi: 10.2807/1560-7917.ES.2019.24.4.1800346

Hyden P, Pietzka A, Lenkh A, Murer A, Springer B, Blaschitz M, Indra A, Huhulescu S, Allerberger F, Ruppitsch W and Sensen C W 2016 Whole genome sequence-based serogrouping of Listeria monocytogenes isolates J. Biotechnol. 235 181–86

Brown E, Dessai U, McGarry S and Gerner-Smidt P 2019 Use of Whole-Genome Sequencing for Food Safety and Public Health in the United States Foodborne Pathog. Dis. doi: 10.1089/fpd.2019.2662

Jagadeesan B, Baert L, Wiedmann M and Orsi R H 2019 Comparative Analysis of Tools and Approaches for Source Tracking Listeria monocytogenes in a Food Facility Using Whole-Genome Sequence Data Front. Microbiol. 10 947 doi: 10.3389/fmicb.2019.00947

Hyden P, Grim C, Pietzka A, Blaschitz M, Indra A, Sensen CW, Allerberger, F, Rattei T and Ruppitsch W 2016 Comparison of SNP based and cgMLST based typing of Listeria monocytogenes isolates from Seeliger’s historical “Special Listeria Culture Collection”. Poster; Boston, USA; ASM Microbe

Gilchrist C A, Turner S D, Riley M F, Petri W A Jr and Hewlett E L 2015 Whole-genome sequencing in outbreak analysis Clin. Microbiol. Rev. 28 541–63

Pietzka A, Allerberger F, Murer A, Lenkh A, Stöger A, Cabal Rosel A, Huhulescu S, Maritschnik S, Springer B, Lepuschitz S, Ruppitsch W and Schmid D 2019 Whole Genome Sequencing Based Surveillance of L. monocytogenes for Early Detection and Investigations of Listeriosis Outbreaks Front. Public Health. 7 139 doi: 10.3389/fpubh.2019.00139

Lepuschitz S, Ruppitsch W, Pekard-Amenitsch S, Forsythe S J, Cormican M, Mach R L, Piérard D, Allerberger F and EUCRONI Study Group 2019 Multicenter Study of Cronobacter sakazakii Infections in Humans, Europe, 2017 Emerg. Infect. Dis. 25 515–22 doi:10.3201/eid2503.181652

Ruppitsch W, Prager R, Halbedel S, Hyden P, Pietzka A, Huhulescu S, Lohr D, Schönberger K, Aichinger E, Hauri A, Stark K, Vygen S, Tietze E, Allerberger F and Wilking H 2015 Ongoing outbreak of invasive listeriosis, Germany, 2012 to 2015 Euro Surveill 20:pii=30094

Gelbíčová T, Zobaničková M, Tomášťková Z, Van Walle I, Ruppitsch W and Karpišková R 2018 An outbreak of listeriosis linked to turkey meat products in the Czech Republic, 2012-2016 Epidemiol. Infect. 146 1407–12

Elson R, Awofisayo-Okuyelu A, Greener T, Swift C, Painset A, Amar C F L, Newton A, Aird
H, Swindlehurst M, Elviss N, Foster K, Dallman TJ, Ruggles R and Grant K 2019 Utility of Whole Genome Sequencing To Describe the Persistence and Evolution of *Listeria monocytogenes* Strains within Crabmeat Processing Environments Linked to Two Outbreaks of Listeriosis *J. Food Protect.* 82 30–38

[20] Zuber I, Lakicevic B, Pietzka A, Milanov D, Djordjevic V, Karabasil N, Teodorovic V, Ruppitsch W and Dimitrijevic M 2019 Molecular characterization of *Listeria monocytogenes* isolates from a small-scale meat processor in Montenegro 2011-2014 *Food Microbiol.* 79 116–22 doi: 10.1016/j.fm.2018.12.005

[21] EFSA (European Food Safety Authority) and ECDC (European Centre for Disease Prevention and Control), 2018. Multi-country outbreak of *Listeria monocytogenes* serogroup IVb, multilocus sequence type 6, infections linked to frozen corn and possibly to other frozen vegetables – first update. EFSA supporting publication 2018:EN-1448. 19 pp. doi:10.2903/sp.efsa.2018.EN-1448

[22] European Centre for Disease Prevention and Control and European Food Safety Authority. Multi country outbreak of *Listeria monocytogenes* sequence type 8 infections linked to consumption of salmon products –25 October 2018. Stockholm and Parma: ECDC/EFSA; 2018

[23] Dallman T, Inns T, Jombart T, Ashton P, Loman N, Chatt C, Messelhauesser U, Rabsch W, Simon S, Nikisins S, Bernard H, le Hello S, Jourdan da-Silva N, Kornschober C, Mossong J, Hawkey P, de Pinna E, Grant K and Cleary P 2016 Phylogenetic structure of European *Salmonella Enteritidis* outbreak correlates with national and international egg distribution network *Microb. Genom.* 2 e000070 doi: 10.1099/mgen.0.000070.

[24] Donachie A, Melillo T, Bubba L, Hartman H and Borg M L 2018 National outbreak of *Salmonella Give* linked to a local food manufacturer in Malta, October 2016 *Epidemiol. Infect.* 146 1425–32

[25] Mellmann A, Harmsen D, Cummings C A, Zentz E B, Leopold S R, Rico A, Prior K, Szczepanowski R, Ji Y, Zhang W, McLaughlin SF, Henkhaus J K, Leopold B, Bielaszewska M, Prager R, Brzoska P M, Moore R L, Guenther S, Rothenberg J M, Karch H 2011 Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology *PLoS One* 6 e22751 doi: 10.1371/journal.pone.0022751

[26] Van Walle I, Björkman J T, Cormican M, Dallman T, Mossong J, Pietzka A, Ruppitsch W, Takkinen J and European Listeria Wgs Typing Group 2018 Retrospective validation of whole genome sequencing-enhanced surveillance of listeriosis in Europe, 2010 to 2015 *Euro Surveill.* 23 pii =1900368

[27] Pichtling A W, Bettengill J B, Luo Y, Baugh J D, Rand H and Strain E 2018 Interpreting Whole-Genome Sequence Analyses of Foodborne Bacteria for Regulatory Applications and Outbreak Investigations *Front Microbiol.* 9 1482 doi: 10.3389/fmicb.2018.0148

[28] Espenhain L, Riess M, Müller L, Colombe S, Ethelberg S, Litrep E, Jernberg C, Kühlmann-Berenson S, Lindblad M, Höve N K, Torpdahl M and Mörk M J 2019 Cross-border outbreak of *Yersinia enterocolitica* O3 associated with imported fresh spinach, Sweden and Denmark, March 2019 *Euro Surveill.* 24 pii =1900368

[29] Sudarikov K, Tyakht A and Alexeev D 2017 Methods for The Metagenomic Data Visualization
and Analysis *Curr. Issues Mol. Biol.* 24 37–58 doi: 10.21775/cimb.024.037

[32] Walsh A M, Crispie F, O’Sullivan O, Finnegan L, Claesson M J and Cotter P D 2018 Species classifier choice is a key consideration when analysing low-complexity food microbiome data *Microbiome.* 650 doi: 10.1186/s40168-018-0437-0

[33] Ghurye J S, Cepeda-Espinoza V and Pop M 2016 Metagenomic Assembly: Overview, Challenges and Applications *Yale J. Biol. Med.* 89(3) 353–362 PubMed PMID: 27698619; PubMed Central PMCID: PMC5045144

[34] Zolfo M, Tett A, Jousson O, Donati C and Segata N 2017 MetaMLST: multi-locus strain-level bacterial typing from metagenomic samples *Nucleic. Acids. Res.* 45(2) e7 doi: 10.1093/nar/gkw837

[35] Highlander S K 2012 High throughput sequencing methods for microbiome profiling: application to food animal systems *Anim. Health Res. Rev.* 13 40–53

[36] Oniciuc E A, Likotrafiti E, Alvarez-Molina A, Prieto M, Santos J A, Alvarez-Ordóñez A 2018 The Present and Future of Whole Genome Sequencing (WGS) and Whole Metagenome Sequencing (WMS) for Surveillance of Antimicrobial Resistant Microorganisms and Antimicrobial Resistance Genes across the Food Chain. *Genes (Basel).* 2018 9 268 doi: 10.3390/genes9050268. Erratum in: Genes (Basel). 2018 p 9