EcoSal Plus
Cellular and Molecular Biology of
E. coli, Salmonella, and the Enterobacteriaceae

DOMAIN 6 EVOLUTION AND GENOMICS

Salmonella Genomics in Public Health and Food Safety

ERIC W. BROWN,[a] REBECCA BELL,[a] GUODONG ZHANG,[a] RUTH TIMME,[a] JIE ZHENG,[a] THOMAS S. HAMMACK,[a] AND MARC W. ALLARD[a]

[a]Center for Food Safety and Applied Nutrition, U.S. Food and Drug Administration, College Park, Maryland, USA

Received: 18 February 2021
Accepted: 16 March 2021
Published: 14 June 2021
Editor: David Rasko, University of Maryland School of Medicine, Baltimore, Maryland, USA
Citation: Brown EW, Bell R, Zhang G, Timme R, Zheng J, Hammack TS, Allard MW. 2021. *Salmonella* genomics in public health and food safety. EcoSal Plus 9:eESP-0008-2020. https://doi.org/10.1128/ecosalplus.ESP-0008-2020.
Address correspondence to Eric W. Brown, eric.brown@fda.hhs.gov.

ABSTRACT The species *Salmonella enterica* comprises over 2,600 serovars, many of which are known to be intracellular pathogens of mammals, birds, and reptiles. It is now apparent that *Salmonella* is a highly adapted environmental microbe and can readily persist in a number of environmental niches, including water, soil, and various plant (including produce) species. Much of what is known about the evolution and diversity of nontyphoidal *Salmonella* serovars (NTS) in the environment is the result of the rise of the genomics era in enteric microbiology. There are over 340,000 *Salmonella* genomes available in public databases. This extraordinary breadth of genomic diversity now available for the species, coupled with widespread availability and affordability of whole-genome sequencing (WGS) instrumentation, has transformed the way in which we detect, differentiate, and characterize *Salmonella enterica* strains in a timely way. Not only have WGS data afforded a detailed and global examination of the molecular epidemiological movement of *Salmonella* from diverse environmental reservoirs into human and animal hosts, but they have also allowed considerable consolidation of the diagnostic effort required to test for various phenotypes important to the characterization of *Salmonella*. For example, drug resistance, serovar, virulence determinants, and other genome-based attributes can all be discerned using a genome sequence. Finally, genomic analysis, in conjunction with functional and phenotypic approaches, is beginning to provide new insights into the precise adaptive changes that permit persistence of NTS in so many diverse and challenging environmental niches.

KEYWORDS whole-genome sequencing, foodborne pathogen isolates, data sharing, ontology, *Salmonella* detection, genomic impact, epidemiology, microbiology, adaptive change, nontyphoidal salmonellae

The genus *Salmonella* represents a group of rod-shaped, Gram-negative, facultatively anaerobic bacteria in the family *Enterobacteriaceae*. The genus is composed of two species, *Salmonella bongori* (originally called *Salmonella* subspecies V) and *Salmonella enterica* (1). The latter species contains a great number (>2,600) of serologically distinct variants, or serovars, known to persist predominantly in mammals, birds, and reptiles. Taxonomically, the

species *Salmonella enterica* is partitioned into seven subspecies, including I (*Salmonella enterica* subsp. *enterica*), II (*S. enterica* subsp. *salamae*), IIIa (*S. enterica* subsp. *arizonae*), IIIb (*S. enterica* subsp. *diarizonae*), IV (*S. enterica* subsp. *indica*), VI (*S. enterica* subsp. *houtenae*), and VII (2, 3). However, nearly all (>99%) of those serovars associated with clinical and veterinary illness are derived from subspecies I (3).

Salmonella is a highly fit environmental microbe, however, and enjoys a distribution that is largely ubiquitous across geographic and biologic reservoirs. The pathogen is well known to persist frequently outside animals in other natural environments, including fresh and marine surface waters, soil, and dust, and as epiphytes on and inside plant materials (4). Since its divergence from *Escherichia coli* more than 100 million years ago, *Salmonella* has undergone widespread evolutionary diversification and niche-specific adaptive change through the acquisition of numerous novel genomic changes, many of which have been acquired as a result of horizontal gene transfer (HGT) and other reticulate evolutionary forces (3).

Salmonella enterica is responsible for 1.4 million cases of foodborne salmonellosis in the United States annually, making it the number one causative agent of bacterial foodborne illnesses. Infection can occur after eating undercooked meat, poultry, and eggs, as well as fresh-cut produce that is readily consumed raw and has been contaminated with *Salmonella* (5). In recent years, more *Salmonella*-related outbreaks have occurred in the United States associated with the consumption of produce than animal-based food commodities. Recent outbreaks of this nature include a massive *Salmonella enterica* serovar Saintpaul outbreak associated with tomatoes, jalapeños, and serrano peppers that sickened over 1,400 individuals in 2007 (6) as well as four separate events involving Maradol papaya in 2017 that included at least 8 different serovars and caused more than 250 known illnesses, including two deaths (7). Additionally, a 2014 cucumber-related outbreak associated with *Salmonella enterica* serovar Newport and *Salmonella enterica* serovar Javiana caused more than 275 reported illnesses and one death (8). These events underscore the notion that numerous serovars of *Salmonella* may have migrated successfully into previously naive niches (i.e., produce-growing niches) and point to a role for the ongoing genetic and epigenetic adaptation of *Salmonella* into food and feed environments. WGS source tracking is allowing us to see novel evidence of exposure of fresh-cut produce to foodborne pathogens in animal reservoirs and through contaminated water and soil amendments.

The importance of genetic and genomic data in understanding the ecological and evolutionary adaptations that drive persistence of foodborne pathogens such as *Salmonella* cannot be overstated. Application of whole-genome sequencing (WGS) for the characterization of *Salmonella* has provided extraordinary insight into epidemiology, biology, evolution, and population structure of *Salmonella* over the past decade (9), including permitting the detailed organization of *Salmonella enterica* into a phylogenetic hierarchy that largely recapitulates the species' natural population structure (10). The WGS of *Salmonella* strains now regularly provides a highly reliable and predictive means to ascribe various phenotypic and diagnostic traits to a specific isolate by means of one analytical workflow—that is, the sequence itself (9). Historically, during surveillance and diagnostics, important phenotypic tests such as serotyping, antimicrobial resistance (AMR) testing, and phage typing (PT) were cumbersome and expensive, but genetic and genomic alternatives are already developed that can provide comparable and quite reliable results simply by analyzing the genomic sequence of a particular *Salmonella* isolate.

The same genomic approaches to phenotypic discovery are also beginning to yield clues regarding the emergence of unique and strongly selected adaptations in *Salmonella*, some of which have transformed *Salmonella* with novel virulence traits and capabilities in its host as well as permitting it to endure in environmentally harsh and unexpected environments and to persist in the face of otherwise lethal assaults from antimicrobials, oxidative agents, and other sanitizers in industrial and health care settings. This has become particularly true in the food production and processing industry, where recent genetic adaptations observed in *Salmonella* may subvert certain controls and preventions and contribute to foodborne illness of public health concern. One example of this includes a recent outbreak strain of *Salmonella enterica* serovar Bareilly isolated from tuna, which was found to harbor a genomic island containing a previously undescribed arsenic resistance operon (11, 12).

S. Bareilly isolates that carry the novel genetic island are significantly more resistance to arsenic than those that do not. Arsenic is a toxic metalloid ubiquitous in the natural environment and often found at higher levels in fish and seafood, which absorb arsenic from surrounding waters and other lower trophic species that are part of their diet. With the novel arsenic resistance element integrated into the chromosome, *Salmonella* Bareilly is now likely to better survive and persist in tuna. This highlights the potential impact of evolutionary change in *Salmonella*, whereby a selective advantage was conferred for survival, persistence, and even growth within human food matrices. Moreover, by pinpointing the underlying differences in phenotype among closely related salmonellae, we are better able to predict *Salmonella*'s environmental responses and subsequently may be able to provide specific and targeted mitigation strategies, recently termed "precision food safety" approaches, for controlling nontyphoidal *Salmonella* serovars (NTS) from further contaminating the food and feed supply (13).

Continued understanding of *Salmonella* fitness, diversification, virulence, and survivability will be essential to our ability to manage, treat, and prevent its contamination of humans and human-associated upstream niches. While much remains to be discerned regarding the ever-changing face of *Salmonella* in its natural habitat, it is now clear that *Salmonella* microbiology has already been advanced tremendously from the information provided by genomic tools and in particular the integration of WGS into traditional microbiological areas of the study of *Salmonella* (9). Here, it is our intent to present the impact of genomics in several key areas of *Salmonella* microbiology, including its phylogenetic partitioning, adaptive changes, environmental persistence, host specificity, virulence, and continued burden on food and feed safety.

TRANSFORMATION OF *SALMONELLA* SEROLOGY AND SEROTYPING THROUGH WGS

Serotyping has long been a key classification for *Salmonella*. The White-Kauffmann-Le Minor (WKL) scheme is the international standard for the designation of *Salmonella* serotypes and is based on serological characterization of the O and H antigens (14). Serotyping remains a critical part of public health investigation into *Salmonella* outbreaks and contamination events. The WKL approach uses the agglutination reaction with antisera against O and H antigenic variants (15). Currently, 46 O antigens and 114 H antigens exist among the known salmonellae, and from the possible combinations, about 2,600 have been isolated and named "serovars" of the species (16). Although it was a staple of *Salmonella* serological characterization for decades, the method rapidly is being replaced by the widespread availability of genomic approaches to serotyping of the species. Moreover, the schema itself (i) relies heavily on antisera that are now available from a dearth of sources, (ii) depends on the availability of highly trained experts to interpret often complex and sometimes ambiguous results, and (iii) requires sufficient time to complete the reactions while remaining in step with an investigation or traceback event. In order to overcome these shortfalls, various molecular testing-based strategies were developed in the early 2000s that used gene sequences as surrogate antigenic markers, including multiple PCR-based approaches (17–19) and a liquid suspension DNA hybridization approach based on X-map (e.g., Luminex) technology (20, 21).

More recently, genomic approaches to *Salmonella* serotyping have risen in popularity and utility. An abundance of genomic data from the species is now widely available for comparison of WGS-derived phenotypes, including serovar status (12). Several of these methods now hold great potential for deployment and use by public health and food safety authorities, including the U.S. FDA, the CDC, Public Health England, and Public Health Canada, to name a few.

An early example of predicting serovar was a WGS-based solution called SeqSero (22). This genomic dashboard tool, launched in 2015, relies solely on the upload of a draft *Salmonella* genome to the SeqSero web tool (http://www.denglab.info/SeqSero) and subsequent receipt of the genomic serotype of the *Salmonella* strain based on the rapid genomic comparison of O and H antigen-encoding genes. Very recently, an enhanced functional update and new version of the software, SeqSero2 (https://github.com/denglab/SeqSero2/branches) (16), was released that is 50× faster than its predecessor with a serovar call accuracy of 98% when evaluated against several large *Salmonella* WGS-based public databases, including National Antimicrobial Resistance Monitoring System (NARMS) and the GenomeTrakr National databases. An additional

genome-based typing tool was also recently developed that targets the serovar-specific spacer regions of the two CRISPR loci (i.e., CRISPR 1 and CRISPR 2) in *Salmonella*. The method, called CRISPR-SeroSeq, provides a multiplexed partial genome-sequencing scheme that can detect and characterize multiple *Salmonella* serovars from a single analysis. This direct serotyping approach can pick up underrepresented serovars in a single environmental sample as low as 0.01% (23), making it particularly attractive for direct serotyping of poultry and poultry processing related swab surveillance samples, where numerous serovars can populate an individual test (24). The software improvements are supported by large curated databases of more known serovar sequences to compare against new uploaded unknowns. This strategy of building curated known reference sequence databases in specific bioprojects will allow investigators to build more genomic tools that predict phenotype from genotype for any genes that are fully characterized and linked to specific phenotypes.

RECOMBINATION, RETICULATE EVOLUTION, AND THE IMPACT OF HGT IN *SALMONELLA*

Salmonella evolves both vertically and horizontally. By vertical evolution, we mean ancestor-to-descendant evolution based on passing genetic variants on to the next generation through inherited genetic changes. In contrast, horizontal evolution refers to genetic variation that was not inherited but rather exchanged between organisms (HGT). Plasmids and phages can be transferred through conjugation to other compatible bacteria. Thus, bacterial isolates that independently acquired horizontal elements (phages and plasmids) may appear closely related when they are distantly related and share only the HGT elements. HGT may confound vertical evolution; that is why these potential elements are initially filtered out when building phylogenetic trees to document closely related isolates that share an ancestor. Variation from recombination also may confound vertical evolution and thus also may be filtered out to more accurately measure the vertical evolutionary signal, as the focus for WGS in food safety is to identify the most closely related isolates and to cluster them for follow-up investigation. Phylogenetic comparisons between different regions of the *Salmonella* genome cemented the key role of HGT in the genetic and evolutionary diversification of *S. enterica* subspecies, serovars, and individual pathogenic clones.

Numerous genetic and genomic studies on population structure and chromosome organization in *Salmonella* have repeatedly demonstrated that HGT has driven the emergence of highly adapted strains of *Salmonella* in the environment, on the farm, and in the food supply (25). While numerous mechanisms and pressures can drive HGT among the salmonellae, the hypermutable phenotype has underscored numerous examples of reticulate evolution on the *Salmonella* chromosome. Methyl-directed mismatch repair (MMR) defects, leading to the "mutator," or hypermutable, phenotype, are found in more than 1% of the isolates within naturally occurring populations of *S. enterica* and at even greater frequencies in the food supply, where oxidative and other antimicrobial stressors often abound (26). Up to 73% of the MMR defects found in feral settings are the result of lesions within *mutS*, resulting in increased nucleotide substitution rates, enhanced DNA transposition, and, perhaps most importantly, a relaxation of the internal barriers that normally restrict homeologous recombination following the horizontal acquisition of foreign DNA (27–30). Phylogenetic analyses were conducted of the *Salmonella* reference (SAR) collections (i.e., SARA, SARB, and SARC), which were largely considered to represent the extent of genetic variability within the species but are now known to represent a subset of that diversity (2, 31, 32). Work first suggested by single-gene studies revealed striking levels of phylogenetic discordance between trees derived from *mutS* alleles and whole-chromosome trees of the same strains based on multilocus enzyme electrophoresis (MLEE) analysis (28, 29). These findings support the notion that HGT helped forge current relationships among *Salmonella* and other enteric pathogens in this region and throughout numerous other locales in the *Salmonella* chromosome. Indeed, as evidenced in other studies employing genomic scanning approaches, such as WGS, microarray analysis, and multilocus sequence typing (MLST), the substantial impact that HGT has had on structuring the chromosome of *Salmonella enterica* is indisputable (33, 34). Estimates based on such studies indicate that more than one-quarter of the *Salmonella* genome may have been brought about by HGT and reticulate evolutionary forces (33), although this number is likely conservative, based on current views.

It is now evident that HGT has played a key role in structuring many other regions of the *Salmonella* chromosome as well. Notably, *Salmonella* pathogenicity islands (SPI) were likely acquired through HGT (35–37).

For example, SPI-1, comprising the genes encoding a type III secretion system, was probably acquired early in *Salmonella* evolution (38), yet several *inv-spa* alleles seem to have converged horizontally more recently between *S. enterica* groups IV and VII (28, 39). Additionally, genes comprising the *inv-spa* gene cluster appear to have undergone extensive allelic shuffling among the group I salmonellae (40). Also, type 1 pilin genes, encoding fimbrial adhesins, retain unusually low GC content and obscured phylogenies relative to other *fim* genes (41). Other studies focusing on housekeeping genes in *Salmonella* have reported evolutionary histories for these genes that are strikingly decoupled from *S. enterica* strain history (28, 42–44).

In comparing across and within subspecies of *Salmonella*, a recombination "gradient" has been noted wherein lateral DNA transfer appears to be inversely correlated with genetic variability among strains. Subsequently, a genetic boundary may exist that tolerates free recombination of DNA sequences within a framework delimited by sequence variation and niche diversity of individual *Salmonella* strains. This has been documented previously through the observation of intragenic (or patch-like) recombination events among more diverged *Salmonella* subspecies and assortative (whole-allele) recombination events, responsible for extensive whole-allele reassortment, among more genetically homogeneous populations of group I *Salmonella* strains (28, 29). It is notable that the latter strains all share a niche primarily found in warm-blooded mammals (2).

Largely due to the recent availability of genomic data, our understanding in reconstructing the HGT of important features, including those involved in virulence, drug resistance, and other adaptations that foster an enhanced fitness for *Salmonella* persistence in the farm-to-fork continuum is expanding at a pace that we could not have foreseen at the turn of the millennium (12). The extent and effects of recombination have now been noted for both typhoid-causing salmonellae and NTS in important clinical and environmental niches and across both core genome regions and the *Salmonella* mobilome (45–47), including AMR determinants, a suite of chemical resistance operons (48), and numerous *Salmonella* genomic island (SGI) regions (49).

It is important to recall, however, that reticulate evolutionary pressures do not subside once selectively advantageous traits are gained. Rather, HGT likely continues between the most closely related salmonellae (29). To this end, phylogenetic studies also have revealed the often-underappreciated role of HGT and recombination in the homogenization of allele structure within closely related populations of *S. enterica* (40), as well as a genetically panmictic structure for restriction-modification (R-M) genes among group I salmonellae. The latter finding, noting open exchange of R-M (i.e., *hsd*) alleles, constitutes phylogenetic evidence of the compatibility of *S. enterica* subspecies I R-M complexes, likely accounting for the documented successful HGT of entire gene sequences among closely (e.g., intrasubspecies) related strains, as DNA exchange between strains that share or recently shared common R-M alleles would not be subject to substantial restriction (50).

The phylogenetic data demonstrate that HGT has been a frequent and regular phenomenon among the salmonellae, and this may explain, in part, why *Salmonella* possesses such unique evolutionary histories for numerous genes, operons, and islands within its chromosome. With the complete genome sequences of literally hundreds of thousands of *Salmonella* isolates now available, such refined genomic and phylogenetic analyses should aid in arriving at the final verdict on the impact of recombination on *Salmonella* chromosome structure. Certainly, a greater recognition of precisely how HGT and homeologous and homologous recombination have forged the genomes of *Salmonella* pathogens should enhance the accuracy of our risk assessment of these pathogens as well as providing solutions for better detection and characterization of this devastating foodborne disease-causing agent.

THE IMPACT OF GENOMICS ON THE DETECTION AND CHARACTERIZATION OF AMR IN *SALMONELLA*

The AMR *S. enterica* strains are a significant source of enteric foodborne illness and a public health concern (51). In particular, the presence of AMR genes associated with NTS in the food and feed supply presents an important challenge to controlling human and veterinary illness associated with the consumption of contaminated food and feed commodities (52, 53). The AMR genes are found across large numbers of *Salmonella* genomes, with comparable averages for *Salmonella* associated with produce and animal food products of 72% and 74%, respectively (54). This high prevalence of AMR genes has been

attributed specifically to an elevated presence of aminoglycoside resistance genes along with tetracycline and sulfonamide genes, in most food isolates (54).

An automated AMR typing approach that relies on genomic data has been developed to meet the challenge of AMR characterization in *Salmonella* associated with the food supply. Consistent with other WGS tools now available for predicting phenotype from genotype, AMR genotypes can now be readily targeted from WGS and the resultant phenotype predicted (55). The importance of these *Salmonella* AMR tests cannot be overstated, as tens of thousands are conducted each year by the federal government and their public health partners, primarily through the NARMS program (https://www.fda.gov/animalveterinary/safetyhealth/antimicrobialresistance/nationalantimicrobialresistancemonitoringsystem/), a consortium of state and federal agencies that monitor AMR in meats and clinically obtained isolates of *Salmonella* and other enteric bacterial foodborne pathogens. To this end, The National Center for Biotechnology Information (NCBI), in collaboration with experts on AMR in *Salmonella*, provide online WGS tools for predicting AMR genotypes (https://www.ncbi.nlm.nih.gov/pathogens/). The NCBI has produced AMRFinder Plus, a tool that identifies AMR genes using a high-quality curated AMR gene reference database (55). The Bacterial AMR Reference Gene Database consists of up-to-date gene nomenclature, a set of hidden Markov models (HMMs), and a curated protein family hierarchy. Predictive assessments of AMRFinderPlus revealed genotype-to-phenotype concordance for *Salmonella* AMR profiles of more than 98%, pointing to a margin of error of less than 2% and providing evidence that AMRFinderPlus is a highly accurate WGS-based AMR gene detection system (55) for each draft genome uploaded and released through the WGS pathogen portal at NCBI. The NCBI Pathogen Detection web site also provides investigators with detailed guidance on how to upload corresponding phenotypic antibiogram metadata with their draft genomic data so that improved calling of AMR genotypes can ensue.

However, other questions surrounding AMR salmonellae remain, including those concerning the geographic distribution of AMR genotypes (see the FDA CVM Resistome Tracker), the genomic diversity present in known AMR genes, and how much differential expression and phenotypic variation is present in known genes and allelic AMR variants. Traditional antibiogram testing is generally limited to a smaller subset of antibiotics, whereas a genomic screen interrogates all known AMR genes supported in the database. A search of the NCBI Pathogen Detection website identifies a *Salmonella* genome with up to 29 AMR genes present and over 200 isolates with up to 20 AMR genes present. The AMRFinderPlus database has been expanded to predict genes related to stress and virulence. These publicly available reference databases allow any investigator to create rapid PCR or sequencing panel screens for specific genes of particular interest to stakeholders. This is a strategy to reduce the costs of genotype-to-phenotype predictions by targeting gene panels for food safety, industry surveillance, and clinical diagnostics.

Salmonella found in the environment, food, and animals: adaptive fitness and persistence. Two types of clinical manifestations were associated with *Salmonella* serovars, including invasive, life-threatening systemic disease, referred to as typhoid fever, and self-limited gastroenteritis caused by NTS (56–58) found in foods, animals, and the environment. However, 5% of individuals infected with NTS develop bacteremia (59), and disease manifestations are substantially different between different serovars (60–62). Multiple NTS were found in asymptomatic food-producing livestock, including poultry, sheep, cattle, and swine (63–65), indicating that *Salmonella* persistence and carriage in livestock are very common, possibly since *E. coli* and *Salmonella* diverged from a common ancestor. Multiple sets of *Salmonella* genes are involved in prolonged infection and persistence (66). Distinct sets of fimbriae contribute to the intestinal persistence and colonization in different animal species (67, 68). Besides fimbrial adhesins, other adhesins, including the autotransporter adhesins MisL, SadA, and ShdA, the type I secretion system-secreted adhesins SiiE and BapA, and curli biogenesis (*csg*) adhesins, were found to play a role in colonization and persistence in mouse gastrointestinal tract (69–73). Recent works suggested that biofilm formation is involved in *Salmonella* gallbladder persistence. *Salmonella* gallbladder colonization triggers upregulation of the O-antigen capsule-encoding operon (*yihU-yshA* and *yihV-yihW*) in an *agfD*-independent manner, which is specifically required for biofilm formation on cholesterol gallstones (74). Iron is an essential nutrient for human and animals. A major host defense against infection is nutritional immunity, e.g., via sequestration of metals, including iron (75), to prevent pathogen

growth (76). The siderophore ABC transporter FepBDGC is responsible for primary ferric ion import in *Salmonella*. It was shown that the *fep* system together with the ferric-iron-binding siderophores enterobactin and salmochelin is required for persistent *Salmonella* infection in mice (77). Comparative genomic analysis of serial isolates associated with long-term epidemics revealed mutation rates from 1.9×10^{-7} substitution site^{-1} year^{-1} to 1.49×10^{-6} substitution site^{-1} year^{-1} in the core genome of *Salmonella enterica* serovar Typhimurium (78–80). Multiple nonsynonymous single nucleotide polymorphisms (SNPs) were found in global virulence regulators, including DksA, RpoS, HilD, MelR, and BarA, and metabolic pathways, providing an adaptative advantage during persistence in the host (79, 81). Meanwhile, the mean genome-wide rate of nonsynonymous to synonymous substitutions (*dN*/*dS*) was less than 1 during the short-term evolution of *Salmonella*, indicating that the underlying substitution rate is subject to purifying selection (78). In contrast to the relatively stable core genome, considerable variation in composition of mobile genetic elements, including prophages and plasmids, was identified within the same clone in the course of an epidemic (78–81). All these changes contributed to clinically relevant differences in phenotype and virulence, further emphasizing the critical importance of integrated genotypic data sets in understanding of biological variability in *Salmonella* epidemiology.

Host adaptation. Primarily, when infecting its host, *Salmonella* exists in the intestinal tract as a gastrointestinal pathogen with limited duration and disease progression. However, some serovars have adapted to cause an invasive disseminated disease. During this process, these serovars also have lost the ability to infect a broad range of hosts, becoming much more host adaptive (HA) or host restrictive (HR). Host-adaptive serovars tend to have one or two main animals that they naturally infect but are capable of infecting other hosts given the opportunity. Host-restricted serovars have one main host and rarely or never naturally infect a different host. Well-known examples of HR serovars include *Salmonella enterica* serovar Typhi, *Salmonella enterica* serovar Paratyphi A, *Salmonella enterica* serovar Gallinarum/Pullorum, and *Salmonella enterica* serovar Abortusovis (Table 1) (82). The disease caused by these serovars in their natural host is typically characterized by fever and septicemia, with very little or no gastrointestinal disease. Similarly, HA serovars, such as *Salmonella enterica* serovar Choleraesuis and *Salmonella enterica* serovar Dublin, cause severe systemic disease in their natural host and humans which is also characterized by fever and septicemia with little diarrheal symptoms (82). This is in contrast to most *Salmonella* serotypes, which exhibit a broad or unrestricted host range and cause severe gastroenteritis. In recent years, though, there has been an emergence of HA in certain clones of some unrestricted-host-range serotypes which cause invasive disease mainly in immunocompromised patients (Table 1) (83, 84).

The precise mechanisms of how HA and HR evolved are not well understood but seem to encompass three major steps: the gain of genetic information, the loss of genetic information (genome reduction), and the increase of pseudogenes within the genome. During the evolution of HR, *S.* Typhi gained several pathogenicity islands (SPI-7, -15, -17, and -18), which includes the Vi capsular antigen, which allows it to avoid being killed by host complement and prevents phagocytosis (85). Conversely, *S.* Typhi and other HA/HR serovars tend to be auxotrophic for specific amino acids and vitamins (82) and have lost large numbers of genes in anaerobic metabolic pathways necessary for growth in the inflamed gut (86). Also, many have a reduced number of virulence factors commonly found in broad-host-range serovars (87). Along with the loss of genes, all HA and HR serovars have large numbers of pseudogenes compared to unrestricted-host-range serovars (85–88). Many of the genes seen to be degraded in HR and HA serovars tend to be involved in motility or chemotaxis, to encode type III secretion effectors, or to encode structures involved in attachment to host cells, such as fimbriae and other adhesins (85). In addition, the allelic variation found within the HA and HR group is reduced compared to host generalists, suggesting their more recent emergence in specialized host species (87). Allelic variation also likely has played a role in the host tropism in salmonellae, which may be a first step toward host restriction. For example, Yue and colleagues studied allelic diversity of several fimbrial adhesins and found patterns of alleles associated with different host types (89–91).

Some clones of *Salmonella* are adapting to a more restricted host range. Two prominent examples of this include invasive *S.* Typhimurium ST313 and invasive *Salmonella enterica* serovar Enteritidis, which have emerged in sub-Saharan Africa (83, 84). Genetic analysis

TABLE 1 Characteristics of host-restricted, host-adapted, and unrestricted invasive *Salmonella* serovars[a]

Serovar	Host	Auxotrophic requirement[b]	Disease	Missing genes/ pseudogenes
Host-restricted serovars[c]				
Typhi	Humans, chimpanzees	Tryptophan, cobalamin (vitamin B_{12})	Typhoid fever	*sseI*, *gtgE*, *sopA*, *sseK2*, *ratB*, *sadA*, *stfH*, *sopD2*, *gtgA*, *ompD*, *steB*, *sopE2*, *shdA*, *sinH*, *bapA*, *avrA*, *misL*, *cigR*, *sseK1*
Paratyphi A	Humans	Cystine, arginine, cobalamin (vitamin B_{12})	Enteric fever	
Gallinarum/Pullorum	Poultry	Cystine, leucine, aspartic acid, thiamine, cobalamin (vitamin B_{12})	Fowl typhoid, Pullorum disease	*sseI*, *gtgE*, *sopA*, *sseK2*, *ratB*, *sadA*, *sirP*, *sifB*, *fliC*, *sspH2*
Abortusovis	Ovines	Cystine, nicotinic acid (vitamin B_{12}), thiamine	Abortions, newborn mortality	
Typhisuis	Swine	Cystine	Chronic paratyphoid	
Abortusequi	Equines		Abortions, newborn mortality	
Host-adapted serovars[d]				
Choleraesuis	Swine, humans		Swine paratyphoid	*sseK2*, *sspH2*, *shdA*, *avrA*, *sadA*, *sopE*
Dublin	Bovines, humans, ovines	Nicotinic acid (vitamin B_3)	enteric and invasive disease, abortions	*srfN*, *fliC*, *sseK2*, *shdA*, *bapA*, *sadA*
Invasive NTS strain or serovar				
Typhimurium ST313	Humans	Cobalamin (vitamin B_{12})	Invasive disease	*sseI*, *shdA*, *siiE*, *sspH2*, *sadA*
Enteritidis	Humans	Cobalamin (vitamin B_{12})	Invasive disease	*sseI*, *sspH2*, *shdA*, *sadA*, *siiE*, *fliC*, *sseK2*, *sinH*

[a]Data are from references 82 and 87.
[b]Auxotrophy is a common characteristic in host-restricted and host-adapted serovars.
[c]Host-restricted disease tends to be systemic with little or no gastroenteritis.
[d]Host-adapted disease does not produce severe enteritis and is followed by systemic dissemination.

of *S.* Typhimurium ST313 has shown signatures of adaptation. For example, a number of pseudogenes have been identified in ST313, compared to other *S.* Typhimurium strains, and are similar to the pseudogenes found in *S. Typhi* and *S.* Paratyphi A, including *ratB*, *ttdA*, and *sseI* (79, 80, 84). Also similar to *S. Typhi* and *S.* Paratyphi A, *S. Typhi*murium ST313 shows a loss in metabolic capacity (80). Interestingly, ST313 has a limited ability to form biofilms, which may lead to reduced fitness to survive outside the human host (92); this in turn may explain the lack of an identified zoonotic reservoir and evidence for human-to-human spread of this pathogen (93). In parallel, analysis of *S.* Enteritidis isolates linked to invasive disease have identified two clades circulating in sub-Saharan Africa, one localized to western areas and one to central/eastern areas (83). These clades show signatures of adaptation similar to those of other host-adapted or restricted serovars, namely, multiple pseudogenes in metabolic pathways along with an accumulation of nonsynonymous SNPs in membrane protein genes (83). It should also be noted these invasive serovars are multidrug resistant (MDR) and have most likely gained the genes for drug resistance since the divergence with their most recent common ancestor (83, 84).

Examples from poultry-adapted serovars. (i) *Salmonella enterica* serovar Enteritidis. *Salmonella* Enteritidis is a host-promiscuous serovar that is predominantly

associated with gastroenteritis. Separating from *S.* Dublin, rather being a single clade itself, *S.* Enteritidis is more structurally complex. The WGS phylogeny of *S.* Enteritidis suggests the presence of at least four clades with three epidemic clades and one clade from which *S.* Gallinarum/*S.* Pullorum complex directly evolved (83, 88). Among the three epidemic clades, one is the classic or global epidemic clade (83, 94) including the most commonly isolated MLSTs and PTs associated with enterocolitis in human. The other two *S. Enteritidis* clades emerged from Africa and are strongly associated with multidrug resistance and invasive disease (83). Given the genetic and phenotypic heterogeneity within *S.* Enteritidis clades, multiple signatures of differential host adaptation are observed in the context of genome evolution. Analysis of the accessory genome, consisting of 14,015 predicted genes, showed the acquisition of a novel prophage region closely related to *Enterobacter* phage P88 and an enlarged virulence plasmid with acquired MDR genes in both African clades. Additionally, strains from both African lineages harbored pseudogenes which are concentrated in common metabolic pathways, as observed in other host-restricted invasive *Salmonella* serovars. Some of the overlaps are astonishing, including reduced metabolic activity in cobalamin and propanediol utilization and also ornithine decarboxylase activity, indicative of the role of gene loss/pseudogene formation in the adaptation from a gut to a systemic lifestyle.

(ii) *Salmonella enterica* serovar Heidelberg. *Salmonella* Heidelberg is primarily a poultry-adapted serovar of *Salmonella* that can colonize and infect multiple hosts. Infections with *S.* Heidelberg are more likely to be invasive and associated with greater risk for severe disease than other serovars (95). The pan-genome (*pan*, from the Greek word $\pi\alpha\nu$, meaning "whole") includes a core genome containing genes ubiquitous in all strains and an accessory genome composed of genes absent from one or more strains and genes that are unique to each strain. The pan-genome of *S. enterica* subspecies I is predicted to have 42 or 43 different fimbrial gene clusters (FGCs), which have been implicated in host colonization and adaptation. With acquisition and deletion of FGCs, the evolutionary pathway has led to four clades of *S. enterica* subspecies I. *S.* Heidelberg resides in clade 1b with two other serovars, *Salmonella enterica* serovar Virchow and *Salmonella enterica* serovar Hadar, carrying the highest numbers of FGCs among all other serovars (95). Accumulation of different FGCs may improve the efficiency of specific host colonization and broaden the host range. Comparative genomics of outbreak-related bovine strains and isolates from other resources have predicted that the gain of Saf fimbrial genes may have contributed to the increased bovine colonization (96). Moreover, *S.* Heidelberg isolates presented divergent MDR genes with strong phylogeographic signature (97) and displayed a broad diversity of phage-related genes, with some unique to different poultry farms (98). Phage and plasmid HGT may facilitate the dissemination of MDR (95, 99, 100) and contribute to the fitness of *S.* Heidelberg in different poultry farm environments (98).

(iii) *Salmonella enterica* serovar Kentucky. *S.* Kentucky is among the *S. enterica* serovars most frequently isolated from poultry in the United States (101, 102) and has been increasingly isolated from dairy cattle as well (103). However, it is less commonly identified as a source of human salmonellosis than other serovars commonly detected in poultry, such as *S.* Enteritidis and *S.* Heidelberg. Although *S.* Kentucky contains five pathogenicity islands (SPI-1 to -5), like other serovars in *S. enterica*, the lack of full-length SPI-2-associated genes and fimbrial genes (104, 105) might compromise its virulence in humans. The rise of *S.* Kentucky as the dominant serovar in poultry may be due to the acid response phenotype (106) and a metabolic advantage conferred by the acquisition of the ColV plasmid for scavenging scarce energy sources available in the chicken cecum (107). Phylogenetic analysis indicated that *S.* Kentucky is polyphyletic (57, 108), with two highly divergent ST complexes. ST152 and ST198 are the most frequently isolated *S.* Kentucky sequence types in each ST complex globally. ST198 is reported to be MDR and causes gastroenteritis in humans (109), while ST152 is rarely associated with human disease (108). Comparative genomics between ST152 and ST198 isolates found significant differences in gene content and core genome nucleotide sequence divergence. The roles of several genomic elements in ST198, such as a sialic acid transport region, inositol catabolism, and a homolog of the Typhi colonization factor, need to be further evaluated for host-associated colonization (108). It also is noted that MDR is mostly conferred by plasmids in poultry-associated *S.* Kentucky ST152 isolates (102, 108), while it is associated with the acquisition of

Salmonella genomic island 1 (SGI1), plasmids, and mutations in the core genome of ST198 isolates (108–110).

Patterns of *Salmonella* serovar evolution in egg and poultry production. Egg- and poultry-associated products have been frequently implicated in foodborne gastroenteritis caused by *Salmonella* serovars (111). Predominant *Salmonella* serovars in commercial poultry have undergone significant shifts over the last several decades. Several bacterial factors could contribute to such shifts in *Salmonella* populations in poultry, including competitive exclusion and genetic factors that facilitate *Salmonella* colonization in poultry. The prevalence of *Salmonella* serovars among poultry can be dated back to the early 1900s. *S.* Gallinarum biovars Pullorum and Gallinarum caused pullorum disease and fowl typhoid in poultry, respectively, posing a serious economic threat to the poultry industry at that time (112). In the 1980s, *S.* Enteritidis O9,12:g,m emerged as a major public health problem in Europe and the Americas (113). *S.* Enteritidis did not spread to domestic fowl until much later after the initial introduction into poultry flocks through its rodent animal reservoir (114–116). *S.* Enteritidis shared the immunodominant O antigen (O9) on the cell surface with Gallinarum, which may have contributed to the exclusion of *S.* Enteritidis earlier (117, 118). One of the reasons for the increased spread of *S.* Enteritidis could be the eradication of *S.* Gallinarum, which may have opened an ecological niche for *S.* Enteritidis to fill (119).

The prevalence of *S.* Enteritidis has declined in chicken and egg products in the United States since the mid-1990s due to multiple factors (120, 121). The recent emergence and spread of *S.* Heidelberg and *S.* Kentucky in poultry could be attributed to the acquisition of virulence plasmids which harbor genes for iron acquisition, colicin production, and disinfectant and heavy metal resistance via HGT, providing a selective advantage in the avian environment (107, 122, 123). Additionally, *S.* Enteritidis and *S.* Heidelberg share a common immunodominant surface O antigen (O12) (124). The prevalence of *S.* Heidelberg is partially due to the shared O12 antigen as well, as it competes for the same ecological niche with *S.* Enteritidis. Similar to *S.* Gallinarum and *S.* Enteritidis, immunization of chickens specifically against infection with serovar *S.* Enteritidis (125, 126) led to the decrease of *S.* Enteritidis and expansion of serovar *S.* Heidelberg in poultry populations. The evolution of *Salmonella* genomes related to poultry together with intervention strategies in poultry population marked a path toward the shifts in *Salmonella* serovars in egg and poultry production.

***Salmonella* virulence and genomic evolution.** Compared to *Escherichia coli*, which has a bigger pan-genome (127), *S. enterica* has a smaller pan-genome, which indicates that the rate of discovery of new genomic regions would decrease for each new genome of the species sequenced (128–130). The *S. enterica* pan-genome and core genome have been examined based on different sets of available genomes (56, 57, 127–132). A recent study of 4,893 genomes of *S. enterica* identified a pan-genome of 25.3 Mbp, a strict core of 1.5 Mbp present in all genomes, and a conserved core of 3.2 Mbp found in at least 96% of these genomes. Given an average gene size of 1,000 bp, the core genome has ~1,500 genes and includes ~3,200 genes in the conserved core genome, with a much larger pan-genome of ~25,300 genes (131). Worley et al. (132) described a core genome for *Salmonella* that included 2,278 genes present only once in each genome and of the same length, without indels, comprising 2,036,954 bp, which is less than half of the known reference genome of *S. enterica* serovar Typhimurium LT2, comprising 4,857,450 bp. Larger *Salmonella* genomes have been reported.

S. enterica phylogeny based on WGS indicated that important acquisitions from a virulence perspective included acquisition of SPI-1, which enables invasion of host cells, by the most recent common ancestor of all *Salmonella* subspecies, and SPI-2, for replication in macrophages, during species divergence of *S. enterica* from *S. bongori* (56). SPIs play a crucial role in the pathogenesis of *S. enterica* infections. So far, 24 SPIs have been described and characterized. Of all SPIs reported in *Salmonella*, only SPI-1, SPI-4, SPI-5, and SPI-9 were acquired by *Salmonella* prior to speciation (56). The acquisition of SPI by HGT confers rapid gain of complex virulence functions from other species. Although several common motifs are present among SPI, the distribution, size, structure, and function of these SPIs can be markedly different among subspecies, serovars, and/or strains. One such example is SPI-3, which has at least four different versions and no identical copies within a single version (128). Other

examples are SPI-13 and SPI-8. SPI-13 is conserved in many serovars in clade A and clade B, except the Typhi clade (containing *S.* Typhi and *S.* Paratyphi A) in clade A, while SPI-8 is carried only by the Typhi clade at the same genomic location (56, 132, 133). Different roles for SPI-13 and SPI-8 have been reported in intracellular macrophage survival (134). Moreover, several SPIs, such as SPI-7, are excisable from *Salmonella* chromosome by site-specific recombination, which can be of great epidemiological importance and may be a mechanism to regulate gene expression during human infection as well (135).

Type III secretion systems (T3SS), encoded by SPI-1 and SPI-2, respectively, play important roles in gastrointestinal disease and systemic infection. Comparative genomic analysis reveals that all serovars of *S. enterica* encode a subset of core effectors, and additional effectors are sporadically distributed among different serovars (132, 136), suggesting that they are critical for virulence in different hosts. In addition, a third T3SS is responsible for the flagellum-based motility of the pathogen (56). *Salmonella* expresses a characteristic intracellular transcriptomic signature in different cell types. Simultaneous expression of three T3SSs suggested a time-dependent transcriptional adaptation to the environment (137). The fluctuations in expression of *mgtBC*, *pstACS*, and *iro* genes, for magnesium, phosphate, and iron uptake, and T3SS could reflect bacterial response to host cells during infection (137). Regulations of these virulence factors can occur both globally and locally, forming complex feedback and feed-forward regulatory loops. For example, transcription of *hilA*, encoding the activator of the T3SS-1 structural genes, is activated by three AraC-like proteins, HilD, HilC, and RtsA, which each bind the promoter of *hilA* to directly enhance transcription (138, 139).

Many environmental signals, including low oxygen and high osmolarity, and regulatory systems such as the small RNAs (sRNAs) FnrS and AcrZ are integrated into this circuit to precisely regulate SPI-1 expression (140). By base pairing with target mRNA or protein, sRNAs modulate expression of distinct regulons and key transcription factors and play an important role in major stress response and virulence networks in *Salmonella* (141). Moreover, recent advances in *Salmonella* pathogenicity showed that *Salmonella* can cause infection in a T3SS-1-independent manner, which is mediated by large outer membrane proteins called invasins, namely, Rck and PagN, which allow salmonellae to invade host cells (142). Comparative genomics revealed four gene clusters encoding T6SS, acquired by independent lateral transfer events, located in different genomic islands, including SPI-6, SPI-19, SPI-20, and SPI-21 (143). Interestingly, *S.* Enteritidis has a degenerate genetic element lacking about 22 T6SS-related open reading frames (ORF) on SPI-19 with respect to *S.* Gallinarum, suggesting the role of T6SS in *Salmonella* evolution and host specificity.

Besides O and H antigens, *Salmonella* carries another group of surface proteins designated FGCs. Some types of FGCs are restricted to a host, and some are carried by all *Salmonella* serovars, suggesting a potential role for fimbriae in host specificity. Comparison of 90 genomes and 60 plasmids of *Salmonella* revealed a fimbriome consisting of 35 different FGCs in the *Salmonella* pan-genome, each carrying the structural subunits and biogenesis genes of a fimbria (90). The *Salmonella* fimbriome was extremely diverse due to the extensive FGC deletion and acquisition through HGT and to a high level of allelic variation in predicted or known adhesins which parallel *Salmonella* evolution toward host range modulation, survival, and persistence in specific niches, as well as strain virulence (90, 144).

Salmonella genomic evolution in the environment.

Salmonellae possess multiple traits that permit survival in a diverse set of environments, such as soils, sediments, waters, and plant surfaces (145). In order to endure in these environments, *Salmonella* must be able to overcome several stresses, including extremes of temperature, pH, salt/osmotic pressure, moisture, exposure to UV, and predators, to name a few (145, 146). Long-term persistence of these pathogens in the environment has been documented. For example, *Salmonella* introduced into corn crop soil through naturally contaminated poultry litter was found a year later (147). In surface waters from the Eastern Shore of Virginia, *Salmonella* with the same pulsed-field gel electrophoresis (PFGE) pattern was isolated over multiple years (4, 148). Some mechanisms that *Salmonella* uses for survival in the environment are similar to those used during an infection (145). Interestingly even some virulence factors have been shown to be important to environmental survival. For example, Maserati et al. (149) demonstrated a role for *sopD* and *sseD* in desiccation tolerance and

survival. As more naturally occurring isolates are recovered and sequenced from various environments, more unique adaptations for survival in those environments may be discovered.

Genomic evolution and phylogeny of *Salmonella*. FDA phylogenetic methods of evaluating *Salmonella* phylogeny enhance vertical evolution. The Center for Food Safety and Applied Nutrition (CFSAN) SNP pipeline is designed and validated to cluster isolates and to look for closely related shared ancestry (150). The HGT elements are examined in addition to the vertical signal, with the vertical elements defining the phylogenetic clusters and the HGT elements defining any pathogenicity, virulence, or AMR genes present among the isolates of interest.

Salmonella enterica is represented by >2,600 serovars, making it difficult to fully place in a phylogenetic context with a single analysis and resultant phylogeny. Investigators have attempted to capture all of the known diversity by using MLST (PubMLST [https://pubmlst.org/salmonella/]) (151), and several MLST schemes are available or have been proposed (34, 152, 153). Large phylogenies have been built and diversity within the genus *Salmonella* has been described using more of the WGS available data. Worley et al. (132) combined 445 isolate genomes from 266 distinct serovars and from 52 countries to build a comprehensive WGS phylogeny. An important finding from the study was that more than 10% of the examined serovars, nearly three dozen, designated by SeroSeq (24) were either polyphyletic or paraphyletic. These results suggest that the serovar markers have moved across the genus horizontally, though a clear timeline has not been established or proposed. This WGS study reported on two previously unidentified *S. enterica* subsp. *enterica* clades labeled C and D, to add to the two other major lineages, A and B, that other *Salmonella* phylogenetics works have identified. Gifsy-1- and Gifsy-2-like phages appear more prevalent in clade A. Most virulence genes are widely distributed across *S. enterica*, suggesting extensive, frequent HGT and a more dynamic hypothesis for bacterial evolution of the species. For antigen evolution, observations suggest that the genes responsible for O groups and phase 1 flagellar antigen traits are not evolving in a vertical fashion, suggesting an HGT role.

In addition, nearly four dozen phages were identified among the *Salmonella* isolates examined, with over three quarters of isolates having an associated phage. Numerous isolates with multiple phages were observed often, and one isolate had up to 6 intact phages. Gifsy-1 and Fels-2 phages were commonly observed, with the Gifsy-2 phage also being common for this largely North American and Asian sample set. Plasmid replicons are commonly found in *Salmonella* isolates, with IncFI and IncFII regularly being present among many genomes of the more than two dozen other plasmid replicons identified. Virulence factors (154–156) from type III secretion system (T3SS) SPI-1, including *sipA*, *sipB*, *sipC*, and *sptP*, were identified in all *Salmonella* genomes examined, as were the SPI-2 T3SS genes *spiC* and *ssaB*. Other *Salmonella* virulence factors varied in clade presence or absence, suggesting a complex evolutionary pattern.

CRISPR-Cas systems have been identified in numerous *Salmonella* genomes, with the alignments revealing mixed homology across serovars with an increase of shared spacers toward the ancestral end of the CRISPR array. Spacer alignments have revealed degradation of many internal spacers (57). The median number of spacers in CRISPR 1 and 2 are 13 and 14, respectively, with the largest array of 113 spacers being reported for *Salmonella enterica* serovar Mbandaka.

Applications from long-read sequencing. There are several different sequencing technologies that produce longer reads. Read lengths of 10,000 to 100,000 bp and longer have been described. Once these longer reads exceeded roughly 11,000 bp, more *Salmonella* genomes could be more easily and completely sequenced, having spanned a common major repeat. These closed genomes included the plasmids and phages associated with the bacterial genome. A closed genome refers to sequences that produce a single contig for each chromosome and mobile element present in the isolate. Numerous groups with access to these sequencing technologies began closing genomes and plasmids with a focus on fully describing the synteny of the genes on the chromosome and plasmids. Knowing the specific order and presence of genes allowed investigators to determine new AMR genes and the pathogenicity and virulence genes associated with each unique isolate and plasmid. For most foodborne-pathogen genomes, the presence of a known AMR gene conferred the phenotype of resistance.

Sequences with multiple contigs may interrupt some genes near long repetitive regions, confounding their identification. Genes may appear to be absent because they are only partially sequenced. The value of having a fully closed and characterized genome allowed investigators to see all the genes that are present with more certainty, so that they can match gene presence with phenotype presence. This also allowed discovery of new AMR genes when the gene was absent but the phenotype was present and the discovery of allelic variants when the gene was present but the phenotype was absent. This strategy of closing genomes to better characterize *Salmonella* will work as well for any other gene that investigators wish to characterize phenotypically, such as those involved in resistance to disinfectants and resistance to desiccation or heat (157). Long-read sequencers have various levels of error, so some reads require polishing with a higher-accuracy sequencing chemistry if the application being used is MLST or SNP-based phylogenetics (158). Even long reads with higher error rates can be useful, as the data can provide accurate scaffolds to determine gene synteny and can be fully applied to BLAST searches that discover the presence and absence of genes and plasmids against reference databases. An additional benefit is the ability to sequence in the field due to the mobile rapid nature of some of these tools.

Application of metagenomics in pathogen detection in food. A typical *Salmonella* isolation and identification take 5 to 7 days using the isolate-centric workflows described in the United States Food and Drug Administration's bacteriological analytical manual (BAM) (https://www.fda.gov/food/foodscienceresearch/laboratorymethods/ucm2006949.htm) and the U.S. Department of Agriculture Food Safety and Inspection Service's microbiology laboratory guidebook (MLG) (https://www.fsis.usda.gov/wps/portal/fsis/topics/science/laboratories-and-61procedures/guidebooks-and-methods/microbiology-laboratory-guidebook/microbiology-62laboratory-guidebook). Further characterization of the isolates can increase the laboratory turnaround time even more. The routine use of WGS substantially reduces time and cost for public health laboratories. However, current standard laboratory procedures for WGS, from regrowth of the pathogen to actual sequencing, take up to 5 days to complete. That means that the entire process from the collection of a contaminated food sample to the determination of the pathogen genome sequences can take up to 10 to 12 days. Unlike clinical samples, food samples often contain very low levels of *Salmonella* cells. In addition, the presence of competitive or antagonist organisms against salmonellae in the food microbiome can pose serious challenges for effective culture enrichment. Metagenomics, the collection of genomes and genes from the members of a microbiota obtained through shotgun sequencing of DNA extracted from a sample (159), however, is now beginning to provide a path forward in the use of WGS technology for *Salmonella* detection *in situ* in food and environmental backgrounds (160).

Despite the promise of metagenomic sequencing becoming a one-stop solution in food microbiology laboratories, it still has several challenges to overcome. One of the greatest challenges is the sensitivity and specificity of the current metagenomic methods for direct detection of low levels of pathogen of interest from high-background food microbiomes. Additionally, the choice of the extraction and sequencing protocols and the type of controls and analysis using metagenomic software tools are other major challenges to implementing and standardizing metagenomics for routine use in food microbiology laboratories (161, 162). Recent studies using a hybrid "quasi-metagenomic" approach demonstrated detection and subtyping of Shiga toxin-producing *Escherichia coli* (STEC) from spinach (163), *Listeria monocytogenes* in ice cream (164), and *Salmonella enterica* on cilantro (165), black peppercorn, peanut butter, and lettuce (166) and in wheat flour (167). Quasi-metagenomic sequencing is a direct sequencing of microbiological enrichments (the first step in culture-based detection methods). Sequencing of the modified microbiomes of food and environmental samples can provide high-resolution sequencing data for foodborne pathogen detection and subtyping, expediting source tracking by up to 4 to 6 days (168–170). These studies also underscore the value of metagenomics as a tool to evaluate and rationalize culture enrichment methods. Metagenomics reveals which species grow along the enrichment timeline and documents changes in species composition with various perturbations to enrichment protocols. Long-read sequencers also show promise for metagenomics methods, with the longer reads more accurately characterizing the species present in a microbiome (171).

Integration of genomics, investigation, and epidemiology. Foodborne contamination events and outbreaks are investigated by numerous federal and state partners, including the FDA, USDA Food Safety and Inspection Service (FSIS), CDC, and NCBI. Investigations are supported by three lines of potential evidence. The first lines of evidence often come from the laboratory, where WGS provides genetic support for a phylogenetic cluster that links food, environmental, and clinical isolates. By focusing on the most closely related isolates at the tips of the phylogenetic tree, WGS clusters a subset of the isolates that are monophyletic and share an ancestor. These subclusters are often used to separate outbreak signals from background noise, to unravel the complexities of foodborne contaminations, to support and prioritize epidemiological data, and to carry out site investigations. WGS unravels the complexity of a polyclonal outbreak by breaking the investigation into smaller solvable parts. Each lineage within a polyclonal outbreak or contamination event is treated as an independent pathogen and piece of evidence tying a specific food commodity or firm to a clinical case. Epidemiological evidence may determine whether the patients with clinical cases have been exposed to a common contaminant found at a firm. The FDA inspection may provide positive cultures of the foodborne pathogens contaminating the facility. The FDA relies on field inspectors to recover the diversity of pathogens present in a contaminated facility. For FDA compliance, it is often the inspection results that determine whether a contamination event is polyclonal. Also, multiple WGS clusters may each independently point back to the same firm being responsible for the contaminant exposure. The power and prediction of the full investigation comes from integrating the various relevant pieces of evidence, including those from laboratory and epidemiological investigation.

Genomic methods are always superior to lower-resolution subtyping methods when the goals are source tracking, root cause investigation, and infectious disease control. Having more data is better for numerous reasons. The superior performance of WGS methods is the reason why states and federal agencies have adopted WGS for all investigations of foodborne illness (12). WGS is best suited to integrate all case information, provided that its use is not delayed. We have not seen any WGS evidence to suggest that having more data creates any problems, and in fact, the reverse is always the case, in that having more data provides higher resolution, which more clearly defines a contamination case and the explicit genetic changes that have occurred among the isolates sequenced (116). As more experience is gained using WGS, the examples and evidence continue to show the powerful predictive role that genomics plays in investigating contamination events (11, 62, 83, 100, 108). Methods of reduced resolution generally increase false inclusions, which are particularly problematic for ecological and epidemiological models when clinical, food, and environmental isolates are included that were not part of the same contamination event. False inclusions misdirect investigations and reduce the power of prediction, ultimately delaying removal of the contaminant from the food supply.

High-resolution SNP analysis resolves all isolates down to the very tips of the tree (150). Phylogenetic trees are hierarchical, showing greater and greater resolution from the base to the tip of the tree. If exposure data suggest a common contaminant and or food vehicle at a particular node that is supported by WGS data, then that node on the phylogenetic tree can be set as the case definition and scope of the outbreak. Often there is clear evidence for a cluster break based on the number and/or positions of SNPs that define a lineage and the bootstrap scores for the node. WGS provides additional evidence about the amount of genetic diversity that has accrued during a contamination event. It is the high-resolution WGS data, combined with detailed and structured metadata, that may be used by artificial intelligence (AI) and machine learning (ML) tools to make even more predictive models for the accurate prediction of food, animal source (172), and or geographic location. Published WGS data have shown that most *Salmonella* and *Listeria* isolates exhibit a very strong phylogeographic signal that is highly predictive (173), based on the ability to predict with high probability whether a pathogen comes from the same facility, for isolates acquired during inspection. We also know that isolates from clinical sources show similar levels of genetic variability, suggesting that they would show similar probabilities if comparable evidence was available to predict the sources of their illness.

Risk assessment and risk management predictions also benefit from WGS data (174). FDA and GenomeTrakr partners are including more detailed structured metadata

food ontology (175) (FoodOn, GenEpiOn, MixS, and IFSAC) to support efforts to foster innovation in AI and ML. We have already seen numerous WGS examples of the power to predict country of origin (11), growing region (176, 177), and even implicated egg farms (178). As we see phylogeographic structure in most of the trees we build, it is likely that AI and ML will contribute additional future predictions to support contamination and outbreak investigations. FDA investigators currently watch approximately 4,000 of the more than 40,000 clusters at the NCBI Pathogen Detection web site for isolates that cluster with FDA foodborne pathogen genomes. This includes data from roughly 340,000 *Salmonella* genomes, a number which has grown from less than 1,000 in 2012.

Cladistic methodology is one approach used to build phylogenetic trees using parsimony methods. Cladistics is uniquely valuable in optimizing characters on a phylogenetic tree to predict when character variation occurs. Using these methods, investigators can predict the unique changes that define a lineage. The nucleotide changes that modify the coded amino acid (nonsynonymous changes) may also modify the protein and affect the phenotype. By combining cladistics, character optimization, and WGS, investigators may be able to identify genotype-to-phenotype changes that specific bacterial lineages have acquired and that allow foodborne pathogens to survive and contaminate foods, animals, and the environment (177). In several examples, investigators have predicted which genomic changes correlate with outbreaks in Italian-style meats (178) and in eggs (115, 116, 179), with the underlying phenotype predictions uncovering known pathogenicity and virulence gene variants and/or the ability to infect the chicken host. These general methods will continue to be valuable for constructing genotype-to-phenotype hypotheses.

Global genomic standards. Harmonization of test protocols from different organizations, e.g., FDA, International Organization for Standardization (ISO), AOAC International, and Association Française de Normalisation (AFNOR), has been pursued recently to facilitate global data sharing and comparison when dealing with worldwide public health problems (180, 181). The need to validate newly developed or alternative methods in comparison with established and standard protocols, such as FDA, USDA, AOAC, ISO, and AFNOR methods, has become urgent in recent years, in order to make sure that proper methods are used in all case investigations, survey, surveillance, monitoring, and outbreak investigations. The ISO creates standards to facilitate trade by forming trust that is based on consensus among groups of experts in government, industry, and academia. There are more than 21,000 ISO standards that address a wide variety of topics, including food microbiology. These standards develop trust among trading partners by standardizing the activities in which they are involved. For example, ISO develops consensus positions on food microbiological standard methods in ISO Technical Committee (TC) 34/Subcommittee (SC) 9. TC 34 is devoted to foods, and SC 9 is devoted to microbiology. Within TC 34/SC 9, there is working group (WG) 25, "Whole-genome sequencing for typing and genomic characterization."

The WG 25 recently completed a committee draft (23418; "Whole-genome sequencing for typing and genomic characterization of foodborne bacteria—general requirements and guidance"). The purpose of this standard is to address both the laboratory and bioinformatic components of WGS for foodborne microorganisms. The overall goal of this standard is to provide consistency in the approach to WGS regardless of the sequencing instrument, so that sequencing results will be comparable throughout the world. The standard is in three parts: laboratory operations, validation, and metadata. Within these areas, the standard covers handling of bacterial cultures; genomic DNA isolation; sequencing library preparation, sequencing, and assessment of raw DNA sequence read quality and storage; bioinformatics analysis for determining genetic relatedness, genetic content and predicting phenotype, and bioinformatics pipeline validation; metadata capture and sequence repository deposition; and validation of the end-to-end WGS workflow. These parameters are the minimum necessary for generating and analyzing WGS data obtained from foodborne bacteria.

ACKNOWLEDGMENTS

We thank our partners at the United States Food and Drug Administration. We thank the many collaborators in the GenomeTrakr network, CDC PulseNet, USDA-FSIS, who provide their data for broad use and applications, and we acknowledge the data support that we all receive at the FDA and the NCBI.

Work was funded by internal research funding from the U.S. Food and Drug Administration.

We declare no conflicts of interest, financial or otherwise.

REFERENCES

1. Grimont PA, Weill F-X. 2007. Antigenic formulae of the Salmonella serovars, p 166. WHO Collaborating Centre for Reference and Research on Salmonella, Paris, France.

2. Boyd EF, Wang F-S, Whittam TS, Selander RK. 1996. Molecular genetic relationships of the salmonellae. *Appl Environ Microbiol* **62:**804–808. https://doi.org/10.1128/AEM.62.3.804-808.1996.

3. Winfield MD, Groisman EA. 2004. Evolution and ecology of Salmonella. *EcoSal Plus.* https://doi.org/10.1128/ecosalplus.6.4.6.

4. Bell RL, Zheng J, Burrows E, Allard S, Wang CY, Keys CE, Melka DC, Strain E, Luo Y, Allard MW, Rideout S, Brown EW. 2015. Ecological prevalence, genetic diversity, and epidemiological aspects of Salmonella isolated from tomato agricultural regions of the Virginia Eastern Shore. *Front Microbiol* **7:**415. https://doi.org/10.3389/fmicb.2015.00415.

5. CDC. 2007. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food–10 states, 2006. *MMWR Morb Mortal Wkly Rep* **56:**336–339.

6. Klontz KC, Klontz JC, Mody RK, Hoekstra RM. 2010. Analysis of tomato and jalapeño and Serrano pepper imports into the United States from Mexico before and during a national outbreak of Salmonella serotype Saintpaul infections in 2008. *J Food Prot* **73:**1967–1974. https://doi.org/10.4315/0362-028X-73.11.1967.

7. CDC. 2017. Multistate outbreak of *Salmonella* infections linked to imported Maradol papayas (final update). https://www.cdc.gov/salmonella/kiambu-07-17/.

8. Angelo KM, Chu A, Anand M, Nguyen TA, Bottichio L, Wise M, Williams I, Seelman S, Bell R, Fatica M, Lance S, Baldwin D, Shannon K, Lee H, Trees E, Strain E, Gieraltowski L, Centers for Disease Control and Prevention. 2015. Centers for Disease Control and Prevention (CDC). Outbreak of Salmonella Newport infections linked to cucumbers–United States, 2014. *MMWR Morb Mortal Wkly Rep* **64:**144–147.

9. Allard MW, Bell R, Ferreira CM, Gonzalez-Escalona N, Hoffmann M, Muruvanda T, Ottesen A, Ramachandran P, Reed E, Sharma S, Stevens E, Timme R, Zheng J, Brown EW. 2018. Genomics of foodborne pathogens for microbial food safety. *Curr Opin Biotechnol* **49:**224–229. https://doi.org/10.1016/j.copbio.2017.11.002.

10. Alikhan N-F, Zhou Z, Sergeant MJ, Achtman M. 2018. A genomic overview of the population structure of Salmonella. *PLoS Genet* **14:**e1007261. https://doi.org/10.1371/journal.pgen.1007261.

11. Hoffmann M, Luo Y, Monday SR, Gonzalez-Escalona N, Ottesen AR, Muruvanda T, Wang C, Kastanis G, Keys C, Janies D, Senturk IF, Catalyurek UV, Wang H, Hammack TS, Wolfgang WJ, Schoonmaker-Bopp D, Chu A, Myers R, Haendiges J, Evans PS, Meng J, Strain EA, Allard MW, Brown EW. 2016. Tracing origins of the Salmonella bareilly strain causing a food-borne outbreak in the United States. *J Infect Dis* **213:**502–508. https://doi.org/10.1093/infdis/jiv297.

12. Allard MW, Strain E, Melka D, Bunning K, Musser SM, Brown EW, Timme R. 2016. Practical value of food pathogen traceability through building a whole-genome sequencing network and database. *J Clin Microbiol* **54:**1975–1983. https://doi.org/10.1128/JCM.00081-16.

13. Kovac J, Bakker Hd, Carroll LM, Wiedmann M. 2017. Precision food safety: a systems approach to food safety facilitated by genomics tools. *Trends Anal Chem* **96:**52–61. https://doi.org/10.1016/j.trac.2017.06.001.

14. LeMinor L. 1970. Supplement no. 13 to the Kauffman-White scheme. *Ann Inst Pasteur (Paris)* **119:**438–443.

15. Le Minor L, Véron M, Popoff M. 1982. A proposal for Salmonella nomenclature. *Ann Microbiol (Paris)* **133:**245–254.

16. Diep B, Barretto C, Portmann AC, Fournier C, Karczmarek A, Voets G, Li S, Deng X, Klijn A. 2019. Salmonella serotyping; comparison of the traditional method to a microarray-based method and an in silico platform using whole genome sequencing data. *Front Microbiol* **11:**2554. https://doi.org/10.3389/fmicb.2019.02554.

17. Kim S, Frye JG, Hu J, Fedorka-Cray PJ, Gautom R, Boyle DS. 2006. Multiplex PCR-based method for identification of common clinical serotypes of Salmonella enterica subsp. enterica. *J Clin Microbiol* **44:**3608–3615. https://doi.org/10.1128/JCM.00701-06.

18. Liu B, Zhang L, Zhu X, Shi C, Chen J, Liu W, He X, Shi X. 2011. PCR identification of Salmonella serogroups based on specific targets obtained by comparative genomics. *Int J Food Microbiol* **144:**511–518. https://doi.org/10.1016/j.ijfoodmicro.2010.11.010.

19. Salazar JK, Wang Y, Yu S, Wang H, Zhang W. 2015. Polymerase chain reaction-based serotyping of pathogenic bacteria in food. *J Microbiol Methods* **110:**18–26. https://doi.org/10.1016/j.mimet.2015.01.009.

20. McQuiston JR, Waters RJ, Dinsmore BA, Mikoleit ML, Fields PI. 2011. Molecular determination of H antigens of Salmonella by use of a microsphere-based liquid array. *J Clin Microbiol* **49:**565–573. https://doi.org/10.1128/JCM.01323-10.

21. Moore MM, Nucci MJ, Madson SM, Wagley GS, Keys CE, Brown EW, McQuiston JR, Fields PI. 2019. Evaluation of a bead-based Salmonella molecular serotyping method for Salmonella isolated from food and environmental samples. *J Food Prot* **82:**1973–1987. https://doi.org/10.4315/0362-028X.JFP-18-600.

22. Zhang S, Yin Y, Jones MB, Zhang Z, Deatherage Kaiser BL, Dinsmore BA, Fitzgerald C, Fields PI, Deng X. 2015. Salmonella serotype determination utilizing high-throughput genome sequencing data. *J Clin Microbiol* **53:**1685–1692. https://doi.org/10.1128/JCM.00323-15.

23. Cox NA, Berrang ME, House SL, Medina D, Cook KL, Shariat NW. 2019. Population analyses reveal preenrichment method and selective enrichment media affect Salmonella serovars detected on broiler carcasses. *J Food Prot* **82:**1688–1696. https://doi.org/10.4315/0362-028X.JFP-19-166.

24. Thompson CP, Doak AN, Amirani N, Schroeder EA, Wright J, Kariyawasam S, Lamendella R, Shariat NW. 2018. High-resolution identification of multiple Salmonella serovars in a single sample by using CRISPR-SeroSeq. *Appl Environ Microbiol* **84:**e01859-18. https://doi.org/10.1128/AEM.01859-18.

25. Kozyreva VK, Ilina EN, Malakhova MV, Carattoli A, Azizov IS, Tapalski DV, Kozlov RS, Edelstein MV. 2014. Long-term dissemination of CTX-M-5-producing hypermutable Salmonella enterica serovar typhimurium sequence type 328 strains in Russia, Belarus, and Kazakhstan. *Antimicrob Agents Chemother* **58:**5202–5210. https://doi.org/10.1128/AAC.02506-14.

26. LeClerc JE, Li B, Payne WL, Cebula TA. 1996. High mutation frequencies among Escherichia coli and Salmonella pathogens. *Science* **274:**1208–1211. https://doi.org/10.1126/science.274.5290.1208.

27. LeClerc JE, Payne WL, Kupchella E, Cebula TA. 1998. Detection of mutator subpopulations in Salmonella typhimurium LT2 by reversion of his alleles. *Mutat Res* **400:**89–97. https://doi.org/10.1016/S0027-5107(98)00069-4.

28. Brown EW, Kotewicz ML, Cebula TA. 2002. Detection of recombination among Salmonella enterica strains using the incongruence length difference test. *Mol Phylogenet Evol* **24:**102–120. https://doi.org/10.1016/S1055-7903(02)00222-1.

29. Brown EW, Mammel MK, LeClerc JE, Cebula TA. 2003. Limited boundaries for extensive horizontal gene transfer among Salmonella pathogens. *Proc Natl Acad Sci U S A* **100:**15676–15681. https://doi.org/10.1073/pnas.2634406100.

30. Kotewicz ML, Li B, Levy DD, LeClerc JE, Shifflet AW, Cebula TA. 2002. Evolution of multi-gene segments in the mutS-rpoS intergenic region of Salmonella enterica serovar Typhimurium LT2. *Microbiology (Reading)* **148:**2531–2540. https://doi.org/10.1099/00221287-148-8-2531.
31. Boyd EF, Wang F-S, Beltran P, Plock SA, Nelson K, Selander RK. 1993. Salmonella reference collection B (SARB): strains of 37 serovars of subspecies I. *J Gen Microbiol* **139:**1125–1132. https://doi.org/10.1099/00221287-139-6-1125.
32. Beltran P, Plock SA, Smith NH, Whittam TS, Old DC, Selander RK. 1991. Reference collection of strains of the Salmonella typhimurium complex from natural populations. *J Gen Microbiol* **137:**601–606. https://doi.org/10.1099/00221287-137-3-601.
33. Porwollik S, McClelland M. 2003. Lateral gene transfer in Salmonella. *Microbes Infect* **5:**977–989. https://doi.org/10.1016/S1286-4579(03)00186-2.
34. Fricke WF, Mammel MK, McDermott PF, Tartera C, White DG, Leclerc JE, Ravel J, Cebula TA. 2011. Comparative genomics of 28 Salmonella enterica isolates: evidence for CRISPR-mediated adaptive sublineage evolution. *J Bacteriol* **193:**3556–3568. https://doi.org/10.1128/JB.00297-11.
35. Groisman EA, Ochman H. 2000. The path to Salmonella. *ASM News* **66:**21–27.
36. Hacker J, Kaper JB. 2000. Pathogenicity islands and the evolution of microbes. *Annu Rev Microbiol* **54:**641–679. https://doi.org/10.1146/annurev.micro.54.1.641.
37. Bäumler AJ. 1997. The record of horizontal gene transfer in Salmonella. *Trends Microbiol* **5:**318–322. https://doi.org/10.1016/S0966-842X(97)01082-2.
38. Kingsley RA, Baumler AJ. 2000. Host adaptation and the emergence of infectious disease: the Salmonella pathogen. *Mol Microbiol* **36:**1006–1014. https://doi.org/10.1046/j.1365-2958.2000.01907.x.
39. Boyd EF, Li J, Ochman H, Selander RK. 1997. Comparative genetics of the inv-spa invasion complex of Salmonella enterica. *J Bacteriol* **179:**1985–1991. https://doi.org/10.1128/JB.179.6.1985-1991.1997.
40. Brown EW, Bell RL, Allard MW, Gonzalez-Escalona N, Perlloni A, LeClerc JE, Cebula TA. 2012. Reticulate evolution among the group I salmonellae: an ongoing role for horizontal gene transfer, p 209–230. *In* Kumar Y (ed), *Salmonella*—a diversified superbug. InTechOpen, London, United Kingdom.
41. Boyd EF, Hartl DL. 1999. Analysis of the type 1 pilin gene cluster fim in Salmonella: its distinct evolutionary histories in the 5′ and 3′ regions. *J Bacteriol* **181:**1301–1308. https://doi.org/10.1128/JB.181.4.1301-1308.1999.
42. Nelson K, Selander RK. 1994. Intergeneric transfer and recombination of the 6- phosphogluconate dehydrogenase gene (gnd) in enteric bacteria. *Proc Natl Acad Sci U S A* **91:**10227–10231. https://doi.org/10.1073/pnas.91.21.10227.
43. McQuiston JR, Herrera-Leon S, Wertheim BC, Doyle J, Fields PI, Tauxe RV, Logsdon JM. 2008. Molecular phylogeny of the salmonellae: relationships among Salmonella species and subspecies determined from four housekeeping genes and evidence of lateral gene transfer events. *J Bacteriol* **190:**7060–7067. https://doi.org/10.1128/JB.01552-07.
44. Octavia S, Lan R. 2006. Frequent recombination and low level of clonality within Salmonella enterica subspecies I. *Microbiology (Reading)* **152:**1099–1108. https://doi.org/10.1099/mic.0.28486-0.
45. Mebrhatu MT, Cenens W, Aertsen A. 2014. An overview of the domestication and impact of the Salmonella mobilome. *Crit Rev Microbiol* **40:**63–75. https://doi.org/10.3109/1040841X.2012.755949.
46. Tanner JR, Kingsley RA. 2018. Evolution of Salmonella within hosts. *Trends Microbiol* **26:**986–998. https://doi.org/10.1016/j.tim.2018.06.001.
47. Ilyas B, Tsai CN, Coombes BK. 2017. Evolution of Salmonella-host cell interactions through a dynamic bacterial genome. *Front Cell Infect Microbiol* **7:**428. https://doi.org/10.3389/fcimb.2017.00428.
48. Bae YM, Baek SY, Lee SY. 2012. Resistance of pathogenic bacteria on the surface of stainless steel depending on attachment form and efficacy of chemical sanitizers. *Int J Food Microbiol* **153:**465–473. https://doi.org/10.1016/j.ijfoodmicro.2011.12.017.
49. Hall RM. 2010. Salmonella genomic islands and antibiotic resistance in Salmonella enterica. *Future Microbiol* **5:**1525–1538. https://doi.org/10.2217/fmb.10.122.
50. Sharp PM, Kelleher JE, Daniel AS, Cowan GM, Murray NE. 1992. Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. *Proc Natl Acad Sci U S A* **89:**9836–9840. https://doi.org/10.1073/pnas.89.20.9836.
51. Seiffert SN, Hilty M, Perreten V, Endimiani A. 2013. Extended-spectrum cephalosporin-resistant Gram-negative organisms in livestock: an emerging problem for human health? *Drug Resist Update* **16:**22–45. https://doi.org/10.1016/j.drup.2012.12.001.
52. Kaushik M, Kumar S, Kapoor RK, Gulati P. 2019. Integrons and antibiotic resistance genes in water-borne pathogens: threat detection and risk assessment. *J Med Microbiol* **68:**679–692. https://doi.org/10.1099/jmm.0.000972.
53. Aarestrup FM. 2015. The livestock reservoir for antimicrobial resistance: a personal view on changing patterns of risks, effects of interventions and the way forward. *Philos Trans R Soc Lond B Biol Sci* **370:**20140085. https://doi.org/10.1098/rstb.2014.0085.
54. Pettengill JB, Tate H, Gensheimer K, Hsu CH, Ihrie J, Markon AO, McDermott PF, Zhao S, Strain E, Bazaco MC. 2020. Distribution of antimicrobial resistance genes across Salmonella enterica isolates from animal and nonanimal foods. *J Food Prot* **83:**295–304. https://doi.org/10.4315/0362-028X.JFP-19-310.
55. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu CH, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob Agents Chemother* **63:**e00483-19. https://doi.org/10.1128/AAC.00483-19.
56. Desai PT, Porwollik S, Long F, Cheng P, Wollam A, Clifton SW, Weinstock GM, McClelland M. 2013. Evolutionary genomics of Salmonella enterica subspecies. *mBio* **4:**e00579-12. https://doi.org/10.1128/mBio.00579-12.
57. Timme RE, Pettengill JB, Allard MW, Strain E, Barrangou R, Wehnes C, Van Kessel JS, Karns JS, Musser SM, Brown EW. 2013. Phylogenetic diversity of the enteric pathogen Salmonella enterica subsp. enterica inferred from genome-wide reference-free SNP characters. *Genome Biol Evol* **5:**2109–2123. https://doi.org/10.1093/gbe/evt159.
58. Coburn B, Grassl GA, Finlay BB. 2007. Salmonella, the host and disease: a brief review. *Immunol Cell Biol* **85:**112–118. https://doi.org/10.1038/sj.icb.7100007.
59. Acheson D, Hohmann EL. 2001. Nontyphoidal salmonellosis. *Clin Infect Dis* **32:**263–269. https://doi.org/10.1086/318457.
60. Marzel A, Desai PT, Nissan I, Schorr YI, Suez J, Valinsky L, Reisfeld A, Agmon V, Guard J, McClelland M, Rahav G, Gal-Mor O. 2014. Integrative analysis of Salmonellosis in Israel reveals association of Salmonella enterica serovar 9,12:l,v:− with extraintestinal infections, dissemination of endemic S. enterica Serovar

Typhimurium DT104 biotypes, and severe underreporting of outbreaks. *J Clin Microbiol* **52:**2078–2088. https://doi.org/10.1128/JCM.00399-14.
61. Jones TF, Ingram LA, Cieslak PR, Vugia DJ, Tobin-D'Angelo M, Hurd S, Medus C, Cronquist A, Angulo FJ. 2008. Salmonellosis outcomes differ substantially by serotype. *J Infect Dis* **198:**109–114. https://doi.org/10.1086/588823.
62. Kariuki S, Onsare RS. 2015. Epidemiology and genomics of invasive nontyphoidal Salmonella infections in Kenya. *Clin Infect Dis* **61:**S317–24. https://doi.org/10.1093/cid/civ711.
63. Gong J, Zhang J, Xu M, Zhu C, Yu Y, Liu X, Kelly P, Xu B, Wang C. 2014. Prevalence and fimbrial genotype distribution of poultry Salmonella isolates in China (2006 to 2012). *Appl Environ Microbiol* **80:**687–693. https://doi.org/10.1128/AEM.03223-13.
64. Stipetic K, Chang YC, Peters K, Salem A, Doiphode SH, McDonough PL, Chang YF, Sultan A, Mohammed HO. 2016. The risk of carriage of Salmonella spp. and Listeria monocytogenes in food animals in dynamic populations. *Vet Med Sci* **2:**246–254. https://doi.org/10.1002/vms3.39.
65. Kagambèga A, Lienemann T, Aulu L, Traoré AS, Barro N, Siitonen A, Haukka K. 2013. Prevalence and characterization of Salmonella enterica from the feces of cattle, poultry, swine and hedgehogs in Burkina Faso and their comparison to human Salmonella isolates. *BMC Microbiol* **13:**253. https://doi.org/10.1186/1471-2180-13-253.
66. Gal-Mor O. 2019. Persistent infection and long-term carriage of typhoidal and nontyphoidal salmonellae. *Clin Microbiol Rev* **32:**e00088-18. https://doi.org/10.1128/CMR.00088-18.
67. Weening EH, Barker JD, Laarakker MC, Humphries AD, Tsolis RM, Bäumler AJ. 2005. The Salmonella enterica serotype Typhimurium lpf, bcf, stb, stc, std, and sth fimbrial operons are required for intestinal persistence in mice. *Infect Immun* **73:**3358–3366. https://doi.org/10.1128/IAI.73.6.3358-3366.2005.
68. Clayton DJ, Bowen AJ, Hulme SD, Buckley AM, Deacon VL, Thomson NR, Barrow PA, Morgan E, Jones MA, Watson M, Stevens MP. 2008. Analysis of the role of 13 major fimbrial subunits in colonisation of the chicken intestines by Salmonella enterica serovar Enteritidis reveals a role for a novel locus. *BMC Microbiol* **8:**228. https://doi.org/10.1186/1471-2180-8-228.
69. Dorsey CW, Laarakker MC, Humphries AD, Weening EH, Bäumler AJ. 2005. Salmonella enterica serotype Typhimurium MisL is an intestinal colonization factor that binds fibronectin. *Mol Microbiol* **57:**196–211. https://doi.org/10.1111/j.1365-2958.2005.04666.x.
70. Raghunathan D, Wells TJ, Morris FC, Shaw RK, Bobat S, Peters SE, Paterson GK, Jensen KT, Leyton DL, Blair JMA, Browning DF, Pravin J, Flores-Langarica A, Hitchcock JR, Moraes CTP, Piazza RMF, Maskell DJ, Webber MA, May RC, MacLennan CA, Piddock LJ, Cunningham AF, Henderson IR. 2011. SadA, a trimeric autotransporter from Salmonella enterica serovar Typhimurium, can promote biofilm formation and provides limited protection against infection. *Infect Immun* **79:**4342–4352. https://doi.org/10.1128/IAI.05592-11.
71. Kingsley RA, van Amsterdam K, Kramer N, Bäumler AJ. 2000. The shdA gene is restricted to serotypes of Salmonella enterica subspecies I and contributes to efficient and prolonged fecal shedding. *Infect Immun* **68:**2720–2727. https://doi.org/10.1128/IAI.68.5.2720-2727.2000.
72. Griessl MH, Schmid B, Kassler K, Braunsmann C, Ritter R, Barlag B, Stierhof YD, Sturm KU, Danzer C, Wagner C, Schäffer TE, Sticht H, Hensel M, Muller YA. 2013. Structural insight into the giant Ca(2)(+)-binding adhesin SiiE: implications for the adhesion of Salmonella enterica to polarized epithelial cells. *Structure* **21:**741–752. https://doi.org/10.1016/j.str.2013.02.020.
73. White AP, Gibson DL, Grassl GA, Kay WW, Finlay BB, Vallance BA, Surette MG. 2008. Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in Salmonella enterica serovar Typhimurium. *Infect Immun* **76:**1048–1058. https://doi.org/10.1128/IAI.01383-07.
74. Crawford RW, Gibson DL, Kay WW, Gunn JS. 2008. Identification of a bile-induced exopolysaccharide required for Salmonella biofilm formation on gallstone surfaces. *Infect Immun* **76:**5341–5349. https://doi.org/10.1128/IAI.00786-08.
75. Schaible UE, Kaufmann SH. 2004. Iron and microbial infection. *Nat Rev Microbiol* **2:**946–953. https://doi.org/10.1038/nrmicro1046.
76. Kehl-Fie TE, Skaar EP. 2010. Nutritional immunity beyond iron: a role for manganese and zinc. *Curr Opin Chem Biol* **14:**218–224. https://doi.org/10.1016/j.cbpa.2009.11.008.
77. Nagy TA, Moreland SM, Andrews-Polymenis H, Detweiler CS. 2013. The ferric enterobactin transporter Fep is required for persistent Salmonella enterica serovar typhimurium infection. *Infect Immun* **81:**4063–4070. https://doi.org/10.1128/IAI.00412-13.
78. Hawkey J, Edwards DJ, Dimovski K, Hiley L, Billman-Jacobe H, Hogg G, Holt KE. 2013. Evidence of microevolution of Salmonella Typhimurium during a series of egg-associated outbreaks linked to a single chicken farm. *BMC Genomics* **14:**800. https://doi.org/10.1186/1471-2164-14-800.
79. Octavia S, Wang Q, Tanaka MM, Sintchenko V, Lan R. 2015. Genomic variability of serial human isolates of Salmonella enterica serovar Typhimurium associated with prolonged carriage. *J Clin Microbiol* **53:**3507–3514. https://doi.org/10.1128/JCM.01733-15.
80. Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN, Kariuki S, Msefula CL, Gordon MA, de Pinna E, Wain J, Heyderman RS, Obaro S, Alonso PL, Mandomando I, MacLennan CA, Tapia MD, Levine MM, Tennant SM, Parkhill J, Dougan G. 2012. Intracontinental spread of human invasive Salmonella Typhimurium pathovariants in sub-Saharan Africa. *Nat Genet* **44:**1215–1221. https://doi.org/10.1038/ng.2423.
81. Marzel A, Desai PT, Goren A, Schorr YI, Nissan I, Porwollik S, Valinsky L, McClelland M, Rahav G, Gal-Mor O. 2016. Persistent infections by nontyphoidal Salmonella in humans: epidemiology and genetics. *Clin Infect Dis* **62:**879–886. https://doi.org/10.1093/cid/civ1221.
82. Uzzau S, Brown DJ, Wallis T, Rubino S, Leori G, Bernard S, Casadesús J, Platt DJ, Olsen JE. 2000. Host adapted serotypes of Salmonella enterica. *Epidemiol Infect* **125:**229–255. https://doi.org/10.1017/S0950268899004379.
83. Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, Wigley P, Barquist L, Langridge GC, Feltwell T, Harris SR, Mather AE, Fookes M, Aslett M, Msefula C, Kariuki S, Maclennan CA, Onsare RS, Weill F-X, Le Hello S, Smith AM, McClelland M, Desai P, Parry CM, Cheesbrough J, French N, Campos J, Chabalgoity JA, Betancor L, Hopkins KL, Nair S, Humphrey TJ, Lunguya O, Cogan TA, Tapia MD, Sow SO, Tennant SM, Bornstein K, Levine MM, Lacharme-Lora L, Everett DB, Kingsley RA, Parkhill J, Heyderman RS, Dougan G, Gordon MA, Thomson NR. 2016. Distinct Salmonella Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. *Nat Genet* **48:**1211–1217. https://doi.org/10.1038/ng.3644.
84. Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, Harris D, Clarke L, Whitehead S, Sangal V, Marsh K, Achtman M, Molyneux ME, Cormican M, Parkhill J, MacLennan CA, Heyderman RS, Dougan G. 2009. Epidemic multiple drug resistant Salmonella Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res* **19:**2279–2287. https://doi.org/10.1101/gr.091017.109.

85. Gal-Mor O, Boyle EC, Grassl GA. 2014. Same species, different diseases: how and why typhoidal and non-typhoidal Salmonella enterica serovars differ. *Front Microbiol* **5:**391. https://doi.org/10.3389/fmicb.2014.00391.
86. Nuccio SP, Bäumler AJ. 2014. Comparative analysis of Salmonella genomes identifies a metabolic network for escalating growth in the inflamed gut. *mBio* **5:**e00929-14. https://doi.org/10.1128/mBio.00929-14.
87. Rakov AV, Mastriani E, Liu S, Schifferli DM. 2019. Association of Salmonella virulence factor alleles with intestinal and invasive serovars. *BMC Genomics* **20:**429. https://doi.org/10.1186/s12864-019-5809-8.
88. Langridge GC, Fookes M, Connor TR, Feltwell T, Feasey N, Parsons BN, Seth-Smith HM, Barquist L, Stedman A, Humphrey T, Wigley P, Peters SE, Maskell DJ, Corander J, Chabalgoity JA, Barrow P, Parkhill J, Dougan G, Thomson NR. 2015. Patterns of genome evolution that have accompanied host adaptation in Salmonella. *Proc Natl Acad Sci U S A* **112:**863–868. https://doi.org/10.1073/pnas.1416707112.
89. De Masi L, Yue M, Hu C, Rakov AV, Rankin SC, Schifferli DM. 2017. Cooperation of adhesin alleles in Salmonella-host tropism. *mSphere* **2:**e00066-17. https://doi.org/10.1128/mSphere.00066-17.
90. Yue M, Rankin SC, Blanchet RT, Nulton JD, Edwards RA, Schifferli DM. 2012. Diversification of the Salmonella fimbriae: a model of macro- and microevolution. *PLoS One* **7:**e38596. https://doi.org/10.1371/journal.pone.0038596.
91. Yue M, Han X, De Masi L, Zhu C, Ma X, Zhang J, Wu R, Schmieder R, Kaushik RS, Fraser GP, Zhao S, McDermott PF, Weill FX, Mainil JG, Arze C, Fricke WF, Edwards RA, Brisson D, Zhang NR, Rankin SC, Schifferli DM. 2015. Allelic variation contributes to bacterial host specificity. *Nat Commun* **6:**8754. https://doi.org/10.1038/ncomms9754.
92. Van Puyvelde S, Pickard D, Vandelannoote K, Heinz E, Barbé B, de Block T, Clare S, Coomber EL, Harcourt K, Sridhar S, Lees EA, Wheeler NE, Klemm EJ, Kuijpers L, Mbuyi Kalonji L, Phoba MF, Falay D, Ngbonda D, Lunguya O, Jacobs J, Dougan G, Deborggraeve S. 2019. An African Salmonella Typhimurium ST313 sublineage with extensive drug-resistance and signatures of host adaptation. *Nat Commun* **10:**4280. https://doi.org/10.1038/s41467-019-11844-z.
93. Singletary LA, Karlinsey JE, Libby SJ, Mooney JP, Lokken KL, Tsolis RM, Byndloss MX, Hirao LA, Gaulke CA, Crawford RW, Dandekar S, Kingsley RA, Msefula CL, Heyderman RS, Fang FC. 2016. Loss of multicellular behavior in epidemic African nontyphoidal Salmonella enterica serovar Typhimurium ST313 strain D23580. *mBio* **7:**e02265-15. https://doi.org/10.1128/mBio.02265-15.
94. Zheng J, Pettengill J, Strain E, Allard MW, Ahmed R, Zhao S, Brown EW. 2014. Genetic diversity and evolution of Salmonella enterica serovar Enteritidis strains with different phage types. *J Clin Microbiol* **52:**1490–1500. https://doi.org/10.1128/JCM.00051-14.
95. Crump JA, Medalla FM, Joyce KW, Krueger AL, Hoekstra RM, Whichard JM, Barzilay EJ, Emerging Infections Program NARMS Working Group. 2011. Antimicrobial resistance among invasive nontyphoidal Salmonella enterica isolates in the United States: National Antimicrobial Resistance Monitoring System, 1996 to 2007. *Antimicrob Agents Chemother* **55:**1148–1154. https://doi.org/10.1128/AAC.01333-10.
96. Antony L, Behr M, Sockett D, Miskimins D, Aulik N, Christopher-Hennings J, Nelson E, Allard MW, Scaria J. 2018. Genome divergence and increased virulence of outbreak associated Salmonella enterica subspecies enterica serovar Heidelberg. *Gut Pathog* **10:**53. https://doi.org/10.1186/s13099-018-0279-0.
97. van den Berg RR, Dissel S, Rapallini MLBA, van der Weijden CC, Wit B, Heymans R. 2019. Characterization and whole genome sequencing of closely related multidrug-resistant Salmonella enterica serovar Heidelberg isolates from imported poultry meat in the Netherlands. *PLoS One* **14:**e0219795. https://doi.org/10.1371/journal.pone.0219795.
98. Deblais L, Lorentz B, Scaria J, Nagaraja KV, Nisar M, Lauer D, Voss S, Rajashekara G. 2018. Comparative genomic studies of Salmonella Heidelberg isolated from chicken- and turkey-associated farm environmental samples. *Front Microbiol* **9:**1841. https://doi.org/10.3389/fmicb.2018.01841.
99. Oladeinde A, Cook K, Orlek A, Zock G, Herrington K, Cox N, Lawrence JP, Hall C. 2018. Hotspot mutations and ColE1 plasmids contribute to the fitness of Salmonella Heidelberg in poultry litter. *PLoS One* **13:**e0202286. https://doi.org/10.1371/journal.pone.0202286.
100. Hoffmann M, Zhao S, Pettengill J, Luo Y, Monday SR, Abbott J, Ayers SL, Cinar HN, Muruvanda T, Li C, Allard MW, Whichard J, Meng J, Brown EW, McDermott PF. 2014. Comparative genomic analysis and virulence differences in closely related Salmonella enterica serotype Heidelberg isolates from humans, retail meats, and animals. *Genome Biol Evol* **6:**1046–1068. https://doi.org/10.1093/gbe/evu079.
101. National Antimicrobial Resistance Monitoring System for Enteric Bacteria. 2008. Veterinary isolates final report, slaughter isolates, 2006. USDA, Washington, DC.
102. Fricke WF, McDermott PF, Mammel MK, Zhao S, Johnson TJ, Rasko DA, Fedorka-Cray PJ, Pedroso A, Whichard JM, Leclerc JE, White DG, Cebula TA, Ravel J. 2009. Antimicrobial resistance-conferring plasmids with similarity to virulence plasmids from avian pathogenic Escherichia coli strains in Salmonella enterica serovar Kentucky isolates from poultry. *Appl Environ Microbiol* **75:**5963–5971. https://doi.org/10.1128/AEM.00786-09.
103. Blau DM, McCluskey BJ, Ladely SR, Dargatz DA, Fedorka-Cray PJ, Ferris KE, Headrick ML. 2005. Salmonella in dairy operations in the United States: prevalence and antimicrobial drug susceptibility. *J Food Prot* **68:**696–702. https://doi.org/10.4315/0362-028X-68.4.696.
104. Dhanani AS, Block G, Dewar K, Forgetta V, Topp E, Beiko RG, Diarra MS. 2015. Genomic comparison of non-typhoidal Salmonella enterica serovars Typhimurium, Enteritidis, Heidelberg, Hadar and Kentucky isolates from broiler chickens. *PLoS One* **10:**e0128773. https://doi.org/10.1371/journal.pone.0128773.
105. Cheng Y, Pedroso AA, Porwollik S, McClelland M, Lee MD, Kwan T, Zamperini K, Soni V, Sellers HS, Russell SM, Maurer JJ. 2015. rpoS-regulated core genes involved in the competitive fitness of Salmonella enterica serovar Kentucky in the intestines of chickens. *Appl Environ Microbiol* **81:**502–514. https://doi.org/10.1128/AEM.03219-14.
106. Joerger RD, Sartori CA, Kniel KE. 2009. Comparison of genetic and physiological properties of Salmonella enterica isolates from chickens reveals one major difference between serovar Kentucky and other serovars: response to acid. *Foodborne Pathog Dis* **6:**503–512. https://doi.org/10.1089/fpd.2008.0144.
107. Johnson TJ, Thorsness JL, Anderson CP, Lynne AM, Foley SL, Han J, Fricke WF, McDermott PF, White DG, Khatri M, Stell AL, Flores C, Singer RS. 2010. Horizontal gene transfer of a ColV plasmid has resulted in a dominant avian clonal type of Salmonella enterica serovar Kentucky. *PLoS One* **5:**e15524. https://doi.org/10.1371/journal.pone.0015524.
108. Haley BJ, Kim SW, Pettengill J, Luo Y, Karns JS, Van Kessel JAS. 2016. Genomic and evolutionary analysis of two Salmonella enterica serovar Kentucky sequence types isolated from bovine and poultry sources in North America. *PLoS One* **11:**e0161225. https://doi.org/10.1371/journal.pone.0161225.

109. Le Hello S, Hendriksen RS, Doublet B, Fisher I, Nielsen EM, Whichard JM, Bouchrif B, Fashae K, Granier SA, Jourdan-Da Silva N, Cloeckaert A, Threlfall EJ, Angulo FJ, Aarestrup FM, Wain J, Weill FX. 2011. International spread of an epidemic population of Salmonella enterica serotype Kentucky ST198 resistant to ciprofloxacin. *J Infect Dis* **204:**675–684. https://doi.org/10.1093/infdis/jir409.

110. Hawkey J, Le HS, Doublet B, Granier SA, Hendriksen RS, Fricke WF, Ceyssens PJ, Gomart C, Billman-Jacobe H, Holt KE, Weill FX. 2019. Global phylogenomics of multidrug-resistant Salmonella enterica serotype Kentucky ST198. *Microb Genom* **5:** e000269. https://doi.org/10.1099/mgen.0.000269.

111. Painter JA, Hoekstra RM, Ayers T, Tauxe RV, Braden CR, Angulo FJ, Griffin PM. 2013. Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998-2008. *Emerg Infect Dis* **19:**407–415. https://doi.org/10.3201/eid1903.111866.

112. Bullis KL. 1977. The history of avian medicine in the U.S. II. Pullorum disease and fowl typhoid. *Avian Dis* **21:**422–429. https://doi.org/10.2307/1589326.

113. Rodrigue DC, Tauxe RV, Rowe B. 1990. International increase in Salmonella enteritidis: a new pandemic? *Epidemiol Infect* **105:**21–27. https://doi.org/10.1017/S0950268800047609.

114. Edwards PR, Bruner DW, Moran AB. 1948. Further studies on the occurrence and distribution of Salmonella types in the United States. *J Infect Dis* **83:**220–231. https://doi.org/10.1093/infdis/83.3.220.

115. Guard-Petter J, Henzler DJ, Rahman MM, Carlson RW. 1997. On-farm monitoring of mouse-invasive Salmonella enterica serovar enteritidis and a model for its association with the production of contaminated eggs. *Appl Environ Microbiol* **63:**1588–1593. https://doi.org/10.1128/AEM.63.4.1588-1593.1997.

116. Guard J, Cao G, Luo Y, Baugher JD, Davison S, Yao K, Hoffmann M, Zhang G, Likens N, Bell R, Zheng J, Brown E, Allard MW. 2020. Genome sequence analysis of 91 Salmonella Enteritidis isolates from mice caught on poultry farms in the mid-1990s. *Genomics* **112:**528–544. https://doi.org/10.1016/j.ygeno.2019.04.005.

117. Feberwee A, de Vries TS, Hartman EG, de Wit JJ, Elbers AR, de Jong WA. 2001. Vaccination against Salmonella enteritidis in Dutch commercial layer flocks with a vaccine based on a live Salmonella gallinarum 9R strain: evaluation of efficacy, safety, and performance of serologic Salmonella tests. *Avian Dis* **45:**83–91. https://doi.org/10.2307/1593015.

118. Rabsch W, Hargis BM, Tsolis RM, Kingsley RA, Hinz KH, Tschäpe H, Bäumler AJ. 2000. Competitive exclusion of Salmonella enteritidis by Salmonella gallinarum in poultry. *Emerg Infect Dis* **6:**443–448. https://doi.org/10.3201/eid0605.000501.

119. Baumler AJ, Hargis BM, Tsolis RM. 2000. Tracing the origins of Salmonella outbreaks. *Science* **287:**50–52. https://doi.org/10.1126/science.287.5450.50.

120. Cogan TA, Humphrey TJ. 2003. The rise and fall of Salmonella Enteritidis in the UK. *J Appl Microbiol* **94:**114–119S. https://doi.org/10.1046/j.1365-2672.94.s1.13.x.

121. Mumma GA, Griffin PM, Meltzer MI, Braden CR, Tauxe RV. 2004. Egg quality assurance programs and egg-associated Salmonella enteritidis infections, United States. *Emerg Infect Dis* **10:**1782–1789. https://doi.org/10.3201/eid1010.040189.

122. Han J, Lynne AM, David DE, Tang H, Xu J, Nayak R, Kaldhone P, Logue CM, Foley SL. 2012. DNA sequence analysis of plasmids from multidrug resistant Salmonella enterica serotype Heidelberg isolates. *PLoS One* **7:**e51160. https://doi.org/10.1371/journal.pone.0051160.

123. Foley SL, Nayak R, Hanning IB, Johnson TJ, Han J, Ricke SC. 2011. Population dynamics of Salmonella enterica serotypes in commercial egg and poultry production. *Appl Environ Microbiol* **77:**4273–4279. https://doi.org/10.1128/AEM.00598-11.

124. Popoff MY, Bockemuhl J, Brenner FW. 2000. Supplement 1998 (no. 42) to the Kauffmann-White scheme. *Res Microbiol* **151:**63–65. https://doi.org/10.1016/S0923-2508(00)00126-1.

125. de Freitas Neto OC, Mesquita AL, de Paiva JB, Zotesso F, Berchieri Júnior A. 2008. Control of Salmonella enterica serovar Enteritidis in laying hens by inactivated Salmonella Enteritidis vaccines. *Braz J Microbiol* **39:**390–396. https://doi.org/10.1590/S1517-83822008000200034.

126. Theuß T, Woitow G, Bulang M, Springer S. 2018. Demonstration of the efficacy of a Salmonella Enteritidis live vaccine for chickens according to the current European Pharmacopoeia monograph. *Heliyon* **4:**e01070. https://doi.org/10.1016/j.heliyon.2018.e01070.

127. Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, Calteau A, Chiapello H, Clermont O, Cruveiller S, Danchin A, Diard M, Dossat C, Karoui ME, Frapy E, Garry L, Ghigo JM, Gilles AM, Johnson J, Le Bouguénec C, Lescat M, Mangenot S, Martinez-Jéhanne V, Matic I, Nassif X, Oztas S, Petit MA, Pichon C, Rouy Z, Ruf CS, Schneider D, Tourret J, Vacherie B, Vallenet D, Médigue C, Rocha EP, Denamur E. 2009. Organised genome dynamics in the Escherichia coli species results in highly diverse adaptive paths. *PLoS Genet* **5:**e1000344. https://doi.org/10.1371/journal.pgen.1000344.

128. Jacobsen A, Hendriksen RS, Aaresturp FM, Ussery DW, Friis C. 2011. The Salmonella enterica pan-genome. *Microb Ecol* **62:**487–504. https://doi.org/10.1007/s00248-011-9880-1.

129. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM. 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome." *Proc Natl Acad Sci U S A* **102:**13950–13955. https://doi.org/10.1073/pnas.0506758102.

130. Tettelin H, Riley D, Cattuto C, Medini D. 2008. Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* **11:**472–477. https://doi.org/10.1016/j.mib.2008.09.006.

131. Laing CR, Whiteside MD, Gannon VPJ. 2017. Pan-genome analyses of the species Salmonella enterica, and identification of genomic markers predictive for species. *Front Microbiol* **8:**1345. https://doi.org/10.3389/fmicb.2017.01345.

132. Worley J, Meng J, Allard MW, Brown EW, Timme RE. 2018. Salmonella enterica phylogeny based on whole-genome sequencing reveals two new clades and novel patterns of horizontally acquired genetic elements. *mBio* **9:**e02303-18. https://doi.org/10.1128/mBio.02303-18.

133. Haneda T, Ishii Y, Danbara H, Okada N. 2009. Genome-wide identification of novel genomic islands that contribute to Salmonella virulence in mouse systemic infection. *FEMS Microbiol Lett* **297:**241–249. https://doi.org/10.1111/j.1574-6968.2009.01686.x.

134. Espinoza RA, Silva-Valenzuela CA, Amaya FA, Urrutia ÍM, Contreras I, Santiviago CA. 2017. Differential roles for pathogenicity islands SPI-13 and SPI-8 in the interaction of Salmonella Enteritidis and Salmonella Typhi with murine and human macrophages. *Biol Res* **50:**5. https://doi.org/10.1186/s40659-017-0109-8.

135. Nieto PA, Pardo-Roa C, Salazar-Echegarai FJ, Tobar HE, Coronado-Arrázola I, Riedel CA, Kalergis AM, Bueno SM. 2016. New insights about excisable pathogenicity islands in Salmonella and their contribution to virulence. *Microbes Infect* **18:**302–309. https://doi.org/10.1016/j.micinf.2016.02.001.
136. Jennings E, Thurston TLM, Holden DW. 2017. Salmonella SPI-2 type III secretion system effectors: molecular mechanisms and physiological consequences. *Cell Host Microbe* **22:**217–231. https://doi.org/10.1016/j.chom.2017.07.009.
137. Hautefort I, Thompson A, Eriksson-Ygberg S, Parker ML, Lucchini S, Danino V, Bongaerts RJ, Ahmad N, Rhen M, Hinton JC. 2008. During infection of epithelial cells Salmonella enterica serovar Typhimurium undergoes a time-dependent transcriptional adaptation that results in simultaneous expression of three type 3 secretion systems. *Cell Microbiol* **10:**958–984. https://doi.org/10.1111/j.1462-5822.2007.01099.x.
138. Olekhnovich IN, Kadner RJ. 2002. DNA-binding activities of the HilC and HilD virulence regulatory proteins of Salmonella enterica serovar Typhimurium. *J Bacteriol* **184:**4148–4160. https://doi.org/10.1128/JB.184.15.4148-4160.2002.
139. Olekhnovich IN, Kadner RJ. 2007. Role of nucleoid-associated proteins Hha and H-NS in expression of Salmonella enterica activators HilD, HilC, and RtsA required for cell invasion. *J Bacteriol* **189:**6882–6890. https://doi.org/10.1128/JB.00905-07.
140. Kim K, Golubeva YA, Vanderpool CK, Slauch JM. 2019. Oxygen-dependent regulation of SPI1 type three secretion system by small RNAs in Salmonella enterica serovar Typhimurium. *Mol Microbiol* **111:**570–587. https://doi.org/10.1111/mmi.14174.
141. Ryan D, Mukherjee M, Suar M. 2017. The expanding targetome of small RNAs in Salmonella Typhimurium. *Biochimie* **137:**69–77. https://doi.org/10.1016/j.biochi.2017.03.005.
142. Velge P, Wiedemann A, Rosselin M, Abed N, Boumart Z, Chaussé AM, Grépinet O, Namdari F, Roche SM, Rossignol A, Virlogeux-Payant I. 2012. Multiplicity of Salmonella entry mechanisms, a new paradigm for Salmonella pathogenesis. *Microbiologyopen* **1:**243–258. https://doi.org/10.1002/mbo3.28.
143. Bao H, Zhao JH, Zhu S, Wang S, Zhang J, Wang XY, Hua B, Liu C, Liu H, Liu SL. 2019. Genetic diversity and evolutionary features of type VI secretion systems in Salmonella. *Future Microbiol* **14:**139–154. https://doi.org/10.2217/fmb-2018-0260.
144. Baumler AJ, Gilde AJ, Tsolis RM, van der Velden AW, Ahmer BM, Heffron F. 1997. Contribution of horizontal gene transfer and deletion events to development of distinctive patterns of fimbrial operons during evolution of Salmonella serotypes. *J Bacteriol* **179:**317–322. https://doi.org/10.1128/JB.179.2.317-322.1997.
145. Spector MP, Kenyon WJ. 2012. Resistance and survival strategies of Salmonella enterica to environmental stresses. *Food Res Int* **45:**455–481. https://doi.org/10.1016/j.foodres.2011.06.056.
146. Jacobsen CS, Bech TB. 2012. Soil survival of Salmonella and transfer to freshwater and fresh produce. *Food Res Int* **45:**557–566. https://doi.org/10.1016/j.foodres.2011.07.026.
147. Hruby CE, Soupir ML, Moorman TB, Pederson C, Kanwar R. 2018. Salmonella and fecal indicator bacteria survival in soils amended with poultry manure. *Water Air Soil Pollut* **229:**32. https://doi.org/10.1007/s11270-017-3667-z.
148. Greene S, Daly E, Talbot E, Demma L, Holzbauer S, Patel NJ, Hill TA, Walderhaug MO, Hoekstra RM, Lynch MF, Painter JA. 2008. Recurrent multistate outbreak of Salmonella Newport associated with tomatoes from contaminated fields, 2005. *Epidemiol Infect* **136:**157–165. https://doi.org/10.1017/S095026880700859X.
149. Maserati A, Fink RC, Lourenco A, Julius ML, Diez-Gonzalez F. 2017. General response of Salmonella enterica serovar Typhimurium to desiccation: a new role for the virulence factors sopD and sseD in survival. *PLoS One* **12:**e0187692. https://doi.org/10.1371/journal.pone.0187692.
150. Davis S, Pettengill JB, Luo Y, Payne J, Shpuntoff A, Rand H, Strain E. 2015. CFSAN SNP pipeline: an automated method for constructing SNP matrices from nextgeneration sequence data. *Peer J Comput Sci* **1:**e20. https://doi.org/10.7717/peerj-cs.20.
151. Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11:**595. https://doi.org/10.1186/1471-2105-11-595.
152. den Bakker HC, Moreno Switt AI, Govoni G, Cummings CA, Ranieri ML, Degoricija L, Hoelzer K, Rodriguez-Rivera LD, Brown S, Bolchacova E, Furtado MR, Wiedmann M. 2011. Genome sequencing reveals diversification of virulence factor content and possible host adaptation in distinct subpopulations of Salmonella enterica. *BMC Genomics* **12:**425. https://doi.org/10.1186/1471-2164-12-425.
153. Lienau EK, Blazar JM, Wang C, Brown EW, Stones R, Musser S, Allard MW. 2013. Whole genome sequence phylogenetic analysis identifies gene gains that define Salmonella enterica subspecies I. *PLoS One* **8:**e76821. https://doi.org/10.1371/journal.pone.0076821.
154. NARMS. 2020. Global resistome data. https://www.fda.gov/animal-veterinary/national-antimicrobial-resistance-monitoring-system/global-salmonella-resistome-data.
155. Hardt WD, Urlaub H, Galán JE. 1998. A substrate of the centisome 63 type III protein secretion system of Salmonella typhimurium is encoded by a cryptic bacteriophage. *Proc Natl Acad Sci U S A* **95:**2574–2579. https://doi.org/10.1073/pnas.95.5.2574.
156. Moreno Switt AIM, Sulakvelidze A, Wiedmann M, Kropinski AM, Wishart DS. 2015. Salmonella phages and prophages: genomics, taxonomy, and applied aspects. *Methods Mol Biol* **1225:**237–287. https://doi.org/10.1007/978-1-4939-1625-2_15.
157. NCBI. 2020. Pathogen detection. https://www.ncbi.nlm.nih.gov/pathogens/.
158. Gonzalez-Escalona N, Allard MW, Brown EW, Sharma S, Hoffmann M. 2019. Nanopore sequencing for fast determination of plasmids, phages, virulence markers, and antimicrobial resistance genes in Shiga toxin-producing Escherichia coli. *PLoS One* **14:**e0220494. https://doi.org/10.1371/journal.pone.0220494.
159. Marchesi JR, Ravel J. 2015. The vocabulary of microbiome research: a proposal. *Microbiome* **3:**31. https://doi.org/10.1186/s40168-015-0094-5.
160. Ottesen AR, Gonzalez A, Bell R, Arce C, Rideout S, Allard M, Evans P, Strain E, Musser S, Knight R, Brown E, Pettengill JB. 2013a. Co-enriching microflora associated with culture based methods to detect Salmonella from tomato phyllosphere. *PLoS One* **8:**e73079. https://doi.org/10.1371/journal.pone.0073079.
161. Bell RL, Jarvis KG, Ottesen AR, McFarland MA, Brown EW. 2016. Recent and emerging innovations in Salmonella detection: a food and environmental perspective. *Microb Biotechnol* **9:**279–292. https://doi.org/10.1111/1751-7915.12359.
162. Carleton HA, Besser J, Williams-Newkirk AJ, Huang A, Trees E, Gerner-Smidt P. 2019. Metagenomic approaches for public health surveillance of foodborne infections: opportunities and challenges. *Foodborne Pathog Dis* **16:**474–479. https://doi.org/10.1089/fpd.2019.2636.
163. Leonard SR, Mammel MK, Lacher DW, Elkins CA. 2015. Application of metagenomic sequencing to food safety: detection of shiga toxin-producing Escherichia coli on fresh bagged spinach. *Appl Environ Microbiol* **81:**8183–8191. https://doi.org/10.1128/AEM.02601-15.
164. Ottesen A, Ramachandran P, Reed E, White JR, Hasan N, Subramanian P, Ryan G, Jarvis K, Grim C, Daquiqan N, Hanes D,

Allard M, Colwell R, Brown E, Chen Y. 2016. Enrichment dynamics of Listeria monocytogenes and the associated microbiome from naturally contaminated ice cream linked to a listeriosis outbreak. *BMC Microbiol* **16:**275. https://doi.org/10.1186/s12866-016-0894-1.

165. Jarvis KG, White JR, Grim CJ, Ewing L, Ottesen AR, Beaubrun JJ, Pettengill JB, Brown E, Hanes DE. 2015. Cilantro microbiome before and after nonselective pre-enrichment for Salmonella using 16S rRNA and metagenomic sequencing. *BMC Microbiol* **15:**160. https://doi.org/10.1186/s12866-015-0497-2.

166. Hyeon J-Y, Li S, Mann DA, Zhang S, Li Z, Chen Y, Deng X. 2017. Quasimetagenomics-based and real-time-sequencing-aided detection and subtyping of Salmonella enterica from food samples. *Appl Environ Microbiol* **84:**e02340-17. https://doi.org/10.1128/AEM.02340-17.

167. Forghani F, Li S, Zhang S, Mann DA, Deng X, den Bakker HC, Diez-Gonzalez F. 2020. Salmonella enterica and Escherichia coli in wheat flour: detection and serotyping by a quasimetagenomic approach assisted by magnetic capture, multiple-displacement amplification, and real-time sequencing. *Appl Environ Microbiol* **86:** e00097-20. https://doi.org/10.1128/AEM.00097-20.

168. Grim CJ, Daquigan N, Lusk PT, Ottesen AR, White JR, Jarvis KG. 2017. High-resolution microbiome profiling for detection and tracking of Salmonella enterica. *Front Microbiol* **8:**1–10. https://doi.org/10.3389/fmicb.2017.01587.

169. Langlais M, Thibodeau A, Philippe F. 2019. A metagenomic analysis of the pre-enrichment step for the isolation of Salmonella spp. from pig feces. *J Microbiol Methods* **157:**43–46. https://doi.org/10.1016/j.mimet.2018.12.014.

170. Hyeon JY, Mann DA, Townsend AM, Deng X. 2018. Quasimetagenomic analysis of Salmonella from food and environmental samples. *J Vis Exp* **140:**e58612. https://doi.org/10.3791/58612.

171. Deshpande SV, Reed TM, Sullivan RF, Kerkhof LJ, Beigel KM, Wade MM. 2019. Offline next generation metagenomics sequence analysis using MinION detection software (MINDS). *Genes* **10:**578. https://doi.org/10.3390/genes10080578.

172. Zhang S, Li S, Gu W, den Bakker H, Boxrud D, Taylor A, Roe C, Driebe E, Engelthaler DM, Allard M, Brown E, McDermott P, Zhao S, Bruce BB, Trees E, Fields PI, Deng X. 2019. Zoonotic source attribution of Salmonella enterica serotype typhimurium using genomic surveillance data, United States. *Emerg Infect Dis* **25:**82–91. https://doi.org/10.3201/eid2501.180835.

173. Wang Y, Pettengill JB, Pightling A, Timme R, Allard MW, Strain E, Rand H. 2018. Genetic diversity of Salmonella and Listeria isolates from food facilities. *J Food Prot* **81:**2082–2089. https://doi.org/10.4315/0362-028X.JFP-18-093.

174. Sanaa M, Pouillot R, Garces-Vega FJ, Strain E, Van Doren JM. 2018. GenomeGraphR: a user-friendly open-source web application for foodborne pathogen whole genome sequencing data integration, analysis, and visualization. *bioRxiv*. https://doi.org/10.1101/495309.

175. Richardson LC, Bazaco MC, Parker CC, Dewey-Mattia D, Golden N, Jones K, Klontz K, Travis C, Kufel JZ, Cole D. 2017. An updated scheme for categorizing foods implicated in foodborne disease outbreaks: a tri-agency collaboration. *Foodborne Pathog Dis* **14:**701–710. https://doi.org/10.1089/fpd.2017.2324.

176. Lienau EK, Strain E, Wang C, Zheng J, Ottesen AR, Keys CE, Hammack TS, Musser SM, Brown EW, Allard MW, Cao G, Meng J, Stones R. 2011. Identification of a salmonellosis outbreak by means of molecular sequencing. *N Engl J Med* **10:**981–982. https://doi.org/10.1056/NEJMc1100443.

177. Allard MW, Brown EW. 2020. Epidemiology needs more interdisciplinary teams with expertise in molecular systematics, public health and food safety. *Cladistics* **36:**345–347. https://doi.org/10.1111/cla.12428.

178. Allard MW, Luo T, Strain E, Li C, Keys CE, Son I, Stones R, Musser SM, Brown EW. 2012. High resolution clustering of *Salmonella enterica* serovar Montevideo strains using a next-generation sequencing approach. *BMC Genomics* **13:**32. https://doi.org/10.1186/1471-2164-13-32.

179. Allard MW, Luo Y, Strain E, Pettengill J, Timme R, Wang C, Li C, Keys CE, Zheng J, Stones R, Wilson MR, Musser SM, Brown EW. 2013. On the evolutionary history, population genetics and diversity among isolates of Salmonella Enteritidis PFGE pattern JEGX01.0004. *PLoS One* **8:**e55254. https://doi.org/10.1371/journal.pone.0055254.

180. Timme RE, Sanchez Leon M, Allard MW. 2019. Utilizing the public GenomeTrakr database for foodborne pathogen traceback. *Methods Mol Biol* **1918:**201–212. https://doi.org/0.1007/978-1-4939-9000-9_17.

181. Timme RE, Strain E, Baugher J, Davis S, Gonzalez-Escalona N, Sanchez-Leon M, Allard MW, Brown EW, Tallent S, Rand H. 2019b. Phylogenomic pipeline validation for foodborne pathogen disease surveillance. *J Clin Microbiol* **57:**e01816-18. https://doi.org/10.1128/JCM.01816-18.