# Foodborne bacterial pathogens: genome-based approaches for enduring and emerging threats in a complex and changing world

Alison E. Mather[1,2†], Matthew W. Gilmour[1,2], Stuart W.J. Reid[3] and Nigel P. French[4]

[1]Quadram Institute Bioscience, Norwich, United Kingdom
[2]University of East Anglia, Norwich, United Kingdom
[3]Royal Veterinary College, Hatfield, United Kingdom
[4]Tāuwharau Ora, School of Veterinary Science, Te Kunenga Ki Pūrehuroa, Massey University, Papaioea, Palmerston North, Aotearoa, New Zealand

[†]email: Alison.Mather@quadram.ac.uk

## Abstract

Foodborne illnesses pose a substantial health and economic burden, presenting challenges in prevention due to the diverse microbial hazards that can enter and spread within food systems. Various factors, including natural, political and commercial drivers, influence food production and distribution. The risks of foodborne illness will continue to evolve in step with these drivers and changes to food systems. For example, climate impacts on water availability during agriculture, changes in food sustainability targets, and evolving customer preferences can all have an impact on the ecology of foodborne pathogens and the agrifood niches that can carry microorganisms. Whole genome and metagenome sequencing, combined with microbial surveillance schemes and insights from the food system, can provide authorities and businesses with transformative information to address risks and implement new food safety interventions across the food chain. In this Review, we describe how genome-based approaches have advanced our understanding of the evolution and spread of enduring bacterial foodborne hazards, as well as their role in identifying emerging foodborne hazards. Furthermore, foodborne hazards exist in complex microbial communities across the entire food chain, and consideration of these co-existing organisms is essential to understanding the entire ecology supporting pathogen persistence and transmission in an evolving food system.

## Introduction

The world is changing at an unprecedented pace, with the current decade being defined by our realisations of the near- and long-term impacts of a changing climate, set amidst the societal and economic influences of the COVID-19 pandemic. The impacts to our food may be just as widespread, with an increasingly globalised food supply reaching 9 billion consumers that ultimately relies on the success of localised production activities.

Food systems comprise widespread but connected activities to produce, distribute, store, and sell foods, with food businesses and regulators seeking economic outcomes through

strong enterprise, health outcomes through safe and nutritious diets, and environmental outcomes through improved land use[1,2]. These food system activities and outcomes are impacted by similarly diverse drivers — for example, the availability of natural resources, changes in consumer preference for different food types, and the flux of supply chains and food pricing[3]. Bacteria are present throughout these systems.

There is a connection between how bacteria transmit through our natural world and the complex agrifood structures used to produce foods. In this Review, we describe how an understanding of microbial ecology can be used to inform and prevent food safety risks. Foodborne pathogens such as *Salmonella enterica* and *Campylobacter* spp. are hazards synonymous with food safety, with food safety itself being an integral part of food security, because 'if it's not safe, it's not food'[4]. As such, the Food and Agriculture Organization (FAO) notes that food safety activities result in a multitude of benefits. These benefits extend beyond merely preventing illness, but also encompass the reduction of food waste and the adaptation of systems to support resilient and healthy ways of producing foods[5]. Within the UK, both food safety and food security feature as key strategies to promote resilience to the impacts of climate change[6].

Considering the connection between ecological and socioeconomic drivers within food systems, we take the view that there are multiple consequences for food safety that may stem from new patterns of microbial transmission within agrifood settings (Figure 1). For example, climate change can alter the prevalence, abundance and distribution of bacterial species in aquatic and marine systems, and may lead to widespread impacts to human, plant and animal health[7,8]. The detection of foodborne pathogens has significantly evolved; the development and increasing affordability of whole genome and metagenome sequencing now make it feasible to apply these techniques within programmes conducting microbial surveillance, outbreak detection, and regulatory testing, allowing the observation of new patterns. Genomics and metagenomics are tools that have been identified to cut across different sectors — including in health, agriculture and the environment, in both high and low resource settings — offering enhanced insights into how to manage foodborne diseases[9].

In this Review, we put these contemporary tools into context in relation to both enduring and emerging foodborne bacterial hazards. We base our definition of emerging hazards on those "that have newly appeared in a population or have existed but are rapidly increasing in incidence or geographic range"[10]; some foodborne or potential foodborne pathogens fitting this definition include Shiga toxin-producing *Escherichia coli* O104:H4 and *Salmonella enterica* subsp. *enterica* serovar Typhimurium sequence type 313 (ST313), which emerged due to changes in food production practices and amplification through the food chain, as well as pathogen evolution[11-14]. We also extend this definition to include hazards which likely have long existed but have only recently been identified due to advances in sequencing and other technologies, examples being Candidatus *Campylobacter infans*[15] (associated with breastfeeding) and *Arcobacter* spp. [16] (Figure 2). Enduring hazards are those that have persisted for decades and, from a scientific standpoint, have often been observed and classified at taxonomic levels such as genus and species, as well as subspecies classifications such as serotype and sequence type. However, the level of observable resolution within subspecies classifications is being re-written by whole genome sequencing

(WGS), and as will be explored in this Review, as are the definitions of enduring hazards that may now be viewed as emerging hazards based upon the revelation of specific genetic traits that explain the expansion of subspecies in agrifood niches. Although this definition is dichotomised, the true nature of foodborne hazards is a continuum, and when viewed under the lens of genomics, what is classified as emerging can shift to enduring, or *vice versa*. The US Centers for Disease Control and Prevention uses WGS to define bacterial strains causing disease as 'Reoccurring' (strains that repeatedly but periodically cause acute outbreaks), 'Emerging' (previously novel or rare strains that are increasingly causing illness) or 'Persisting' (strains that consistently cause disease over time)[17]. In this Review, we will cover how advances in genomics and metagenomics have revolutionised how we identify bacterial foodborne pathogens, track their evolution and spread, and how our understanding of what hazards are present and how to control them can change.

## Bacterial foodborne human pathogens

Foodborne illness remains a significant threat to human health and wealth despite advances in awareness, food hygiene and sanitation. In 2010, the World Health Organization (WHO) estimated that almost 1 in 10 people suffered foodborne illness, with 420,000 dying as a result[18]. This was not evenly distributed across geographical regions or demographics, with Asia, Africa (Figure 3a) and children bearing the heaviest burden. The WHO estimate includes cases caused by 31 foodborne hazards, of which 12 were bacterial. However, bacterial agents have an enormous impact, causing five, six and seven of the top 10 foodborne illnesses, foodborne deaths, and foodborne disability-adjusted life years, respectively[19]; therefore, this Review focuses on foodborne bacterial hazards.

### Enduring hazards

There is good consensus between food and health agencies on the bacterial genera and species considered as key foodborne hazards, with the importance of many being recognised for decades (Figure 3b). Bacterial foodborne pathogens generally cause disease in one of two ways once ingested. First, through colonisation and subsequent penetration through the gastrointestinal mucus, causing damage to the epithelial cells and subsequent diarrhoea. Examples of key pathogens causing disease in this manner are *Salmonella enterica, Campylobacter jejuni,* and *Listeria monocytogenes,* and in some cases, disease progresses to invasive illness or systemic syndromes (for example, bloodstream infections and sepsis). Second, foodborne pathogens can cause disease through the production of toxins, as seen in cases such as *Clostridium perfringens,* Shiga toxin-producing *Escherichia coli* O157, *Bacillus cereus,* and *Clostridium botulinum*. The biology, epidemiology and pathogenesis of these and other foodborne bacterial hazards have been reviewed extensively elsewhere[20,21]. With the increasingly large numbers of whole genome sequences of bacterial pathogens being produced in research and coordinated surveillance programmes, a more nuanced picture is emerging on the population structures and underlying genetic determinants associated with their transmission in their associated niches.

## Emerging hazards

Although bacterial foodborne pathogens like *Salmonella enterica*. and *C. jejuni* have been recognised for decades, the threat they pose to human health is compounded by the complex systemic risks associated with the production and global distribution of current and future foods, influenced by continually changing natural and political environments (Figure 1). These factors not only mean that foodborne hazards may evolve and change, but they also provide opportunities for the emergence and spread of new or previously unrecognised hazards that cause disease.

*Whole genome sequencing — not all foodborne pathogens are created equal*

Historically, surveillance systems for foodborne pathogens relied on techniques such as serotyping or pulsed-field gel electrophoresis, which enabled differentiation of pathogens beyond genus and species levels and facilitated the linking of isolates to common sources in cases where these subspecies characterisations correlated with epidemiological findings[22,23]. These techniques often lack high discriminatory power; the advent of higher resolution WGS data has enhanced our ability to define epidemiologically-related clusters of isolates while also understanding the specific mechanisms and diverse ability of individual bacterial strains to cause disease within species or genus[24]. Bacterial evolution, such as the acquisition of mutations or new genetic elements, can result in increased pathogenicity, virulence, or host range, and subsequent emergence of a new pathogen. For example, successful waves of different subtypes of *S.* Typhimurium, which were originally distinguished by phenotypic methods including serotyping and phage typing including definitive type (DT) 204c and DT104[25], have dominated the cases of human disease caused by this organism. The use of WGS has provided insight into the genetic mechanisms facilitating the success of the various subtypes, including the current pandemic subtype monophasic *S.* Typhimurium ST34. In this case, the acquisition of antimicrobial resistance (AMR) genes, a novel prophage and a genomic island encoding metal tolerance genes likely facilitated the transition of this organism from an emerging hazard, with pigs being the most likely reservoir, to an enduring hazard. Monophasic *S.* Typhimurium ST34 is now one of the most common *Salmonella* subtypes globally, found across multiple host species and countries[26]. Similarly, in the Brazilian poultry industry where antimicrobial selection pressure is high, two distinct serovars of *Salmonella* (*Salmonella enterica* subsp. *enterica* serovar Minnesota and *Salmonella enterica* subsp. *enterica* serovar Heidelberg) became dominant through acquisition of AMR plasmids carrying genes conferring resistance to β-lactams, tetracyclines and sulphonamides[27]. Common genomic features including loss-of-function mutations and gain of pathogenicity islands have also been identified in *Salmonella* serovars adapted to specific hosts[12,28], which offers potential for predicting how future pathogens may emerge.

Other examples are acquisition of a gene encoding Shiga toxin by an enteroaggregative *E. coli* strain, leading to the largest outbreak of Shiga toxin-producing *E. coli* ever recorded[11], and *L. monocytogenes,* a highly heterogeneous organism for which the use of WGS has

identified different evolutionary lineages. These evolutionary lineages have been divided into clonal complexes which are associated with different sources and levels of virulence[29]. The generation of increasingly larger numbers of genome sequences allows the application of genomic epidemiology concepts of source attribution and outbreak response[24], topics that will be explored later in this Review.

Threats to our food system and human health from bacterial pathogens arise at various points across the food chain, from post-primary production to the point of consumption. These threats often manifest in conditions that are very different from the original reservoirs, such as animal gastrointestinal tracts, farm soils or crops[30]. Therefore, key for a successful foodborne pathogen is the ability to survive these diverse conditions, and the use of 'omics technologies is enabling our understanding of the underlying genetic mechanisms. For example, *Cronobacter sakazakii* is a neonatal bacterial pathogen known to persist in low-moisture matrices, such as powdered infant formula. RNA sequencing identified the expression of many genes that were significantly up- or down-regulated as a response to survive desiccation[31]. Phenotypic plasticity has also been observed in major foodborne pathogens, such as *C. jejuni*, where fluoroquinolone resistance has been associated with enhanced survival in the chicken host even in the absence of antimicrobial pressure[32]. Other mechanisms to support survival through the food chain (reviewed elsewhere[33]) can function at the individual bacterium level, such as sporulation[34] and metabolic or transcriptomic changes[35], or at the community or population level, such as single- or multi-species biofilm formation[36,37].

*Contemporary drivers shaping the microbial landscape*

The emergence of foodborne bacterial hazards is driven by many factors, which not only include bacterial evolution itself, but also ecological and socioeconomic drivers such as consumer preference, economy and trade, climate change and technology innovation (Figure 1). Market demands shaped by factors such as shifts in consumer preference is changing the way food is produced in many high-income countries (HICs), leading to new routes for foodborne hazards to enter the food chain. For example, the global demand for sustainable protein and policies to reduce greenhouse gas emissions have led to the development of cell-based foods that derive animal agricultural products such as meat directly from cell cultures; along the production stages, multiple entry points for pathogens have been identified through elicitation of expert opinion[38,39], such as contamination of raw materials and added reagents with bacteria and other microorganisms that can contaminate cells. However, there remains a lack of empirical data. Although little research has been conducted yet on the potential foodborne hazards associated with these new foods, a recent study demonstrated that growth of *Listeria* spp. and *Salmonella enterica* was greater in plant-based milks than in cow milk[40]. A contrasting trend towards increasing consumption of animal source foods is expected in low- and middle-income countries (LMICs) with economic development, which may result in greater numbers of foodborne illness if investment in food safety management does not keep pace[41,42]. The increasingly globalised

nature of food chains can spread and amplify foodborne hazards, which is explored in more detail later in this Review.

Climate change has more extensive impacts than merely the impetus for the development of new food types, affecting the dispersal and persistence of bacterial foodborne hazards[43,44]. In 2022, it was estimated that over half of human pathogenic diseases are exacerbated by climate change, with *Vibrio* spp. as the bacterial foodborne pathogen with the strongest evidence of aggravation by climate hazards[45]. Projections estimate that warming coastal waters will increase the numbers and distribution of vibriosis cases, not only through the extended geographical range of the bacteria but also through increased recreational use of coastal waters[46-48].

Technology innovation in the generation of genome-based data has greatly contributed to our recognition of emerging foodborne hazards. While the increasing cost effectiveness of large scale WGS has revolutionised our ability to understand the biology and characteristics of enduring bacterial pathogens, the majority of foodborne illness is of unknown cause. In the UK, foodborne illness is responsible for an estimated 2.4 million cases, costing in excess of £9bn annually; crucially, £6bn of this cannot be attributed to specific agents and so interventions aimed at reduction are limited by knowledge[49]. The development and use of untargeted, culture-free approaches has been used to investigate both known, and hitherto unknown, bacterial causes of foodborne illness. Whole metagenome sequencing of faecal samples led to the identification of a new candidate species of *Campylobacter*, Candidatus *Campylobacter infans*, which has been found in breastfed infants in sub-Saharan Africa and South Asia[15]. The use of culture-free methods has also highlighted the widespread presence of *Arcobacter* spp. including *Arcobacter cryaerophilus* and *Arcobacter butzleri* in wastewater[50], suggesting previously underappreciated reservoirs of these emerging pathogens (Figure 2). Culture-independent diagnostic tests (CIDTs) provide an agnostic and more rapid approach to identifying potential causative agents of foodborne illness, and facilitate the identification of bacteria which cannot be cultured[51]. However, barriers to its widespread implementation in food safety applications still exist. The proportion of the food metagenome comprised of pathogens is very low[52], and so may not be detected, depending on the amount of DNA from food itself within samples and depth of sequencing. The lack of an isolate presents a barrier to the implementation of genome-based evaluations of outbreaks and transmission at multiple scales[53,54], explored in the next section. Further challenges arise when attempting to classify bacteria present beyond the species level, and to identify the presence of multiple lineages of the same bacterium within a sample; metagenome-assembled genomes (MAGs) can facilitate such identification, particularly with long-read metagenome sequencing, but do not completely solve this problem. As such, 'strain-resolved metagenomics' is an active area of research[55,56].

In terms of food products, all foods will carry bacteria unless specifically sterilised; another advantage of culture-free approaches is that an understanding of the total microbiota — the entire microbial content present — can be obtained. These bacteria may be pathogens, may cause food spoilage or support the persistence of pathogens, or may act as reservoirs of

AMR genes; previous work has shown the concentration of AMR genes on different retail foods to be between $10^4$–$10^{10}$ AMR genes per gram of food[52]. The majority of our knowledge of foodborne pathogens is based on investigation of pure cultures of single strains within the laboratory, but in the food chain pathogens exist within complex microbial communities. These interactions are often critical to the success of foodborne pathogens; for example, previous studies have shown how *C. jejuni* benefits from passive protection of existing *Pseudomonas aeruginosa* biofilms on abiotic surfaces[36] and that multi-species biofilms impart biocide tolerance to *L. monocytogenes*[57].

The novel insights into bacterial foodborne hazards through technological innovation also extends to the analysis of the vast amount of data generated by genome-based approaches. Leveraging the recent advances in artificial intelligence (AI), big data and machine learning, novel genetic factors associated with successful foodborne pathogens have been identified. Such approaches have been used to identify specific subsets of cattle-associated Shiga-toxin producing *E. coli* O157 strains that are more likely to cause human disease, even within previously defined pathogenic lineages[58], and to identify new genes potentially associated with *E. coli* pathogenicity[59,60]. Other applications of machine learning for food safety include the analysis of public search records as a proxy for on-site inspection of restaurants[61].

**Transmission pathways and propagation of pathogens along the food chain**

The prevention and control of both enduring and emerging foodborne bacterial hazards requires an understanding of the factors influencing the propagation of pathogens along the food chain. This extends across multiple scales, both spatial and temporal, encompassing the transmission of pathogens through increasingly interconnected global supply chains, the movement of people and animals across international borders, and localised outbreaks or sporadic cases arising from contaminated food products. Many foodborne pathogens, both enduring and emerging, are amplified in animal reservoirs (maintenance hosts) before contaminating the food supply chain. Hence, the need to adopt a 'One Health' approach[30] in the surveillance of bacterial foodborne hazards that considers sampling across time and space in different environments (Figure 4) to identify robustly the pathways that move pathogens from animal reservoirs to the point of ingestion, infection and disease in human hosts (Figure 5).

Genome-based approaches provide greater resolution for the investigation of transmission at multiple scales, from global to local, and through complex food supply chains and networks. Examples of the application of these approaches are described below. These include determining historical and contemporary patterns of transmission on a global scale, outbreak investigation at local and international scales, and source attribution and source tracking along the food chain.

*Global-scale transmission*

Increased global trade in food products is a major driver of improved food security, notably for countries with rapid population growth and natural resource constraints. However, historical trade in livestock has spread foodborne pathogens between continents and trade increases the potential for contaminated food to be moved further and more rapidly between countries. Genome-based tools provide evidence of the role of historical trade in live cattle in the international transmission of Shiga toxin-producing *E. coli*[62], and insight into the role of transatlantic livestock trade followed by the growth of cattle farming and industrialisation of food production, in the spread of *L. monocytogenes*[63]. Recent advances in machine learning, combined with genome sequencing, have enabled rapid global source tracking of *Salmonella enterica,* identifying the geographical origin of cases arising from contaminated food products in the global supply chain[64].

*Outbreak investigation*

Cases of foodborne illness are often categorised as either outbreak-associated or sporadic, where the latter are not considered to be part of a temporal and/or spatial cluster of cases associated with a common source. By aiding hypothesis-generation and serving as high-quality evidence in public health investigations, high resolution genome-based methods have helped to reduce the size and duration of outbreaks in the United States compared to lower resolution molecular tools[65] (Figure 6). This is achieved by comparison of isolates from clinical cases with isolates from food or food production environments; isolates from food sources that cluster together with human isolates, as a result of differing by a small number of single nucleotide polymorphisms (SNPs) or alleles, are more likely to have a recent common ancestor and be implicated as the source of the outbreak (for example, the large scale outbreak of listeriosis in South Africa[66]). Genome sequencing of presumptive sporadic cases has also identified previously undetected clusters of cases that are more likely to share a common source, providing evidence of covert outbreaks[67,68].

Determining membership of a cluster of cases with a putative common origin can be difficult. A study[69] indicates that three metrics can provide evidence for a common origin of sequenced isolates, thereby indicating their association with of an outbreak and links to food sources. These metrics are the number of SNPs, the location of isolates relative to others in the tree topology (for instance, monophyly provides greater support than paraphyly and polyphyly) and bootstrap support for the branch containing the isolates in question. The number of SNPs or shared alleles that are consistent with a recent common ancestor depends on the organism and period of sampling. A study[29] used the distribution of core genome multilocus sequence typing (cgMLST) allele differences between isolates from known outbreak sets and all cases of listeriosis caused by *L. monocytogenes* to propose a cutoff of <7 allele differences for membership of an outbreak, whereas a recent modelling study identified a wider range of epidemiologically relevant cutoffs not only within *L. monocytogenes* but also a range of SNP or cgMLST allele cutoffs within other common bacterial foodborne hazards[70].

Technology for high throughput sequencing has advanced rapidly, but spatial resolution and speed is important to pinpoint the origin of an outbreak to a specific locale[71]. This

underlines the importance of surveillance, facilitated by global collaboration and data sharing among researchers, public health agencies and industry, including shared data repositories of genome sequences and accompanying metadata[72]. Examples include: National Center for Biotechnology Information[73]; GenomeTrakr Network[74]; Global Microbial Identifier[75]; and PubMLST[76].

Rapid sharing of genomic and associated data helps to resolve local and international outbreaks of enduring pathogens, and aid the characterisation, detection and control of emerging pathogens[11] (for example, aggregate datasets from multiple programmes, informing interventions such as novel vaccine designs[77]).

The ability to apply the techniques described above depends on the availability of large repositories of bacterial whole genome sequences with high quality associated epidemiological metadata, such as date of isolation, location, and host species. Considerable global efforts by publicly-funded and commercial laboratories and public health and food safety authorities have helped to develop the strategic and operational requirements to support widespread implementation and routine use of genomics as a tool for food safety purposes, including sharing of metadata and the standardisation of terminology, bioinformatic analytics, databases and data access (reviewed in refs.[24,51,78,79]). The decreasing cost of sequencing has allowed the implementation of WGS-based surveillance of bacterial foodborne hazards in many high resource settings[23,80,81]. However, given the global nature of travel, food supply chains and bacterial foodborne hazards, the lack of substantial similar data from low resource settings is a barrier to applying these methods in such settings or even understanding the major food safety concerns there[82,83]. Furthermore, the identification of emerging hazards will be even more difficult if they arise in a region with no or limited (sequencing) surveillance capacity. The main barrier to implementation of WGS for food safety in low resource settings identified in a recent Review was lack of governmental dedication of resources for WGS implementation[83]. Additional issues identified were a lack of bioinformatics expertise, computing resources and insufficient resources and infrastructures[82,83], although fundamentally, countries without an established surveillance system are unlikely to see the benefit of adding WGS capacity[83].

The development of more accessible tools for analysing and interpreting genomic data has improved the accessibility, visualisation and presentation of public and private genome sequencing and related metadata. Examples of such platforms include interactive web-based tools such as Nextstrain[84], Microreact[85] and Pathogenwatch[86]. These are dynamic and scalable, and provide a variety of phylogenetic, genomic and epidemiological tools. The developers of Nextstrain and Microreact provide open-source software for end-users to design bespoke applications. As more genomics data with associated metadata and improved bioinformatics and statistical tools become available, the utility of these applications for decision makers will increase. This was evident during the COVID-19 pandemic, where pathogen-specific genomic tools were rapidly developed, providing easy to interpret, interactive websites that were designed to aid public health decision making[87], and which have also been used for foodborne bacterial pathogens[88-90].

*Source attribution and source tracking: determining reservoirs, pathways and risk factors for foodborne illness*

The availability of higher resolution genomic data has resulted in development of new model-based approaches for unravelling the complex chain of events that lead to foodborne illness, from animal reservoirs to the host factors that lead to ingestion, infection and disease (Figure 5). Determining the 'source' of contamination of food with foodborne pathogens helps the implementation of appropriate responses to reduce the burden of foodborne illness. This includes understanding the relative importance of different animal reservoirs and infection pathways, identifying the source of contamination in food processing and determining host-associated risk factors.

Reservoirs and pathways. Determining the relative contribution of different animal reservoirs to the overall burden of disease caused by foodborne pathogens helps the implementation of effective control measures and can lead to measurable improvements in food safety and public health[91]. However, there are often multiple animal reservoirs and transmission pathways involved in cases of illness associated with enduring pathogens such as *C. jejuni* and *Salmonella enterica*. Animal reservoirs include wildlife[92] as well as food-producing animals, and faeces from the same reservoir can contaminate multiple pathways, for example leading to both food and waterborne infections (Figure 5). Many techniques have been devised for determining the relative contribution of different animal reservoirs and pathways, including expert opinion, exposure modelling and, more recently, molecular and genomic methods[93,94].

Advances in modelling have enabled DNA sequence data from isolates from putative sources to be combined with sequence data from clinical cases to determine the relative contribution of, for example, chickens, cattle and sheep, to the burden of campylobacteriosis and salmonellosis[95-97]. Earlier applications of these models used multilocus sequence typing, typically involving loci in seven genes that were amplified and sequenced using Sanger sequencing[98]. Reservoir attribution models have recently been extended to include covariates, highlighting the differences in attribution in urban and rural areas, as well as the changes occurring before and after interventions in the food chain[99].

The advent of WGS has resulted in the availability of considerably more, higher resolution genomic data for source attribution modelling, including cgMLST and k-mer analyses. The application of machine learning tools for reservoir attribution using WGS data has the potential to provide more accurate and precise estimates of source attribution[100,101]. Other approaches include the use of network analysis[102] and genome-wide identification of host associated SNPs[103]. Non-culture based metagenomic approaches have also been used for reservoir attribution, with analysis conducted using machine learning tools[104]. However, many technical challenges remain in utilizing WGS data for source attribution modelling[105].

Processing plant tracking and food source attribution. The application of WGS was described for identifying the ingredient that was the source of *L. monocytogenes* in ice cream and for

determining whether or not a strain of *Salmonella* was 'resident' in a processing plant[69]. Machine learning and *Listeria* spp. genomic data have been used to identify food sources[106] and to distinguish environmental (for example, soil) from food factory isolates[107]. The latter provides insight into barriers that would limit but not prevent transmission between these two niches.

Recent food industry initiatives have explored the use of metagenomic approaches for monitoring hygiene and source tracking in food chains, processing plants and food production facilities. The approach aims to characterise the 'normal' microbiome of food ingredients and final products to provide a framework for identifying anomalies that may be associated with contamination and potential food safety risks[108]. Further applications of metagenomics in food safety, including food and environmental sampling, have been reviewed in ref.[51]. The Review also describes a number of risks and benefits associated with the application of metagenomics for food safety, and the barriers to application, such as standardisation of methods, cost, and lack of representative datasets and bioinformatics expertise.

Host-associated risk factors. Reservoir attribution models have been combined with a traditional case-control study design to identify risk factors for human cases arising from different pathways. These studies provide both reservoir attribution and risk factor analyses, and include epidemiological studies of campylobacteriosis using 7-gene MLST[109] and genome sequencing[110]

## Food safety interventions informed by genomics

Genomics is now widely used by many government agencies to gather the detailed evidence needed to support surveillance activities and take informed action in the multiple transmission pathways affecting food safety. Likewise, food businesses are now widely considering the application of ever-maturing genomics technologies to improve their individualised circumstances and processes to produce food[24,111,112]. Both government and food business have the same motivation — the safety of foods — with overlapping perspectives on how to reduce the risk of exposure to consumers. Government agencies such as those responsible for public health and food safety commonly take the view of the general population, whereas food businesses such as those producing or serving finished foods take the view of their customers. Both views are connected because the opportunity to understand and act on risks in localised production settings informs, and in turn, is informed by, global food systems and the wider population of food consumers (Figure 4).

The consideration or adoption of new technological advances is a core component to the food safety culture within food businesses, as they seek to identify, understand, and then control foodborne risks[113]. Moreover, in many countries it is a regulatory requirement that food businesses establish a food safety management plan founded on the principles of Hazard Analysis and Critical Control Point (HACCP) that include the establishment of monitoring procedures and a process for corrective action. HACCP sampling plans generally involve laboratory testing for microorganisms such as *Salmonella* spp., *L. monocytogenes*, *Campylobacter* spp. and Shiga toxin-producing *E. coli*[114] where detection of these pathogens can necessitate further investigation and action on root causes[115]. Therefore, within a producer's own food safety culture and best practices, the HACCP principles and resulting management plan are the most rigorous measures that food producers use to ensure food safety risks can be identified and prevented, representing localised change. Given that genomics is now becoming feasible to deploy in these settings, similar to their use in public health programmes where the depth of knowledge surpasses that of current routine microbiology surveillance testing, it is likely that any new genomic testing system will likely be applied within the context of HACCP.

Food businesses that consider incorporating genomics into their testing programmes demonstrate an innovative management approach, which reflects a forward-thinking vision with a willingness to think differently about the nature of microbial risks and how they can be controlled with new behaviours and actions. To develop a genomics capacity that is suited for application in food businesses — such that the methodology can sample across diverse food and environmental matrices[116] while being customisable to the individualised circumstances on how a business produces food — it will be important that there is mutual engagement between food businesses and institutes with expertise in genomics to develop most use-case scenarios. As a future perspective and informed by stakeholder engagements with the food industry (partly through the UK Food Safety Research Network), we propose illustrative examples of questions that food businesses may now ask of their processes and the possible actions they may then take (Box 1). The depth of knowledge gained from

genomics offers those responsible for the supply into food chains improved compliance, consumer assurance, and market and brand protection, through the gained confidence in the safety of products going to market[52,114,115].

## Conclusions and perspectives

Bacterial foodborne pathogens remain a significant threat to human health and economies across the world. The enhanced resolution information provided by WGS and metagenomics has advanced our understanding of these hazards, and demonstrated that enduring and emerging foodborne hazards represent a continuum, with bacterial hazards shifting along this continuum. Furthermore, the complex biological, socioeconomic and ecological drivers shaping the evolution and transmission of bacterial pathogens, and the wider microbial communities in which they exist, emphasise the need for surveillance systems that are sensitive enough to identify both enduring and emerging hazards. Key to these surveillance systems is not only sensitivity, but also timeliness, so that pre-emptive action to prevent foodborne illness can be taken rather than using the information to respond to incidents once they have occurred. Currently, such surveillance systems exist predominantly in high resource settings; despite the decreasing cost of technologies, there is still an inconsistency across the international landscape and a disparity in resource availability between high and low resource settings. Although there have been very few economic evaluations of the application of WGS for foodborne pathogen surveillance[117], the net benefit attributed to the introduction of WGS in some high resource settings is estimated to be substantial[65,118]. In a food supply and trading ecosystem that is now global, these technologies also need to be enabled in low resource settings to obtain a true understanding of the nature and risks of foodborne hazards globally. This requires collaboration between high and low resource settings, underpinned by data sharing and common standards for the generation and interpretation of genome-based approaches to food safety, involving researchers, public health agencies and industry across the various One Health compartments.

As routine microbial testing supports the assessment of hazards within food systems and serves as a signal for further action when hazards are detected, there is a need to understand the more complicated scale of information produced by genome-based testing and then to translate those findings into interventions. For example, identification of specific lineages or genes associated with pathogen success or persistence through the use of WGS and other ‘omics technologies can suggest targets to design future intervention strategies. However, the use of metagenomics may identify the presence of pathogens but it cannot determine whether these organisms are viable or if their identification is an artefact[52]. This may have ethical, commercial and legal implications which need to be resolved. Monitoring microbial communities along the food chain may serve as an early warning system for food safety issues or may help identify candidates for biocontrol agents, but requires a step change in thinking about food safety: “...a modified approach to hazard analysis and risk assessment for food safety may be required; for instance, shifting from risk assessments of individual pathogens to assessments of the entire microbial community and food chain”, as stated in ref.[51]. To exploit the full value of genomic and metagenomic data, it is important to remember that these data should not be considered in isolation. Instead, they need to be integrated with epidemiological and other information across all spatial and temporal scales, from local (for example, movement of people and equipment within a food

processing facility) to global (for example, long-term climate change patterns or international trade practices), in order to effect change.

The true burden of foodborne illness is difficult to estimate due to lack of surveillance, as well as underreporting and challenges in definitively identifying causes of foodborne illness[119]. Food safety is a key component of food security, and in a complex and changing world shaped by dynamic ecological and socioeconomic drivers, improved implementation of genome-based technologies offers an opportunity to strengthen the prevention, detection and response to foodborne bacterial hazards throughout the entire food chain.

## References


1 Zurek, M. *et al.* Assessing sustainable food and nutrition security of the EU food system-an integrated approach. *Sustainability* **10**, doi:10.3390/su10114271 (2018).

2 Parsons, K., Hawkes, C. & Wells, R. Brief 2. Understanding the food system: Why it matters for food policy. (Centre for Food Policy, London, 2019).

3 Hasnain, S., Ingram, J. & Zurek, M. Mapping the UK Food System - a report for the UKRI Transforming UK Food Systems Programme. (Environmental Change Institute, University of Oxford, Oxford, 2020).

4 FAO. *If it isn't safe, it isn't food*, <https://www.fao.org/fao-stories/article/en/c/1179647/> (2019).

5 FAO. FAO strategic priorities for food safety within the FAO strategic framework 2022-2031. (Food and Agricultural Organization of the United Nations, Rome, 2023).

6 Committee, C. C. Progress in adapting to climate change; 2023 Report to Parliament. (2023).

7 EFSA, Maggiore, A., Afonso, A., Barrucci, F. & De Sanctis, G. Climate change as a driver of emerging risks for food and feed safety, plant, animal health and nutritional quality. (European Food Safety Authority, 2020).

8 Morgado, M. E. *et al.* Climate change, extreme events, and increased risk of salmonellosis: foodborne diseases active surveillance network (FoodNet), 2004-2014. *Environ Health* **20**, 105, doi:10.1186/s12940-021-00787-y (2021).

9 FAO. Final meeting report: Technical meeting on the impact of whole genome sequencing (WGS) on food safety management: within a One Health approach. (Food and Agriculture Organization of the United Nations, 2016).

10 Morse, S. S. Factors in the emergence of infectious diseases. *Emerg Infect Dis* **1**, 7-15, doi:10.3201/eid0101.950102 (1995).

11 Grad, Y. H. *et al.* Genomic epidemiology of the *Escherichia coli* O104:H4 outbreaks in Europe, 2011. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.1121491109 (2012).

12 Kingsley, R. A. *et al.* Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res* **19**, 2279-2287, doi:10.1101/gr.091017.109 (2009).

13 Seribelli, A. A. *et al.* Phylogenetic analysis revealed that *Salmonella* Typhimurium ST313 isolated from humans and food in Brazil presented a high genomic similarity. *Braz J Microbiol* **51**, 53-64, doi:10.1007/s42770-019-00155-6 (2020).

14 Almeida, F. *et al.* Multilocus sequence typing of *Salmonella* Typhimurium reveals the presence of the highly invasive ST313 in Brazil. *Infect Genet Evol* **51**, 41-44, doi:10.1016/j.meegid.2017.03.009 (2017).

15 Bian, X. *et al. Campylobacter* abundance in breastfed infants and identification of a new species in the Global Enterics Multicenter Study. *mSphere* **5**, doi:10.1128/mSphere.00735-19 (2020).

16 Ramees, T. P. *et al. Arcobacter*: an emerging food-borne zoonotic pathogen, its public health concerns and advances in diagnosis and control - a comprehensive review. *Vet Q* **37**, 136-161, doi:10.1080/01652176.2017.1323355 (2017).
17 CDC. *Reoccurring, emerging, and persisting enteric bacterial strains*, <https://www.cdc.gov/ncezid/dfwed/outbreak-response/rep-strains.html> (2023).
18 WHO. WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group 2007-2015., (World Health Organization, Switzerland, 2015).
19 Devleesschauwer, B., Haagsma, J. A., Mangen, M.-J. J., Lake, R. J. & Havelaar, A. H. in *Food Safety Economics* (ed T. Roberts) (Springer, 2018).
20 Abebe, E., Gugsa, G. & Ahmed, M. Review on major food-borne zoonotic bacterial pathogens. *J Trop Med* **2020**, 4674235, doi:10.1155/2020/4674235 (2020).
21 Bintsis, T. Foodborne pathogens. *AIMS Microbiol* **3**, 529-563, doi:10.3934/microbiol.2017.3.529 (2017).
22 Foley, S. L., Lynne, A. M. & Nayak, R. Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. *Infect Genet Evol* **9**, 430-440, doi:10.1016/j.meegid.2009.03.004 (2009).
23 Deng, X., den Bakker, H. C. & Hendriksen, R. S. Genomic epidemiology: Whole-genome-sequencing-powered surveillance and outbreak investigation of foodborne bacterial pathogens. *Annu Rev Food Sci Technol* **7**, 353-374, doi:10.1146/annurev-food-041715-033259 (2016).
24 Taboada, E. N., Graham, M. R., Carrico, J. A. & Van Domselaar, G. Food safety in the age of next generation sequencing, bioinformatics, and open data access. *Front Microbiol* **8**, 909, doi:10.3389/fmicb.2017.00909 (2017).
25 Threlfall, E. J. Epidemic *Salmonella* Typhimurium DT104 - a truly international multiresistant clone. *J Antimicrob Chemother* **46**, 7-10 (2000).
26 Petrovska, L. *et al.* Microevolution of monophasic *Salmonella* Typhimurium during epidemic, United Kingdom, 2005-2010. *Emerg Infect Dis* **22**, 617-624, doi:10.3201/eid2204.150531 (2016).
27 Alikhan, N. F. *et al.* Dynamics of *Salmonella enterica* and antimicrobial resistance in the Brazilian poultry industry and global impacts on public health. *PLoS Genet* **18**, e1010174, doi:10.1371/journal.pgen.1010174 (2022).
28 Langridge, G. C. *et al.* Patterns of genome evolution that have accompanied host adaptation in *Salmonella*. *Proc Natl Acad Sci U S A* **112**, 863-868, doi:10.1073/pnas.1416707112 (2015).
29 Moura, A. *et al.* Whole genome-based population biology and epidemiological surveillance of *Listeria monocytogenes*. *Nat Microbiol* **2**, 16185, doi:10.1038/nmicrobiol.2016.185 (2016).
30 Ma, L.-C., Zhao, H.-Q., Wu, L. B., Cheng, Z.-L. & Liu, C. Impact of the microbiome on human, animal and environmental health from a One Health perspective. *Science in One Health* **2**, 100037 (2023).
31 Srikumar, S. *et al.* RNA sequencing-based transcriptional overview of xerotolerance in *Cronobacter sakazakii* SP291. *Appl Environ Microbiol* **85**, doi:10.1128/AEM.01993-18 (2019).
32 Luo, N. *et al.* Enhanced in vivo fitness of fluoroquinolone-resistant *Campylobacter jejuni* in the absence of antibiotic selection pressure. *Proc Natl Acad Sci U S A* **102**, 541-546, doi:10.1073/pnas.0408966102 (2005).
33 Marmion, M., Macori, G., Ferone, M., Whyte, P. & Scannell, A. G. M. Survive and thrive: Control mechanisms that facilitate bacterial adaptation to survive manufacturing-related stress. *Int J Food Microbiol* **368**, 109612, doi:10.1016/j.ijfoodmicro.2022.109612 (2022).
34 Gauvry, E. *et al.* Knowledge of the physiology of spore-forming bacteria can explain the origin of spores in the food environment. *Res Microbiol* **168**, 369-378, doi:10.1016/j.resmic.2016.10.006 (2017).

35 NicAogain, K. & O'Byrne, C. P. The role of stress and stress adaptations in determining the fate of the bacterial pathogen *Listeria monocytogenes* in the food chain. *Front Microbiol* **7**, 1865, doi:10.3389/fmicb.2016.01865 (2016).
36 Teh, A. H. T., Lee, S. M. & Dykes, G. A. Association of some *Campylobacter jejuni* with *Pseudomonas aeruginosa* biofilms increases attachment under conditions mimicking those in the environment. *PLoS One* **14**, e0215275, doi:10.1371/journal.pone.0215275 (2019).
37 Whelan, M. V. X. *et al.* Acquisition of fluoroquinolone resistance leads to increased biofilm formation and pathogenicity in *Campylobacter jejuni*. *Sci Rep* **9**, 18216, doi:10.1038/s41598-019-54620-1 (2019).
38 Ong, K. J. *et al.* Food safety considerations and research priorities for the cultured meat and seafood industry. *Compr Rev Food Sci Food Saf* **20**, 5421-5448, doi:10.1111/1541-4337.12853 (2021).
39 FAO & WHO. Food safety aspects of cell-based food. (Food and Agriculture Organization of the United Nations; World Health Organization, Rome, Italy, 2023).
40 Bartula, K., Begley, M., Latour, N. & Callanan, M. Growth of food-borne pathogens *Listeria* and *Salmonella* and spore-forming *Paenibacillus* and *Bacillus* in commercial plant-based milk alternatives. *Food Microbiol* **109**, 104143, doi:10.1016/j.fm.2022.104143 (2023).
41 Li, M. *et al.* Global disease burden of pathogens in animal source foods, 2010. *PLoS One* **14**, e0216545, doi:10.1371/journal.pone.0216545 (2019).
42 Jaffee, S., Henson, S., Unnevehr, L., Grace, D. & Cassou, E. The safe food imperative: Accelerating progrees in low- and middle-income countries. (World Bank, Washington, D.C., 2019).
43 Hellberg, R. S. & Chu, E. Effects of climate change on the persistence and dispersal of foodborne bacterial pathogens in the outdoor environment: A review. *Crit Rev Microbiol* **42**, 548-572, doi:10.3109/1040841X.2014.972335 (2016).
44 Jones, B. A. *et al.* Zoonosis emergence linked to agricultural intensification and environmental change. *Proc Natl Acad Sci U S A*, doi:10.1073/pnas.1208059110 (2013).
45 Mora, C. *et al.* Over half of known human pathogenic diseases can be aggravated by climate change. *Nat Clim Chang* **12**, 869-875, doi:10.1038/s41558-022-01426-1 (2022).
46 Archer, E. J. *et al.* Climate warming and increasing *Vibrio vulnificus* infections in North America. *Sci Rep* **13**, 3893, doi:10.1038/s41598-023-28247-2 (2023).
47 Vezzulli, L. *et al.* Climate influence on *Vibrio* and associated human diseases during the past half-century in the coastal North Atlantic. *Proc Natl Acad Sci U S A* **113**, E5062-5071, doi:10.1073/pnas.1609157113 (2016).
48 Baker-Austin, C. *et al.* Emerging *Vibrio* risk at high latitudes in response to ocean warming. *Nat Clim Chang* **3**, 73-77 (2013).
49 FSA. The burden of foodborne disease in the UK 2018. (Food Standards Agency, 2020).
50 Kristensen, J. M., Nierychlo, M., Albertsen, M. & Nielsen, P. H. Bacteria from the genus *Arcobacter* are abundant in effluent from wastewater treatment plants. *Appl Environ Microbiol* **86**, doi:10.1128/AEM.03044-19 (2020).
51 Billington, C., Kingsbury, J. M. & Rivas, L. Metagenomics approaches for improving food safety: A review. *J Food Prot* **85**, 448-464, doi:10.4315/JFP-21-301 (2022).
52 Bloomfield, S. J. *et al.* Determination and quantification of microbial communities and antimicrobial resistance on food through host DNA-depleted metagenomics. *Food Microbiol* **110**, 104162, doi:10.1016/j.fm.2022.104162 (2023).
53 Carleton, H. A. *et al.* Metagenomic approaches for public health surveillance of foodborne infections: Opportunities and challenges. *Foodborne Pathog Dis* **16**, 474-479, doi:10.1089/fpd.2019.2636 (2019).
54 Ray, L. C. *et al.* Changing diagnostic testing practices for foodborne pathogens, Foodborne Diseases Active Surveillance Network, 2012-2019. *Open Forum Infect Dis* **9**, ofac344, doi:10.1093/ofid/ofac344 (2022).

55 Anyansi, C., Straub, T. J., Manson, A. L., Earl, A. M. & Abeel, T. Computational methods for strain-level microbial detection in colony and metagenome sequencing data. *Front Microbiol* **11**, 1925, doi:10.3389/fmicb.2020.01925 (2020).

56 Kang, X., Luo, X. & Schonhuth, A. StrainXpress: strain aware metagenome assembly from short reads. *Nucleic Acids Res* **50**, e101, doi:10.1093/nar/gkac543 (2022).

57 Rolon, M. L., Voloshchuk, O., Bartlett, K. V., LaBorde, L. F. & Kovac, J. Multi-species biofilms of environmental microbiota isolated from fruit packing facilities promoted tolerance of *Listeria monocytogenes* to benzalkonium chloride. *Biofilm* **7**, 100177, doi:10.1016/j.bioflm.2024.100177 (2024).

58 Lupolova, N., Dallman, T. J., Matthews, L., Bono, J. L. & Gally, D. L. Support vector machine applied to predict the zoonotic potential of *E. coli* O157 cattle isolates. *Proc Natl Acad Sci U S A* **113**, 11312-11317, doi:10.1073/pnas.1606567113 (2016).

59 Im, H., Hwang, S. H., Kim, B. S. & Choi, S. H. Pathogenic potential assessment of the Shiga toxin-producing *Escherichia coli* by a source attribution-considered machine learning model. *Proc Natl Acad Sci U S A* **118**, doi:10.1073/pnas.2018877118 (2021).

60 Burgaya, J. *et al.* The bacterial genetic determinants of *Escherichia coli* capacity to cause bloodstream infections in humans. *PLoS Genet* **19**, e1010842, doi:10.1371/journal.pgen.1010842 (2023).

61 Sadilek, A. *et al.* Machine-learned epidemiology: real-time detection of foodborne illness at scale. *NPJ Digit Med* **1**, 36, doi:10.1038/s41746-018-0045-1 (2018).

62 Franz, E. *et al.* Phylogeographic analysis reveals multiple international transmission events have driven the global emergence of *Escherichia coli* O157:H7. *Clin Infect Dis* **69**, 428-437, doi:10.1093/cid/ciy919 (2019).

63 Moura, A. *et al.* Emergence and global spread of *Listeria monocytogenes* main clinical clonal complex. *Sci Adv* **7**, eabj9805, doi:10.1126/sciadv.abj9805 (2021).

64 Bayliss, S. C. *et al.* Rapid geographical source attribution of *Salmonella enterica* serovar Enteritidis genomes using hierarchical machine learning. *Elife* **12**, doi:10.7554/eLife.84167 (2023).

65 Brown, B., Allard, M., Bazaco, M. C., Blankenship, J. & Minor, T. An economic evaluation of the Whole Genome Sequencing source tracking program in the U.S. *PLoS One* **16**, e0258262, doi:10.1371/journal.pone.0258262 (2021).

66 Thomas, J. *et al.* Outbreak of listeriosis in South Africa associated with processed meat. *N Engl J Med* **382**, 632-643, doi:10.1056/NEJMoa1907462 (2020).

67 Kovanen, S. M. *et al.* Multilocus sequence typing (MLST) and whole-genome MLST of *Campylobacter jejuni* isolates from human infections in three districts during a seasonal peak in Finland. *J Clin Microbiol* **52**, 4147-4154, doi:10.1128/JCM.01959-14 (2014).

68 Yang, C. *et al.* Outbreak dynamics of foodborne pathogen *Vibrio parahaemolyticus* over a seventeen year period implies hidden reservoirs. *Nat Microbiol* **7**, 1221-1229, doi:10.1038/s41564-022-01182-0 (2022).

69 Pightling, A. W. *et al.* Interpreting whole-genome sequence analyses of foodborne bacteria for regulatory applications and outbreak investigations. *Front Microbiol* **9**, 1482, doi:10.3389/fmicb.2018.01482 (2018).

70 Duval, A., Opatowski, L. & Brisse, S. Defining genomic epidemiology thresholds for common-source bacterial outbreaks: a modelling study. *Lancet Microbe* **4**, e349-e357, doi:10.1016/S2666-5247(22)00380-9 (2023).

71 Hoffmann, M. *et al.* Tracing origins of the *Salmonella* Bareilly strain causing a food-borne outbreak in the United States. *J Infect Dis* **213**, 502-508, doi:10.1093/infdis/jiv297 (2016).

72 Dooley, D. M. *et al.* FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. *NPJ Sci Food* **2**, 23, doi:10.1038/s41538-018-0032-6 (2018).

73 NCBI. *National Center for Biotechnology Information datasets*, <https://www.ncbi.nlm.nih.gov/datasets/> (2024).

74 Timme, R. E. *et al.* GenomeTrakr proficiency testing for foodborne pathogen surveillance: an exercise from 2015. *Microb Genom* **4**, doi:10.1099/mgen.0.000185 (2018).

75 Moran-Gilad, J. *et al.* Proficiency testing for bacterial whole genome sequencing: an end-user survey of current capabilities, requirements and priorities. *BMC Infect Dis* **15**, 174, doi:10.1186/s12879-015-0902-3 (2015).

76 Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* **3**, 124, doi:10.12688/wellcomeopenres.14826.1 (2018).

77 Baker, K. S. *et al.* Genomics for public health and international surveillance of antimicrobial resistance. *Lancet Microbe* **4**, e1047-e1055, doi:10.1016/S2666-5247(23)00283-5 (2023).

78 Kovac, J., den Bakker, H., Carroll, L. M. & Wiedmann, M. Precision food safety: A systems approach to food safety facilitated by genomics tools. *Trac-Trends in Analytical Chemistry* **96**, 52-61, doi:10.1016/j.trac.2017.06.001 (2017).

79 WHO. Whole genome sequencing as a tool to strengthen foodborne disease surveillance and response: module 1: introductory module. (World Health Organization, 2023).

80 Jackson, B. R. *et al.* Implementation of nationwide real-time whole-genome sequencing to enhance listeriosis outbreak detection and investigation. *Clin Infect Dis* **63**, 380-386, doi:10.1093/cid/ciw242 (2016).

81 Dallman, T. J. *et al.* Whole-genome sequencing for national surveillance of Shiga toxin-producing *Escherichia coli* O157. *Clin Infect Dis* **61**, 305-312, doi:10.1093/cid/civ318 (2015).

82 Grace, D. Food safety in low and middle income countries. *Int J Environ Res Public Health* **12**, 10490-10507, doi:10.3390/ijerph120910490 (2015).

83 Apruzzese, I. *et al.* Investing in food safety for developing countries: Opportunities and challenges in applying whole-genome sequencing for food safety management. *Foodborne Pathog Dis* **16**, 463-473, doi:10.1089/fpd.2018.2599 (2019).

84 Hadfield, J. *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121-4123, doi:10.1093/bioinformatics/bty407 (2018).

85 Argimon, S. *et al.* Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom* **2**, e000093, doi:10.1099/mgen.0.000093 (2016).

86 Argimon, S. *et al.* A global resource for genomic predictions of antimicrobial resistance and surveillance of *Salmonella* Typhi at pathogenwatch. *Nat Commun* **12**, 2879, doi:10.1038/s41467-021-23091-2 (2021).

87 Gangavarapu, K. *et al.* Outbreak.info genomic reports: scalable and dynamic surveillance of SARS-CoV-2 variants and mutations. *Nat Methods* **20**, 512-522, doi:10.1038/s41592-023-01769-3 (2023).

88 Reuter, S. *et al.* Directional gene flow and ecological separation in *Yersinia enterocolitica*. *Microb Genom* **1**, e000030, doi:10.1099/mgen.0.000030 (2015).

89 Rodrigues, J. A. *et al.* Pangenomic analyses of antibiotic-resistant *Campylobacter jejuni* reveal unique lineage distributions and epidemiological associations. *Microb Genom* **9**, doi:10.1099/mgen.0.001073 (2023).

90 Neves, A. *et al.* The Swiss Pathogen Surveillance Platform - towards a nation-wide One Health data exchange platform for bacterial, viral and fungal genomics and associated metadata. *Microb Genom* **9**, doi:10.1099/mgen.0.001001 (2023).

91 Sears, A. *et al.* Marked campylobacteriosis decline after interventions aimed at poultry, New Zealand. *Emerg Infect Dis* **17**, 1007-1015, doi:10.3201/eid/1706.101272 (2011).

92 Gardner, T. J. *et al.* Outbreak of campylobacteriosis associated with consumption of raw peas. *Clin Infect Dis* **53**, 26-32, doi:10.1093/cid/cir249 (2011).

93 Cody, A. J., Maiden, M. C., Strachan, N. J. & McCarthy, N. D. A systematic review of source attribution of human campylobacteriosis using multilocus sequence typing. *Euro Surveill* **24**, doi:10.2807/1560-7917.ES.2019.24.43.1800696 (2019).

94 Pires, S. M., Vieira, A. R., Hald, T. & Cole, D. Source attribution of human salmonellosis: An overview of methods and estimates. *Foodborne Pathog Dis*, doi:10.1089/fpd.2014.1744 (2014).

95 Mullner, P. *et al.* Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach. *Infect Genet Evol* **9**, 1311-1319, doi:10.1016/j.meegid.2009.09.003 (2009).

96 Sheppard, S. K. *et al. Campylobacter* genotyping to determine the source of human infection. *Clin Infect Dis* **48**, 1072-1078, doi:10.1086/597402 (2009).

97 Mughini-Gras, L. *et al.* Tracing the sources of human salmonellosis: a multi-model comparison of phenotyping and genotyping methods. *Infect Genet Evol* **28**, 251-260, doi:10.1016/j.meegid.2014.10.003 (2014).

98 Wilson, D. J. *et al.* Tracing the source of campylobacteriosis. *PLoS Genet* **4**, e1000203, doi:10.1371/journal.pgen.1000203 (2008).

99 Liao, S. J., Marshall, J., Hazelton, M. L. & French, N. P. Extending statistical models for source attribution of zoonotic diseases: a study of campylobacteriosis. *J R Soc Interface* **16**, 20180534, doi:10.1098/rsif.2018.0534 (2019).

100 Arning, N., Sheppard, S. K., Bayliss, S., Clifton, D. A. & Wilson, D. J. Machine learning to predict the source of campylobacteriosis using whole genome data. *PLoS Genet* **17**, e1009436, doi:10.1371/journal.pgen.1009436 (2021).

101 Munck, N., Njage, P. M. K., Leekitcharoenphon, P., Litrup, E. & Hald, T. Application of whole-genome sequences and machine learning in source attribution of *Salmonella* Typhimurium. *Risk Anal* **40**, 1693-1705, doi:10.1111/risa.13510 (2020).

102 Wainaina, L. *et al.* Source attribution of human campylobacteriosis using whole-genome sequencing data and network analysis. *Pathogens* **11**, doi:10.3390/pathogens11060645 (2022).

103 Jehanne, Q. *et al.* Genome-wide identification of host-segregating single-nucleotide polymorphisms for source attribution of clinical *Campylobacter coli* isolates. *Appl Environ Microbiol* **86**, doi:10.1128/AEM.01787-20 (2020).

104 Duarte, A. S. R. *et al.* Metagenomics-based approach to source-attribution of antimicrobial resistance determinants - Identification of reservoir resistome signatures. *Front Microbiol* **11**, 601407, doi:10.3389/fmicb.2020.601407 (2020).

105 Pasquali, F., Remondini, D., Snary, E. L., Hald, T. & Guillier, L. Editorial: Integrating whole genome sequencing into source attribution and risk assessment of foodborne bacterial pathogens. *Front Microbiol* **12**, 795098, doi:10.3389/fmicb.2021.795098 (2021).

106 Tanui, C. K., Benefo, E. O., Karanth, S. & Pradhan, A. K. A machine learning model for food source attribution of *Listeria monocytogenes*. *Pathogens* **11**, doi:10.3390/pathogens11060691 (2022).

107 Liao, J. *et al.* Comparative genomics unveils extensive genomic variation between populations of *Listeria* species in natural and food-associated environments. *ISME Commun* **3**, 85, doi:10.1038/s43705-023-00293-x (2023).

108 Beck, K. L. *et al.* Monitoring the microbiome for food safety and quality using deep shotgun sequencing. *NPJ Sci Food* **5**, 3, doi:10.1038/s41538-020-00083-y (2021).

109 Mughini-Gras, L. *et al.* Risk factors for human salmonellosis originating from pigs, cattle, broiler chickens and egg laying hens: a combined case-control and source attribution analysis. *PLoS One* **9**, e87933, doi:10.1371/journal.pone.0087933 (2014).

110 Lake, R. J. *et al.* Source attributed case-control study of campylobacteriosis in New Zealand. *Int J Infect Dis* **103**, 268-277, doi:10.1016/j.ijid.2020.11.167 (2021).

111 Amini, S. in *FoodNavigator Europe* (William Reed, 2018).

112 Gerner-Smidt, P. in *Food Safety Magazine* (2021).
113 GFSI. A culture of food safety: A position paper from the Global Food Safety Initiative (GFSI). (2018).
114 Espinoza, M. S. A., Flink, C., Boisen, N., Scheutz, F. & Käsbohrer, A. Microbiological sampling and analyses in the food business operators' HACCP-based self-control programmes. *Frontiers in Food Science and Technology* **3**, doi:10.3389/frfst.2023.1110359 (2023).
115 CFA. Principles of an environmental monitoring program for the management of *Listeria monocytogenes*. (2023).
116 Jagadeesan, B. *et al.* The use of next generation sequencing for improving food safety: Translation into practice. *Food Microbiol* **79**, 96-115, doi:10.1016/j.fm.2018.11.005 (2019).
117 Tran, M. *et al.* Economic evaluations of whole-genome sequencing for pathogen identification in public health surveillance and health-care-associated infections: a systematic review. *Lancet Microbe* **4**, e953-e962, doi:10.1016/S2666-5247(23)00180-5 (2023).
118 Jain, S., Mukhopadhyay, K. & Thomassin, P. J. An economic analysis of *Salmonella* detection in fresh produce, poultry, and eggs using whole genome sequencing technology in Canada. *Food Research International* **116**, 802-809, doi:10.1016/j.foodres.2018.09.014 (2019).
119 WHO. *Food safety*, <https://www.who.int/news-room/fact-sheets/detail/food-safety> (2022).
120 Ogden, N. H., AbdelMalik, P. & Pulliam, J. Emerging infectious diseases: prediction and detection. *Can Commun Dis Rep* **43**, 206-211, doi:10.14745/ccdr.v43i10a03 (2017).
121 Plowright, R. K. *et al.* Pathways to zoonotic spillover. *Nat Rev Microbiol* **15**, 502-510, doi:10.1038/nrmicro.2017.45 (2017).
122 Falloon, P. *et al.* What do changing weather and climate shocks and stresses mean for the UK food system? *Environ Res Lett* **17**, 051001, doi:10.1088/1748-9326/ac68f9 (2022).
123 Elliott, C. in *New Food* (Russell Publishing Ltd, Brasted, Kent, UK, 2022).
124 McLauchlin, J. *et al.* An outbreak of human listeriosis associated with frozen sweet corn consumption: Investigations in the UK. *Int J Food Microbiol* **338**, 108994, doi:10.1016/j.ijfoodmicro.2020.108994 (2021).
125 Kindle, P., Nuesch-Inderbinen, M., Cernela, N. & Stephan, R. Detection, isolation, and characterization of Shiga toxin-producing *Escherichia coli* in flour. *J Food Prot* **82**, 164-167, doi:10.4315/0362-028X.JFP-18-256 (2019).
126 Bloomfield, S. J. *et al.* Genomic analysis of *Salmonella enterica* serovar Typhimurium DT160 associated with a 14-year outbreak, New Zealand, 1998-2012. *Emerg Infect Dis* **23**, 906-913, doi:10.3201/eid2306.161934 (2017).
127 Pijnacker, R. *et al.* An international outbreak of *Salmonella enterica* serotype Enteritidis linked to eggs from Poland: a microbiological and epidemiological study. *Lancet Infect Dis* **19**, 778-786, doi:10.1016/S1473-3099(19)30047-7 (2019).
128 CDC. *A-Z index for foodborne illness*, <https://www.cdc.gov/foodsafety/diseases/index.html> (2021).
129 EFSA & ECDC. The European Union One Health 2021 Zoonoses Report. *EFSA J* **20**, e07666, doi:10.2903/j.efsa.2022.7666 (2022).
130 Wagenaar, J. A., French, N. P. & Havelaar, A. H. Preventing *Campylobacter* at the source: why is it so difficult? *Clin Infect Dis* **57**, 1600-1606, doi:10.1093/cid/cit555 (2013).
131 CDC. *Detecting outbreaks with whole genome sequencing*, <https://www.cdc.gov/amd/how-it-works/detecting-outbreaks-wgs.html> (2019).

## Acknowledgements

A.E.M. and M.W.G. are supported by the Biotechnology and Biological Sciences Research Council (BBSRC) Institute Strategic Programme Microbes and Food Safety BB/X011011/1 and its constituent project BBS/E/F/000PR13634 (Theme 1, Microbial threats from foods in established and evolving food systems). This work was also supported in part by BBSRC grants BB/V018221/1 and BB/X002985/1 (UK Food Safety Research Network). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Author contributions

The authors contributed equally to all aspects of the article.

## Competing interests

International Patent Application No. PCT/GB2023/050906 entitled “Determination and quantification of the microbial communities and antimicrobial resistance genes on food” in the name of Quadram Institute Bioscience has been filed (priority date 05/04/2022) and is currently in the international phase; A.E.M. is an inventor. This relates to the aspect of the Review where it is mentioned that the potential detection of pathogens through metagenomics can be affected by the amount of contaminating host DNA and sequencing depth. N.F. is a member of the International Commission on Microbiological Specifications for Foods (ICMSF) and an Emeritus Director of the New Zealand Food Safety Science and Research Centre. Both organisations receive support from the food industry and government agencies. Both roles are unpaid and advisory, and related to food safety research, and neither organisation influenced the content contributed by N.F. to the Review. The other authors declare no competing interests.

## Peer review information

*Nature Reviews Microbiology* thanks Steven Djordjevic, who co-reviewed with Veronica Jarocki; Jasna Kovac; and the other, anonymous, reviewer(s) for their contribution to the peer

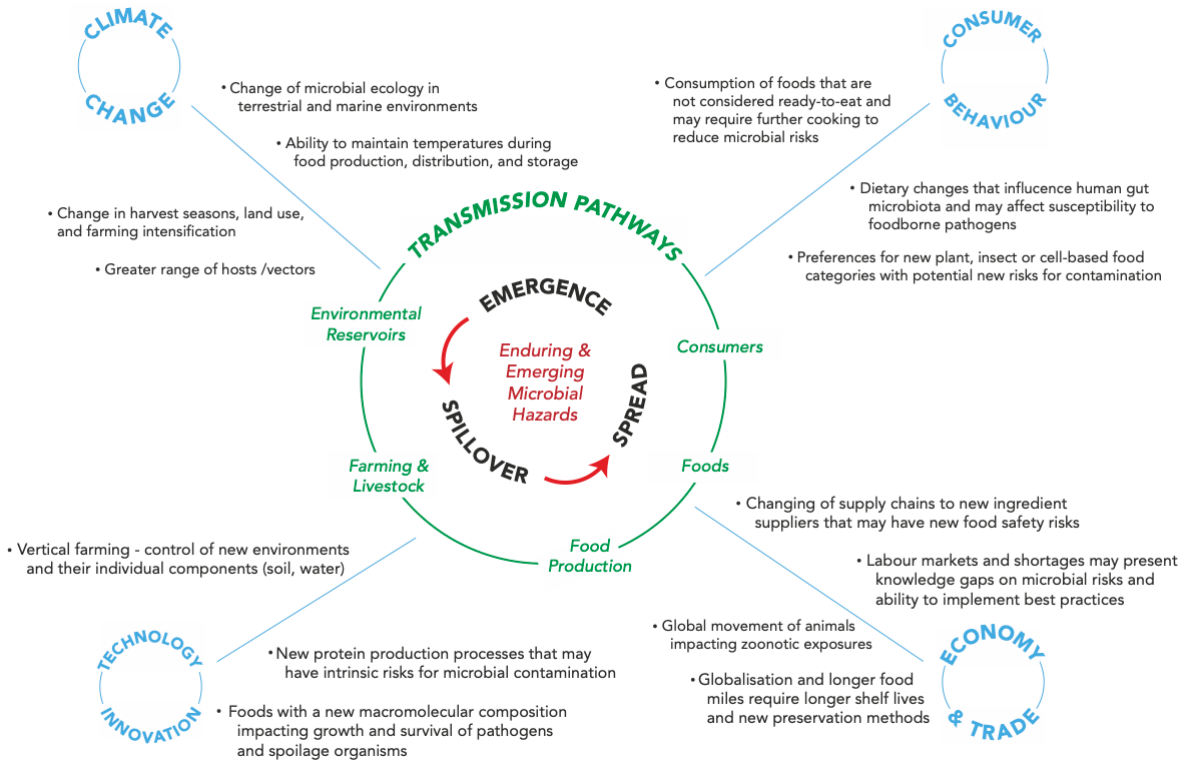



**Figure 1. Ecological and socioeconomic drivers that may shape the microbial landscape of food and food production systems.**

The risk of human exposure to foodborne pathogens and spoilage microorganisms can be amplified due to varied drivers that act across environmental, agricultural, agrifood and consumer niches. These drivers impact and facilitate the emergence and transmission of newly evolved and existing microorganisms within these niches. Drivers include changes to global and local ecologies, innovation in food production technologies, pressures imposed by evolving supply chains and agrifood economies, and the impacts of dietary and consumer choices on food availability, preparation and consumption. Terminology and drivers adapted from refs.[120-123]and examples are cited in the main Review; additionally, examples of potential food risks arising from consumption of foods not considered ready-to-eat include frozen vegetables (with a risk of *Listeria monocytogenes*)[124] and flour (with a risk of Shiga-toxin producing *Escherichia coli*)[125].

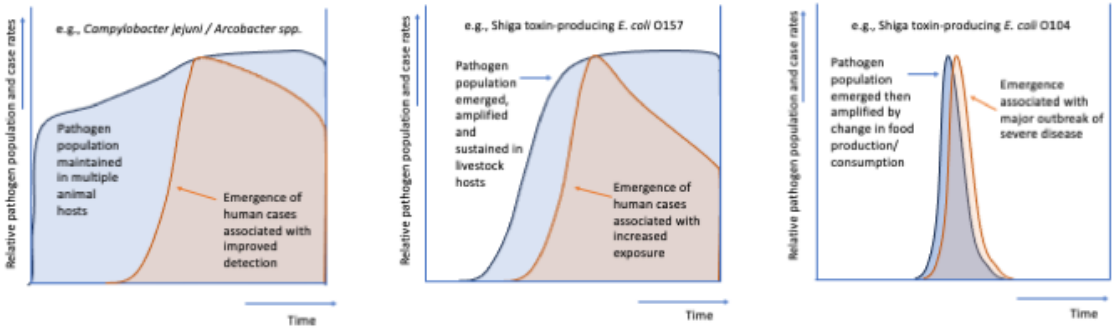
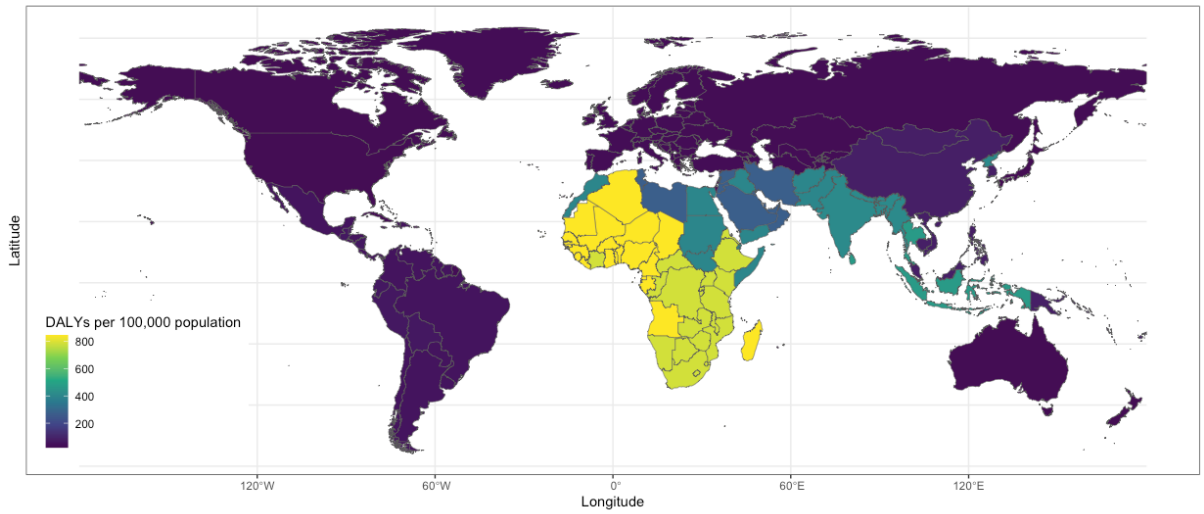



**Figure 2. Relationship between pathogen populations and the dynamics of foodborne illness.** Schematics illustrating putative relationships between pathogen populations (blue) and the dynamics of foodborne illness (red). Emergence may be associated with increased exposure and/or improved detection. Examples of pathogens are indicated in each panel. a) The pathogen population emerged historically and had been amplified as a result of host-species crossover and changes in host dynamics (for example, host population increases and changes in farming practices). The pathogen caused historical human cases of foodborne illness, but these were undetected until improvements in diagnostics led to the apparent emergence. Improved detection resulted in reduced transmission to humans, and potentially reduced transmission in animals. b) Pathogen evolution and changes in host dynamics (for example, movement of farm animals associated with global trade and changes in production practices) resulted in increased exposure to the pathogen via the food chain, which may have coincided with improved detection. Subsequent interventions have reduced human exposure and lowered amplification in animals. c) A pathogen evolved to acquire increased virulence and was amplified following contamination of a food production system. Identification of the source results in interventions that lead to a rapid decline in human cases and the pathogen population. Relative pathogen populations can be inferred from effective population sizes estimated using ancestral state reconstruction methods[126 62 127]. Foodborne illness dynamics can be derived from reported human cases[62].

a)




b)

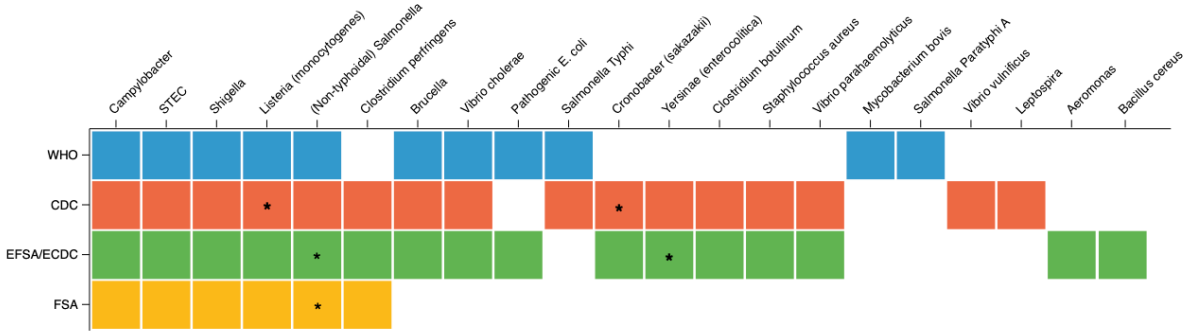



**Figure 3. International impacts and prioritisation of bacterial foodborne illness.**

a) World map showing median rates of foodborne Disability Adjusted Life Years (DALYs) per 100,000 population, caused by bacterial pathogens considered by the World Health Organization Foodborne Disease Burden Epidemiology Reference Group: *Campylobacter* spp., enteropathogenic *Escherichia coli*, enterotoxigenic *E. coli*, Shiga toxin-producing *E. coli* (STEC), non-typhoidal *Salmonella enterica*, *Shigella* spp., *Vibrio cholerae*, *Brucella* spp., *Listeria monocytogenes*, *Mycobacterium bovis*, *Salmonella* Paratyphi A, *Salmonella* Typhi. Data represent the estimated total DALYs from these pathogens in 2010; data obtained from Table 8 in ref.[18]. b) Enduring bacterial foodborne hazards that are widely recognised by public health and food agencies, including the World Health Organization (WHO)[18], UK Food Standards Agency (FSA)[49], US Centers for Disease Control (CDC)[128] and the European Food Safety Authority (EFSA) and European Centre for Disease Prevention and Control (ECDC)[129]. For the CDC and WHO lists of bacterial foodborne hazards, non-Shiga toxin-producing *Escherichia coli* (non-STEC) listed were re-categorised as 'Pathogenic *E. coli*' to match the nomenclature of the European Food Safety Authority (EFSA) and the European Centre for Disease Prevention and Control (ECDC) list. *, indicates when the agency recognises the microorganism to the genus level.

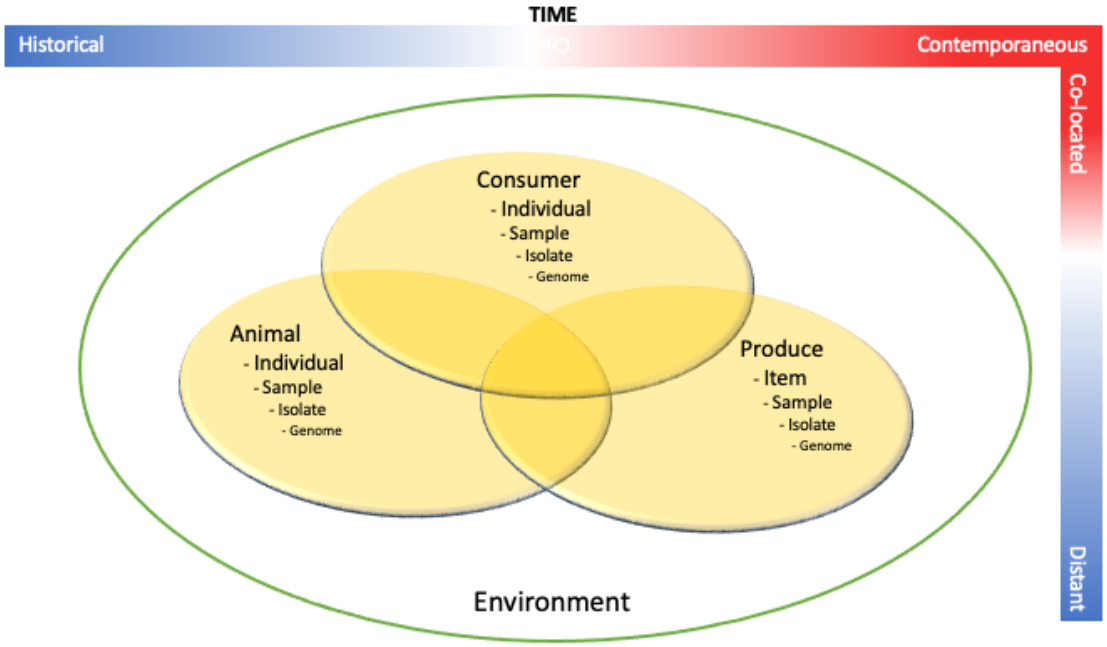



**Figure 4. One Health approach to ecological frameworks and epidemiology in food safety**

A One Health approach to microbiological safety of food requires an ecological framework that embraces the interactions of consumers, animals and produce in the environment in which they interact. The resolution of sampling frame and sample size calculations — the number of samples taken at population level, the number of samples taken per subject, the number of isolates per sample phenotyped and the number of genomes sequenced — are necessary considerations for surveillance design, outbreak investigation and quantification of risk. Notwithstanding the advanced technologies now available, association and causality must still be considered within classical epidemiological parameters of spatial relatedness and temporal directionality; contemporaneous, and co-located data remain the cornerstones of causal inference and plausible source attribution.

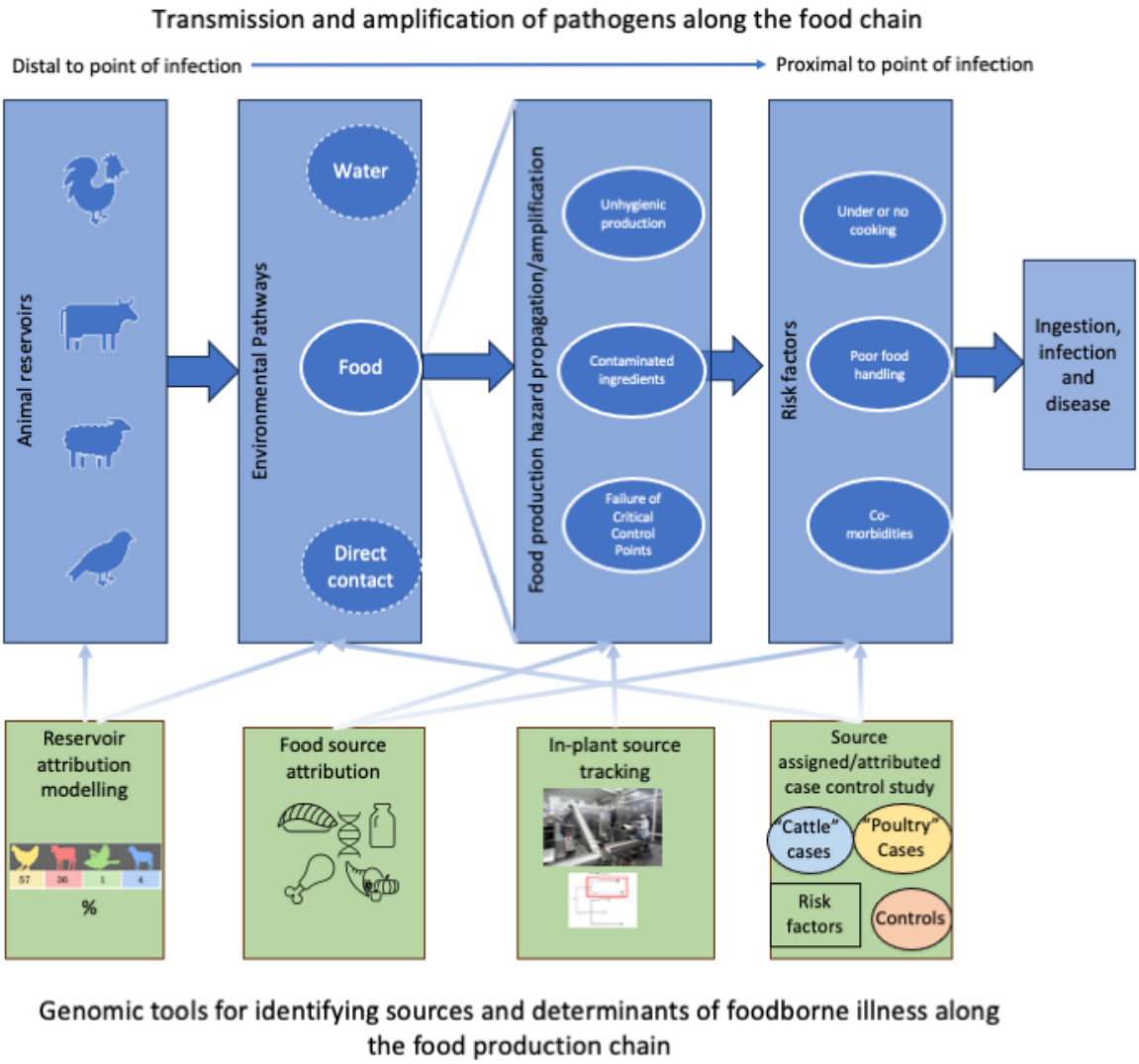



**Figure 5. Tracing bacterial hazards in the food chain.**

Many bacterial hazards that enter the food chain are amplified and maintained in animal reservoirs, including domestic farmed animals and wildlife. Illustrated are the critical steps along the food chain that propagate foodborne pathogens from the animal reservoir to the point of human ingestion, infection and disease. For many enduring pathogens, such as *Campylobacter jejuni* and *Salmonella enterica*, there may be multiple animal reservoirs that can contaminate food and other transmission pathways (for example, water and direct faecal contact as indicated in boxes with dashed lines). Molecular tools, including whole genome sequencing and metagenomics, combined with evolutionary models, are used to determine critical events along the food chain (blue boxes). This includes reservoir attribution models, food processing facility source tracking and source attributed case-control studies (red boxes). Although these methods are focused on understanding the epidemiology and control of enduring pathogens, often defined at the species-level, genomics tools are highly valuable for the investigation of outbreaks of emerging pathogens[11,66]. If the emerging pathogen persists and becomes a sustained, multijurisdictional epidemic or pandemic, large-scale sequencing will facilitate more comprehensive food-chain and epidemiological investigations. Derived from concepts explored in ref.[130].

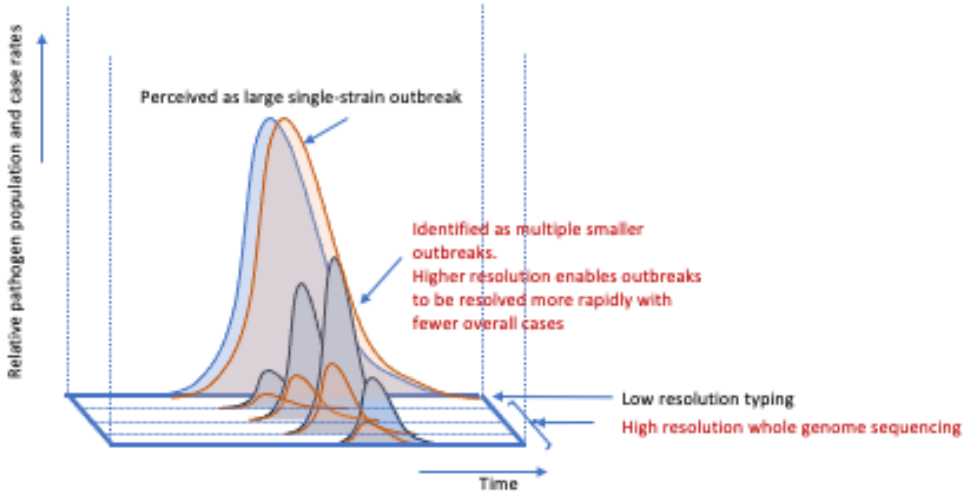



**Figure 6. The impact of high-resolution whole genome sequencing on the detection and resolution of foodborne outbreaks.**
Higher resolution WGS allows more outbreaks to be identified and resolved more rapidly, with fewer overall cases. Pathogen population sizes are shown in blue, and the number of cases of foodborne illness are in red. With low-resolution typing, a single large outbreak is detected (back panel), which may correspond to a seasonal increase in cases or a prolonged, multi-year outbreak. This type of low-resolution typing tends to group epidemiologically-unlinked cases together, resulting in an inaccurate understanding of the underlying epidemiology, and less effective interventions. By contrast, high resolution genome sequence-based typing enables better assignment of cases to discrete smaller outbreaks with a common epidemiological origin (front four panels). The improved assignment enables outbreaks to be resolved more rapidly, leading to fewer cases associated with each outbreak, and a reduction in the overall pathogen population. This example is motivated by the US Centers for Disease Control description of the analysis of outbreaks using Advanced Molecular Detection, comparing low resolution methods with WGS[131].

## Box 1

**Questions that genomics now allows food businesses to ask:**

- Is there a predominant pathogen or strain in a facility? Is this strain spilling over to their food products?
- Are there other microorganisms supporting the pathogens? Do these likewise find their way onto their foods?
- How are the pathogens surviving disinfection measures? Are the 'supporting' microorganisms likewise being selected for with these measures?
- Can the possible spread of microorganisms be tracked within individual facilities (and potentially their foods)?
- Can strains of pathogens be identified that might be of a higher risk than other strains (for example, genetic markers to inform risk, recognising that not all strains of a particular pathogen are equally hazardous)?
- Can the baseline microbial composition of a food process, setting, or ingredient be determined using metagenomics?
- Can the impact of interventions imposed in food businesses be monitored by determining if there is a change in the distribution or composition of microbiota?[79]

**Actions that can now be considered with the depth of knowledge learned from genomics:**

- Modification of chemical disinfection plans based on the observation of susceptible or tolerant microorganisms
- Reconsideration of hygiene best practices related to how equipment and ingredients are handled, placed, and/or move throughout a facility based on evidence of microorganisms residing or spreading in a facility (that is, facility-specific pathways)
- Redesign or use of innovative materials in food non-contact and contact surfaces that may mitigate the persistence or transmission of key microorganisms in these at-risk pathways
- Consideration of novel environmental or food biocontrol strategies (for example, the application of bacteriophages, or the introduction of other microorganisms to change the microbial community structure).
- Implementation of non-chemical interventions at identified critical control points (for example, ultraviolet light or sonication)
- Use of metagenomic information as the guide point to recover key taxa from related samples for further characterisation and study
- Informed and proportional action to urgent levels of risk (for example, observation of strain with characteristics of high concern)