SUPPLEMENT ARTICLE

Molecular Approaches to Understanding Transmission and Source Attribution in Nontyphoidal *Salmonella* and Their Application in Africa

Alison E. Mather,[1] Timothy G. Vaughan,[2,3] and Nigel P. French[4]

[1]Department of Veterinary Medicine, University of Cambridge, United Kingdom; [2]Department of Computer Science, University of Auckland, [3]Allan Wilson Centre for Molecular Ecology and Evolution, and [4]mEpiLab, Infectious Disease Research Centre, Massey University, Palmerston North, New Zealand

Nontyphoidal *Salmonella* (NTS) is a frequent cause of diarrhea around the world, yet in many African countries it is more commonly associated with invasive bacterial disease. Various source attribution models have been developed that utilize microbial subtyping data to assign cases of human NTS infection to different animal populations and foods of animal origin. Advances in molecular microbial subtyping approaches, in particular whole-genome sequencing, provide higher resolution data with which to investigate these sources. In this review, we provide updates on the source attribution models developed for *Salmonella*, and examine the application of whole-genome sequencing data combined with evolutionary modeling to investigate the putative sources and transmission pathways of NTS, with a focus on the epidemiology of NTS in Africa. This is essential information to decide where, what, and how control strategies might be applied most effectively.

Keywords. nontyphoidal *Salmonella*; whole-genome sequencing; source attribution; Africa; bacteremia.

Nontyphoidal *Salmonella* (NTS), which includes all serotypes of *Salmonella enterica* except Typhi, Paratyphi A, Paratyphi B, and Paratyphi C, is a common cause of bacteremia worldwide, particularly in Africa, with the global burden of disease estimated at 3.4 million cases per year [1]. The majority of human NTS infections in industrialized countries cause enterocolitis and are believed to come through the food chain [2]. Source attribution, where human cases are partitioned to specific animal reservoirs and food-related sources [3], has been widely employed as a tool to prioritize interventions [4]. Both epidemiological and microbiological data have been used in source attribution models; other methods include outbreak analysis and intervention studies [3]. The most common source attribution approaches utilize microbial subtyping data, comparing the identities and frequencies of subtypes found in human and other sources.

There are many different methods of microbial subtyping (Figure 1): phenotypic methods, which include serotyping and antimicrobial resistance (AMR) typing, whereas genotypic methods include pulsed-field gel electrophoresis (PFGE) and multilocus sequence typing (MLST). The application of MLST, for example, has allowed the identification of a particular *Salmonella* Typhimurium variant, sequence type (ST) 313, as one of the most common causes of invasive NTS in sub-Saharan Africa, one that is rarely seen outside this region [5]. Barco et al [6] reviewed the advantages and disadvantages of each of these subtyping methods, and their application to *Salmonella* source attribution. Of the molecular tools available, whole-genome sequencing (WGS) provides the greatest resolution for microbial subtyping; the additional information it provides prompted the development of new analytical tools to examine the evolution and dissemination of bacteria and identify the sources and transmission pathways of

Correspondence: Alison E. Mather, BSc, MSc, PhD, Department of Veterinary Medicine, University of Cambridge, Madingley Road, Cambridge CB3 0ES, UK (am2382@cam.ac.uk).

Clinical Infectious Diseases® 2015;61(S4):S259–65

DOI: 10.1093/cid/civ727

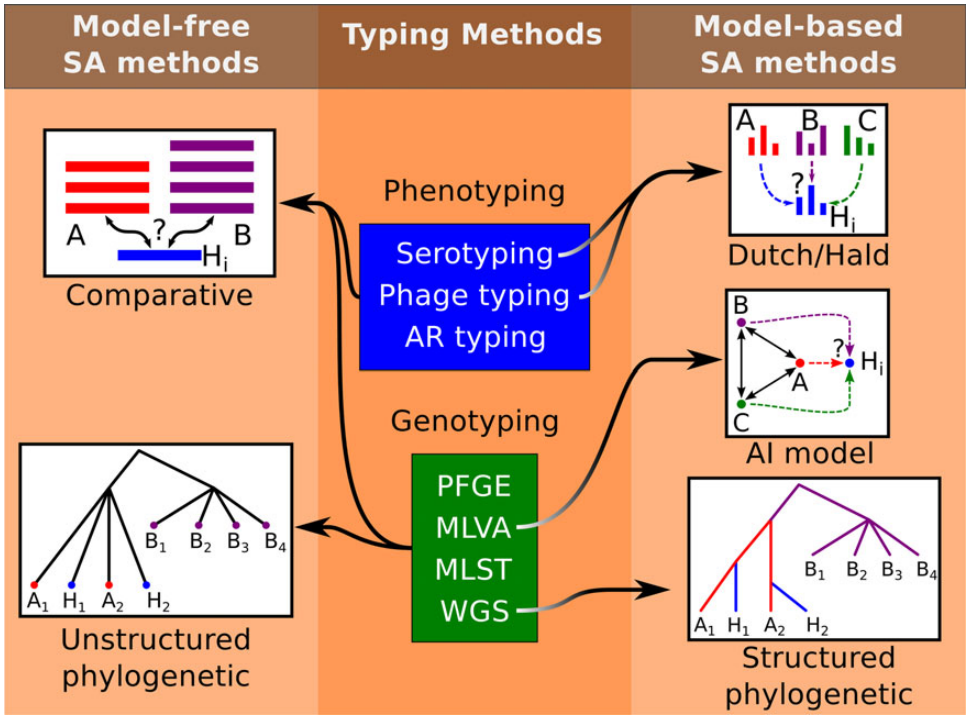

Figure 1. Schematic representation of nontyphoidal *Salmonella* (NTS) source attribution methods discussed in this review. Methods in the left-hand column do not explicitly model the connection between human and animal derived isolates, in contrast to methods on the right. The center column lists typing schemes and the NTS source attribution methods in which they have been used. The comparative methods involve direct visual comparison between human (H) isolate types and the types of isolates derived from putative sources (A, B, . . .). The Dutch/Hald methods quantitatively compare distributions of isolate types obtained from putative sources with those obtained from humans. The asymmetric island method is a population genetic approach that allows for the movement of NTS between putative sources. Unstructured phylogenetic methods use genotypes to produce a dendrogram, the shape of which is used to qualitatively assess potential sources. Structured phylogenetic approaches jointly infer both the phylogenetic relationship of genotyped isolates and the transmission history of the sampled data, allowing for quantitative assessment of potential sources. Abbreviations: AI, asymmetric island; AR, antimicrobial resistance; MLST, multilocus sequence typing; MLVA, multilocus variable-number tandem repeat analysis; PFGE, pulsed-field gel electrophoresis; SA, source attribution; WGS, whole-genome sequencing.

bacterial pathogens, including NTS. As the cost of WGS, particularly for bacteria, is decreasing and is ever more feasible for very large numbers of samples, we focus on this subtyping method.

First, we examine the various source attribution models that have been developed and applied to NTS. Second, we cover studies that have used non-WGS molecular approaches to examine multihost source attribution of NTS in Africa. Finally, we cover the application of WGS to study source attribution, with the focus on NTS in Africa.

SOURCE ATTRIBUTION MODELS THAT UTILIZE ROUTINELY COLLECTED MICROBIAL SUBTYPING DATA

Here we review methods applied to NTS subtyping data where the distribution of subtypes isolated from human clinical samples is compared with the distribution from animal fecal samples and food products. These models provide estimates of the number of human NTS cases attributable to each "source" (usually defined as an animal reservoir or a food product), but do not consider human-to-human transmission. For more detailed reviews of these methods, see [4, 6].

A Bayesian model was developed by Hald et al [7] to estimate the contribution of different food sources to the burden of NTS in Denmark, using serotyping data collected as part of routine surveillance. The Hald model is similar to the "Dutch model" [8], which uses a frequentist approach to attributing the relative contribution of different sources by comparing the frequency of human cases caused by different NTS types with their relative prevalence in animal reservoirs and food products. The original Hald model and a modified version of the Dutch model [8] include weights for the amount of food consumed, and provide estimates of the number of NTS cases attributable to, for example, eggs, broiler chicken, pork, and beef. The inclusion of food weights has been shown subsequently to be important for the

Table 1. Studies Using Non–Whole Genome Sequence Molecular Methods to Examine Nontyphoidal *Salmonella* in Multihost Populations in Africa

Population Examined (No. of Isolates)	Country	Molecular Method Used	Conclusion(s)	Reference
Humans (14), domestic animals (21)	The Gambia	MLST	No relatedness in genotypes from humans and associated animals	Dione et al [15]
Humans (151), animals/ environment/food (78)	Kenya	Plasmid typing, PFGE	No significant relatedness between salmonellae from humans and from animals, environment, food; common plasmid in MDR *Salmonella* Typhimurium and *Salmonella* Enteritidis from humans and some chickens	Kariuki et al [16]
Humans (225), animals/ environment/food (10)	Kenya	Plasmid typing, PFGE	Greater similarity of salmonellae between pediatric cases and their family members than with those from animals, environment, food	Kariuki et al [17]
Humans (19), animals (383)	Burkina Faso	PFGE	*Salmonella* Typhimurium isolates from poultry (n = 4) and humans (n = 13) clustered by PFGE patterns; some clustering of domestic animal and hedgehog isolates	Kagambega et al [18]
Humans (196), captive wild animals (5)	South Africa	PFGE	Similarities between *Salmonella* Enteritidis isolates from humans and captive wild animals	Smith et al [19]

Abbreviations: MDR, multidrug resistant; MLST, multilocus sequence typing; PFGE, pulsed-field gel electrophoresis.

Dutch model but not the Hald model, allowing nonfood sources such as wildlife and water to be considered in the latter [9]. Modified versions of the Hald model have improved performance and robustness (eg, [10]), enabling it to be applied to other foodborne zoonoses (eg, [10]), and modeling temporal trends in attribution [11].

More recently, an alternative source attribution model, the asymmetric island (AI) model [12], was applied to NTS genotyped using multilocus variable-number tandem repeat analysis (MLVA) [13]. In contrast to the Dutch and Hald models, the AI model explores the evolutionary history of the set of human and source isolates and, because it is based on different underlying assumptions to the Dutch and Hald models, it can be used for high-resolution data and in multimodel comparisons [11, 13].

In a further development, the AI and modified Dutch models have been used to refine case-control studies by assigning cases to putative sources before modeling the relative contribution of risk factors to each "source-assigned" set of cases. These studies provide greater power to identify the contribution of more precisely defined risk factors, such as the consumption of raw or undercooked eggs (eg, [14]).

These models are applicable to any disease manifestation, and could be applied to NTS in Africa; cases could be analyzed separately according to whether they are diarrheal or invasive, or considered in a combined analysis to assess whether transmission pathways or sources differ according to disease outcome. The source-assigned case-control study can be a useful way of looking at more precisely defined risk factors for subsets of outcomes, and has the advantage that sources, pathways, and risk factors, the latter 2 encompassing modes of transmission, can be considered together.

NON-WGS MOLECULAR APPROACHES FOR MULTI–HOST SPECIES SOURCE ATTRIBUTION IN NTS IN AFRICA

Source attribution models as described above have not yet been applied to NTS in Africa, but a number of studies using molecular techniques other than WGS, predominantly PFGE or MLST, have examined the role of animal reservoirs (Table 1). In these studies, dendrograms demonstrating the genetic similarity of the isolates were constructed using various methods to cluster the PFGE patterns or concatenated MLST allele sequences [15].

With the exception of one study [18], the predominant observation from these studies is that the salmonellae found in humans were different from those found in animals, although the numbers of isolates examined were generally low. However, these results do not exclude animals as a reservoir of human NTS infections. In the study comparing *Salmonella* Enteritidis from humans and captive wild animals, the similarity of the isolates was suggested to be due to a common dietary component of poultry meat [19]. Parsons et al [20] demonstrated that *Salmonella* Typhimurium ST313, described as adapted to humans [5], can infect chickens and cause invasive disease. Therefore, there is the potential for an unrecognized animal, potentially poultry, reservoir of NTS.

WGS AND SOURCE ATTRIBUTION

In this section, we outline various techniques applied to WGS data in the pursuit of source attribution. Where relevant, we focus on examples where these techniques have been applied

to WGS data of NTS in Africa. In the context of NTS disease in Africa, the human population is believed to be a greater source of NTS than any animal reservoir [17], and therefore, in this review, we are expanding the scope of source attribution to include the contribution of human-to-human transmission and the role of geographical location.

Comparative Genomics

Comparative genomics is an approach in which WGS assemblies of different organisms are directly compared. This can include identifying genes or genetic elements that are present, absent or rearranged between genomes, and a comparison of gene numbers and genome size. In doing so, inferences about the evolutionary history of the bacteria can be made. However, there is always a risk that the particular genomes being compared are not representative of the species or subtype of interest. This approach has proven useful in identifying genetic elements in NTS associated with adaptation to different hosts, which often manifests as gene acquisition and/or functional gene loss. By comparing the genomes of a host generalist, *Salmonella* Enteritidis, and host specialists, *Salmonella* Gallinarum (birds) and *Salmonella* Dublin (cattle), Langridge et al [21] identified that increasing host specialization was associated with gene degradation, particularly in metabolic pathways. These patterns may help identify other serotypes under the process of host specialization, and particular patterns of gene degradation may help source attribution analyses by identifying potential reservoir populations.

Inference by Phylogeny

A common use of WGS data is in the estimation of transmission chains, often assessed through the inspection of a phylogenetic tree, generated from single-nucleotide polymorphism (SNP) information and the genetic distance between isolates. Different methods of tree construction incorporate various models of evolution, including neighbor-joining, maximum likelihood, or Bayesian approaches. Visual examination of the degree to which isolates from different backgrounds cluster can help infer sources and transmission. The utility of this approach has been demonstrated through the investigation of an outbreak of *Salmonella* Heidelberg in the United States involving contaminated ground turkey [22]; the phylogenetic tree illustrated that outbreak isolates were more similar to each other than to nonoutbreak isolates, even though they clustered more closely with nonoutbreak isolates by PFGE pattern and AMR profile. Thus, the finer resolution of the SNP data provided a more accurate indication of the isolates' epidemiological relatedness than the other typing methods.

Importantly, even if direct transmission pathways are present, without further information, it is difficult to ascertain the direction of transmission—clustering of similar isolates from different backgrounds is not sufficient to attribute source. More formal methods of integrating WGS data and informative epidemiological data have been developed (eg, [23]), which can allow for more robust inference.

Ancestral State Reconstruction

WGS data have also been used to evaluate historical transmission, looking at the entire evolutionary history of an organism. Ancestral state reconstruction, which can be achieved through parsimony- or likelihood-based approaches, involves assigning a character state to the ancestral organisms in a phylogenetic tree, represented by the internal nodes and branches [24]. This character may represent geographical location, host type, disease outcome, or any other state that potentially partitions the pathogen population. While the parsimony methods aim to minimize the overall number of state changes, likelihood-based approaches allow the incorporation of different models of evolution. Many of these models assume that samples are collected randomly from the pathogen metapopulation without regard for the value of the character state. Violations of this assumption could lead to incorrect inferences regarding the type state at the root of the tree, and may therefore bias transmission and source attribution studies. Alternative methods based on the structured coalescent (eg, [25]) are not reliant on this assumption, and additionally provide estimates of host-specific effective pathogen population sizes. Likelihood-based methods also exist that are capable of drawing direct inferences regarding important epidemiological parameters such as the basic reproduction number, R_0 [26].

One approach has been to apply discrete phylogenetic diffusion models within a time-scaled Bayesian framework, allowing the reconstruction of ancestral states incorporating a variety of population size change models and evolutionary clocks [27]. This method was used to estimate the cross-species transmission dynamics of *Salmonella* Typhimurium DT104 in animals and humans in Scotland [28]. By reconstructing the ancestral host population states throughout the tree and calculating Markov jumps, which provide estimates of unobserved transmission events, relatively few cross-species transmissions were identified in this setting. Markov rewards, which estimate the amount of time the model spent in each host state, can be used to infer which populations may act as sources and which may act as sinks, that is, receiving immigration of the organism of interest from the source population. In this study [28], the Markov rewards indicated that the model spent significantly more time in the human state, suggesting that the local animal population was unlikely to be the major source of DT104 for the humans.

WGS and NTS Source Attribution in an African Context

The source attribution and WGS examples described above have been conducted predominantly in industrialized countries.

There are not as many examples utilizing WGS in an African NTS context, but where it has been applied, it has provided powerful insights into the epidemiology of NTS in this region.

Comparative Genomics

This approach has used to understand how NTS in Africa has evolved and adapted. Although *Salmonella* Typhimurium in general has a wide host range and causes enterocolitis, there are several subtypes that are more host-restricted [21]. In sub-Saharan Africa, NTS is a frequent cause of invasive disease, in particular, ST313. [5]. Detailed investigation into the genomes of the ST313 variant indicated a genomic signature of host-adaptation similar to that observed in *Salmonella* Typhi, *Salmonella* Paratyphi, and *Salmonella* Gallinarum [5, 29]. This is evidenced by chromosomal degradation in terms of increased numbers of pseudogenes and deletions. This pattern of host adaptation, and the lack of occurrence of ST313 in nonhuman hosts, suggests that human-to-human transmission is the predominant pathway for this pathogen [29].

Inference by Phylogeny

This approach has been applied to examine NTS from human immunodeficiency virus (HIV)–infected and -uninfected adults and children in Malawi [30], where a lack of clustering by age or HIV status suggested that the patients had acquired their NTS from a common source, and/or that NTS was freely circulating among this patient population. In a study of invasive *Salmonella* Bovismorbificans from the same region [31], strains of this organism with highly conserved genomes could be found in the Malawian bacteremia cases and also in 4 isolates from animals in the United Kingdom, despite the human and animal cases being separated by >2 decades and different continents. This further illustrates that similarity of organisms infecting different host populations does not necessarily indicate a causal link between them.

Ancestral State Reconstruction

This method has been used in a manner of source attribution in terms of geography, likely representing human-to-human transmission, by investigating the country origins of the predominant invasive clone in sub-Saharan Africa, *Salmonella* Typhimurium ST313. Okoro et al [32] demonstrated that the majority of these cases were caused by 2 distinct genetic lineages that emerged relatively recently, and were temporally associated with the HIV epidemic. By applying a discrete phylogenetic diffusion model to isolates from 7 different sub-Saharan countries, the authors proposed the ancestral origin for one of the lineages as Malawi and, for the second lineage, the Democratic Republic of the Congo. Numerous estimated transmissions between countries for both lineages were also estimated, followed by local expansion [32]. This study confirmed an epidemic in this region of invasive NTS caused by ST313, with the greater resolution of WGS allowing this ST to be separated into 2 lineages with different evolutionary histories; identification of the pathways of transmission may help inform control strategies.

Considerations for Transmission or Source Attribution Studies With WGS

There are several factors that must be carefully considered when undertaking a source attribution analysis using WGS data. The first relates to the organism of interest itself. Phylogenetic methods naturally depend on a robust phylogeny, which is constructed using the genetic diversity between isolates. With highly similar genomes, such as those that might be expected during an outbreak, it is difficult to identify true phylogenetic relationships; if multiple relationships between isolates are possible, techniques that do not take into account this uncertainty may provide misleading inferences. This is where additional information, such as the sampling date, may be particularly useful. Another potential issue is the diversity of infecting organisms within a host. For diagnostic purposes, typically only one colony of an isolated bacterium is selected for further testing. For acute infections that do not feature a component of long or asymptomatic carriage, the bacterial within-host diversity would be expected to be relatively low. However, for some bacteria such as *Staphylococcus aureus* or *Mycobacterium tuberculosis*, where long periods of latency or carriage can occur, individuals may be infected with a "cloud of diversity", and the population bottlenecks that correspond to a transmission event may make tracing that transmission difficult [33]. What is also difficult to define is a cutoff for the number of SNPs that defines the same strain; in a study that examined 141 *S. aureus* isolates from a single individual (a dog), 78% of isolates were differentiated from each other by ≤2 SNPs [33]. However, the human in that study who was the source of the dog's predominant infection was co-colonized with 2 clades of *S. aureus*; thus, characterization of a single isolate would have only had a 45% chance of correctly identifying this individual as the source.

It is also important to consider whether or not a transmission has occurred. With invasive NTS disease in sub-Saharan Africa, patients typically experience multiple recurrences [34]. Using phylogenetic information derived from WGS, one study found that 78% of these recurrences were due to recrudescence, rather than reinfection [34], and that 13 of the 14 immunocompromised patients experienced recrudescence events. In transmission studies covering short time periods, this phenomenon should not be forgotten, as it could potentially artificially increase the estimated numbers of transmissions.

One of the most critical issues in conducting source attribution, regardless of the typing or analytical method employed, is that of sample selection. If the true reservoirs or sources are not adequately and robustly sampled, then incorrect inferences can

be drawn. At a smaller scale, key individuals within the transmission chain may not be included in the sample—these individuals have been termed "epidemiological dark matter" [35]. This can be extended into estimating the effect of a "ghost population," a reservoir for which no sampling data are available [36]. The impact these gaps have on estimating sources and sinks of infection will depend on whether the scale being examined is local and recent, or over a bacterium's evolutionary history, where sampling is not expected to be as comprehensive. Various methods of estimating the number of gaps in transmission chains have been developed and applied to different pathogens (eg, [37]). Modeling the sensitivity of cross-species transmission estimates to various factors has suggested that limited sampling, unbalanced sampling, recent introduction of the pathogen, and low genetic diversity can potentially bias the estimates of such transmission [38].

Finally, it is important to remember that all of the techniques described above are only powerful tools to identify transmission pathways when there are good epidemiological data available, including spatial, temporal, and host characteristics. The information provided by WGS is thus complementary to epidemiological data, rather than an outright replacement. In resource-poor settings, it may not be possible to carry out WGS on sufficiently large and representative samples to allow inference for transmission and source attribution. However, data from multiple WGS studies can be used to identify markers for host association [39], which in turn could be used to develop cheaper, rapid assays for deploying in Africa and similar settings.

CONCLUSIONS

In this article, we have provided updates on other reviews of source attribution methods for NTS [4, 6]. What we have added to these discussions is an expanded consideration of the use of WGS, and how it can and has been used to examine NTS in different host populations. We focus on NTS in Africa, where the sources and transmission patterns are not currently well characterized, and unlike outside of Africa, human-to-human transmission likely plays a significant role, particularly for *Salmonella* Typhimurium ST313. However, to exclude definitively an animal reservoir of NTS, there is a need to conduct larger, more extensive surveys in animal populations, and also in other potential reservoirs, such as the environment. The choice of analytical approach will depend on resources and the availability of routinely collected surveillance data. If the disease is endemic and there is robust surveillance of multiple potential reservoirs, the Hald or AI models using serotypes or MLST profiles would be informative. Otherwise, or in addition, WGS studies including human cases and targeted animal reservoirs could help identify the relative importance of transmission pathways. The outcomes of these analyses could be used to inform targeted controls in the food chain, if transmission is predominantly zoonotic, or a vaccine strategy, if transmission is predominantly anthroponotic. Furthermore, WGS-based structured phylogenetic studies could also be used to infer R_0 in different hosts, enabling detailed follow-up assessment of vaccine efficacy. Given the public health burden of NTS in Africa and elsewhere, accurately determining where the disease is coming from and how it spreads is essential information to decide where and which control strategies might be applied most effectively.

Notes

Financial support. This work was supported by the Biotechnology and Biological Sciences Research Council (grant number BB/M014088/1 to A. E. M.); the Allan Wilson Centre for Molecular Ecology and Evolution (support to N. P. F.); and the University of Auckland (support to T. G. V.). This publication was made possible by a grant from the Bill & Melinda Gates Foundation (grant number OPP1125993).

Supplement sponsorship. This article appeared as part of the supplement "Invasive *Salmonella* Disease in Africa," sponsored by the University of Otago.

Potential conflicts of interest. All authors report no potential conflicts. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Ao TT, Feasey NA, Gordon MA, Keddy KH, Angulo FJ, Crump JA. Global burden of invasive nontyphoidal *Salmonella* disease, 2010. Emerg Infect Dis **2015**; 21:941–9.
2. Majowicz SE, Musto J, Scallan E, et al. Food safety: the global burden of nontyphoidal *Salmonella* gastroenteritis. Clin Infect Dis **2010**; 50:882–9.
3. Pires SM, Evers EG, van Pelt W, et al. Attributing the human disease burden of foodborne infections to specific sources. Foodborne Pathog Dis **2009**; 6:417–24.
4. Pires SM, Vieira AR, Hald T, Cole D. Source attribution of human salmonellosis: an overview of methods and estimates. Foodborne Pathog Dis **2014**; 11:667–76.
5. Kingsley RA, Msefula CL, Thomson NR, et al. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. Genome Res **2009**; 19:2279–87.
6. Barco L, Barrucci F, Olsen JE, Ricci A. *Salmonella* source attribution based on microbial subtyping. Int J Food Microbiol **2013**; 163:193–203.
7. Hald T, Vose D, Wegener HC, Koupeev T. A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. Risk Anal **2004**; 24:255–69.
8. Mughini-Gras L, Barrucci F, Smid JH, et al. Attribution of human *Salmonella* infections to animal and food sources in Italy (2002–2010): adaptations of the Dutch and modified Hald source attribution models. Epidemiol Infect **2014**; 142:1070–82.
9. Mughini-Gras L, van Pelt W. *Salmonella* source attribution based on microbial subtyping: does including data on food consumption matter? Int J Food Microbiol **2014**; 191:109–15.
10. Mullner P, Jones G, Noble A, Spencer SE, Hathaway S, French NP. Source attribution of food-borne zoonoses in New Zealand: a modified Hald model. Risk Anal **2009**; 29:970–84.
11. Mullner P, Spencer SE, Wilson DJ, et al. Assigning the source of human campylobacteriosis in New Zealand: a comparative genetic and epidemiological approach. Infect Genet Evol **2009**; 9:1311–9.

12. Wilson DJ, Gabriel E, Leatherbarrow AJ, et al. Tracing the source of campylobacteriosis. PLoS Genet **2008**; 4:e1000203.
13. Mughini-Gras L, Smid J, Enserink R, et al. Tracing the sources of human salmonellosis: a multi-model comparison of phenotyping and genotyping methods. Infect Genet Evol **2014**; 28:251–60.
14. Mughini-Gras L, Enserink R, Friesema I, Heck M, van Duynhoven Y, van Pelt W. Risk factors for human salmonellosis originating from pigs, cattle, broiler chickens and egg laying hens: a combined case-control and source attribution analysis. PLoS One **2014**; 9:e87933.
15. Dione MM, Ikumapayi UN, Saha D, et al. Clonal differences between non-typhoidal *Salmonella* (NTS) recovered from children and animals living in close contact in The Gambia. PLoS Negl Trop Dis **2011**; 5: e1148.
16. Kariuki S, Revathi G, Gakuya F, Yamo V, Muyodi J, Hart CA. Lack of clonal relationship between non-typhi *Salmonella* strain types from humans and those isolated from animals living in close contact. FEMS Immunol Med Microbiol **2002**; 33:165–71.
17. Kariuki S, Revathi G, Kariuki N, et al. Invasive multidrug-resistant non-typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic transmission? J Med Microbiol **2006**; 55:585–91.
18. Kagambega A, Lienemann T, Aulu L, et al. 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 **2013**; 13:253.
19. Smith AM, Ismail H, Henton MM, Keddy KH; Network G-SS. Similarities between *Salmonella* Enteritidis isolated from humans and captive wild animals in South Africa. J Infect Dev Ctries **2014**; 8:1615–9.
20. Parsons BN, Humphrey S, Salisbury AM, et al. Invasive non-typhoidal *Salmonella* Typhimurium ST313 are not host-restricted and have an invasive phenotype in experimentally infected chickens. PLoS Negl Trop Dis **2013**; 7:e2487.
21. Langridge GC, Fookes M, Connor TR, et al. Patterns of genome evolution that have accompanied host adaptation in *Salmonella*. Proc Natl Acad Sci U S A **2015**; 112:863–8.
22. Hoffmann M, Zhao S, Pettengill J, et al. Comparative genomic analysis and virulence differences in closely related *Salmonella enterica* serotype Heidelberg isolates from humans, retail meats, and animals. Genome Biol Evol **2014**; 6:1046–68.
23. Morelli MJ, Thebaud G, Chadoeuf J, King DP, Haydon DT, Soubeyrand S. A Bayesian inference framework to reconstruct transmission trees using epidemiological and genetic data. PLoS Comput Biol **2012**; 8: e1002768.
24. Royer-Carenzi M, Pontarotti P, Didier G. Choosing the best ancestral character state reconstruction method. Math Biosci **2013**; 242:95–109.
25. Rasmussen DA, Volz EM, Koelle K. Phylodynamic inference for structured epidemiological models. PLoS Comput Biol **2014**; 10:e1003570.
26. Kuhnert D, Stadler T, Vaughan TG, Drummond AJ. Simultaneous reconstruction of evolutionary history and epidemiological dynamics from viral sequences with the birth-death SIR model. J R Soc Interface **2014**; doi:10.1098/rsif.2013.1106.
27. Lemey P, Rambaut A, Drummond AJ, Suchard MA. Bayesian phylogeography finds its roots. PLoS Comput Biol **2009**; 5:e1000520.
28. Mather AE, Reid SWJ, Maskell DJ, et al. Distinguishable epidemics of multidrug-resistant *Salmonella* Typhimurium DT104 in different hosts. Science **2013**; 341:1514–7.
29. Okoro CK, Barquist L, Connor TR, et al. Signatures of adaptation in human invasive *Salmonella* Typhimurium ST313 populations from sub-Saharan Africa. PLoS Negl Trop Dis **2015**; 9:e0003611.
30. Msefula CL, Kingsley RA, Gordon MA, et al. Genotypic homogeneity of multidrug resistant *S*. Typhimurium infecting distinct adult and childhood susceptibility groups in Blantyre, Malawi. PLoS One **2012**; 7:e42085.
31. Bronowski C, Fookes MC, Gilderthorp R, et al. Genomic characterisation of invasive non-typhoidal *Salmonella* enterica subspecies *enterica* serovar Bovismorbificans isolates from Malawi. PLoS Negl Trop Dis **2013**; 7:e2557.
32. Okoro CK, Kingsley RA, Connor TR, et al. Intracontinental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. Nat Genet **2012**; 44:1215–21.
33. Paterson GK, Harrison EM, Murray GG, et al. Capturing the cloud of diversity reveals complexity and heterogeneity of MRSA carriage, infection and transmission. Nat Commun **2015**; 6:6560.
34. Okoro CK, Kingsley RA, Quail MA, et al. High-resolution single nucleotide polymorphism analysis distinguishes recrudescence and reinfection in recurrent invasive nontyphoidal *Salmonella* Typhimurium disease. Clin Infect Dis **2012**; 54:955–63.
35. Orton RJ, Wright CF, Morelli MJ, et al. Observing micro-evolutionary processes of viral populations at multiple scales. Phil Trans R Soc B **2013**; doi:10.1098/rstb.2012.0203.
36. Beerli P. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. Mol Ecol **2004**; 13:827–36.
37. Mollentze N, Nel LH, Townsend S, et al. A Bayesian approach for inferring the dynamics of partially observed endemic infectious diseases from space-time-genetic data. Proc R Soc B **2014**; doi:10.1098/rspb.2013.3251.
38. Benavides JA, Cross PC, Luikart G, Creel S. Limitations to estimating bacterial cross-species transmission using genetic and genomic markers: inferences from simulation modeling. Evol Appl **2014**; 7:774–87.
39. Sheppard SK, Didelot X, Meric G, et al. Genome-wide association study identifies vitamin B5 biosynthesis as a host specificity factor in Campylobacter. Proc Natl Acad Sci U S A **2013**; 110:11923–7.