



# Whole Genome Sequencing

Recent advances in rapid, high throughput DNA sequencing have revolutionised the study of biology and are now used in microbial risk assessment (MRA) to enhance food safety and security. 'Next Generation Sequencing' (NGS) technologies have a wide range of applications, including the sequencing of microbial genomes and culture-free identification of microbial communities.

- Competent authorities use this tool to:
  - Source track the origin of bacterial contaminants,
  - Identify sporadic incidents of human cases caused by single food sources,
  - Detection, investigation and control of foodborne outbreak.
- Food manufacturers use it to:
  - Identify spoilage communities in processing environments,
  - Characterize food microbiota,
  - Document non-involvement in specified outbreaks,
  - Authenticate the origin of foodstuffs.

These high throughput technologies are capable of rapidly producing genome sequences of key pathogens of concern, including *Salmonella* spp., *Listeria monocytogenes*, *Campylobacter jejuni* and *Cronobacter* spp. and others. The comparison of those sequences and sequences of isolates from clinical patients allows a more rapid detection of foodborne outbreaks at an early stage. Outbreaks can also be traced back to their source, and even geographical origin.

## Microbial Taxonomy

A bacterial species is represented by a 'type' strain with strains showing a high degree of phenotypic and/or genotypic similarity to that 'type' strain regarded as belonging to the same species. Whilst objective measures of relatedness have been proposed (such as % genome hybridization or sequence similarity) currently there is no simple consensual definition of what constitutes species as a taxonomical unit.

A prokaryotic species is defined by:

- a phylogenetic component given as "the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descendants", and
- a taxonomic component given as "a group of related organisms that can be distinguished from similar groups by a constellation of significant genotypic, phenotypic, and/or ecological characteristics."

Within a species, a strain is a specific isolate which can be discriminated from another strain of the same species. Over the last century, several techniques to perform this discrimination have been developed, culminating in whole genome sequencing as an ultimate-resolution technology for discerning strains and inferring their phylogeny.

## Whole Genome Sequencing as a typing technique for Source Tracking

Whole Genome Sequencing (WGS) was first introduced in 2008 for source tracking in the United States. It soon replaced Pulsed Field Gel Electrophoresis (PFGE) to become the Food and Drug Agency's preferred method for identification and tracking of pathogenic isolates from food samples. Numerous food safety authorities now use the technique widely in other regions of the world.

WGS is an analytical technique that can allow the determination of the complete DNA genome sequence of a microorganism. WGS improves the detection, surveillance and response to food-borne diseases and outbreaks. This technology provides uniform typing systems across the environmental, animal, food and human sectors, offering the potential to trace food-borne contamination back to its microbial sources.

WGS is more repeatable and reproducible than previous techniques used, such as PFGE or ribotyping. Due to the digital nature of genomic data, the strain-level identification can be shared and compared between laboratories over the world using open access international database such as GenomeTrakr. Food microbiologists and clinical microbiologists can now pinpoint the origin of a food borne contaminant by comparison with clinical isolates and epidemiological data. As a result, foodborne illnesses are being more easily tracked to a specific location/source.

Regulatory bodies can now identify the origin of small outbreaks with a limited number of clinical isolates or identify a long-lasting contamination events associated with sporadic cases within the population.

## Whole Genome Sequencing and Food Safety: Beyond Identification

WGS is more accurate than a serotype and more discriminatory than PFGE, ribotyping, RAPD assay and RFLP analysis. Supported with epidemiological data, it can help to draw faster relationships between different isolates within a limited timeframe.

WGS provides the ability to:

- Differentiate sources of contamination, even within the same outbreak
- Determine which ingredient was originally contaminated by the pathogen associated with an illness outbreak and therefore narrow the search for the source of a contaminated ingredient
- Determine unexpected vectors for food contamination and provide information for root cause analysis
- Determine phenotypic properties of isolates such as antibiotic resistance, virulence, sensitivity to cleaning agents or adherence to surfaces

## Challenges and limitation of WGS

Since its recent implementation, the use of WGS in epidemiology, at least in North America, Europe and ASPAC is growing exponentially.

PFGE has been used as the golden standard for strain typing for more than 30 years. Due to capacity building (training, equipment, premises, data processing...), this technology cannot not be replaced overnight by a methodology that is neither standardized nor harmonized yet. PFGE is also still more cost effective than WGS for a global use.

Currently, the major challenges with the application of WGS are:

- WGS methods are more costly than other typing methods
- Lack of a WGS technical standard on methodology
- Lack of internationally agreed “cut-off” values that differentiate between strains (single nucleotide polymorphisms (SNP). The acceptable counted SNPs may be specific to each microorganism
- Interpretation of results requires bioinformatics specialists
- So far, no internationally accepted approach to protect WGS information stored in data-banks from misuse.

It should be emphasized that food safety outbreaks or incidents cannot be solved solely based on WGS data but must always be linked to epidemiological data for confirmation.

## What is the dairy industry doing?

WGS allows the identification and characterization of microorganisms with an unprecedented level of sensitivity and specificity. The technology continues to evolve quickly as a technique, but as yet only limited standards are available. IDF is actively collaborating in the ISO working group for establishing a standard to define WGS as a technique with the requirement of specificity, repeatability and reproducibility.

Process environment monitoring, collection of isolates and WGS analysis to identify and discriminate resident from transient strains is recommended to anticipate contamination of finished product and to avoid causing harm to the consumer population. The technique can also be used for the safety demonstration of Microbial Food Cultures.

## References

EFSA (European Food Safety Authority), García Fierro R, Thomas-Lopez D, Deserio D, Liebana E, Rizzi V and Guerra B, 2018. Outcome of EC/EFSA questionnaire (2016) on use of Whole Genome Sequencing (WGS) for food- and waterborne pathogens isolated from animals, food, feed and related environmental samples in EU/EFTA countries. EFSA supporting publication 2018:EN-1432, 49 pages

ISO/TC34/SC9, 2018. Microbiology of the Food Chain — Genomic sequencing of foodborne microorganisms — General requirements and guidance for bacterial genomes. N2133, working document

FAO-WHO, 2016. Applications of Whole Genome Sequencing in food safety management. Technical Background Paper. Available at: <http://www.fao.org/3/a-i5619e.pdf>

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., Pouseele H., Klimke W., Besser J., Brown E., Allard M., Gerner-Smidt P., Implementation of Nationwide Real-time Whole-genome Sequencing to Enhance Listeriosis Outbreak Detection and Investigation. Clin Infect Dis. 2016 Aug 1;63(3):380-6.

Kovac, J., den Bakker, H., Carroll, L.M., Wiedmann, M., 2017. Precision food safety: A systems approach to food safety facilitated by genomics tools. Trends in Analytical Chemistry 96 52-61

Technical University of Denmark - National Food Institute; Istituto Zooprofilattico Sperimentale del Lazio e della Toscana; Federal Institute for Risk Assessment; National Institute of Public Health – National Institute of Hygiene; National Veterinary Research Institute; Public Health England; Animal and Plant Health Agency, and Istituto Zooprofilattico Sperimentale delle Venezie, 2018. Final report of ENGAGE - Establishing Next Generation sequencing Ability for Genomic analysis in Europe. EFSA supporting publication 2018:EN-1431. 252 pages.



International Dairy Federation  
[www.fil-idf.org](http://www.fil-idf.org)