Case-control investigation of invasive Salmonella disease in Africa – comparison of human, animal and household environmental isolates find no evidence of environmental or animal reservoirs of invasive strains

Leonard Koolman1,2#, Reenesh Prakash1,2#, Yohane Diness1, Chisomo Msefula4, Tonney S. Nyirenda4, Franziska Olgemoeller1,3, Blanca Perez-Sepulveda2, Jay C. D. Hinton2, Siân V. Owen2, Nicholas A. Feasey1,3, Philip M. Ashton1,2*, Melita A. Gordon1,2,4

1 Malawi-Liverpool Wellcome Programme, Blantyre, Malawi
2 Institute of Infection, Veterinary & Ecological Sciences, University of Liverpool, United Kingdom
3 Department of Clinical Sciences, Liverpool School of Tropical Medicine, United Kingdom
4 Kamuzu University of Health Sciences, Blantyre, Malawi

# contributed equally

*Corresponding author

Email: philip.ashton@liverpool.ac.uk

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

Background

Invasive *Salmonella* infections are a major cause of morbidity and mortality in Sub-Saharan Africa (SSA), but the sources and transmission routes are uncertain. We investigated potential sources for cases of invasive disease by sampling healthy people, animals, and the environment in index-case and geographically-matched control households.

Methods

Sixty index cases of human invasive *Salmonella* infection were recruited (28 invasive Non-Typhoidal *Salmonella* (iNTS) disease, and 32 typhoid). Each index-case household was geographically matched to a control household. Extensive microbiological sampling included stool sampling from healthy household members, stool or rectal swabs from household-associated animals and boot-sock sampling of the household environment.

Results

1203 samples were taken from 120 households, yielding 43 *Salmonella* isolates from 25 households (overall sample positivity 3.6%). Isolates from households were all NTS and spanned 15 Sequence Types (STs). iNTS disease was caused by 3 STs of *Salmonella Typhimurium*, mainly ST313. Two *S. Typhimurium* isolates from index cases closely matched isolates from their respective asymptomatic household members (2 and 3 SNPs different respectively). There was no overlap of STs causing iNTS disease with environmental or animal isolates, despite recovery of diverse NTS.

Conclusions

The finding of NTS strains from index cases matching household members, coupled with lack of overlap with either animal or environmental isolates, supports a hypothesis that healthy humans are the source of iNTS infections in the household. The breadth of NTS strains found in the household environment across all sites demonstrated the robustness of
sampling and methodology to detect NTS, and suggests a diverse ecology of *Salmonella* in this setting. The lack of *S. Typhi* isolated from the household environment may suggest a need for further methodological development to culture sources of typhoid.
Invasive *Salmonella* infections lead to millions of disability adjusted life years (DALYs) lost every year globally. There are two main types of invasive *Salmonella* infections in Africa; i) typhoid fever, caused by *Salmonella Typhi*, and ii) invasive Non-Typhoidal Salmonella (iNTS) disease, primarily caused in our setting by *Salmonella Typhimurium*. Despite their high disease burden, and despite observed differences in the epidemiology of typhoid compared to iNTS disease, we do not have a good understanding of their reservoirs and transmission. Therefore, we carried out extensive microbiological sampling of the household members and living environments of patients with invasive *Salmonella* infections, and of geographically-matched control households, and investigated the genetic relationships between household *Salmonella* and index-case blood-stream isolates using whole genome sequencing (WGS). The only sample type where we identified *Salmonella* that matched strains causing iNTS disease was the stool of healthy people from the case household, suggesting either person-to-person spread or infection from a common source. Boot-sock sampling of the household environment resulted in the highest yield of *Salmonella* of any of our sampling strategies. None of the other 41 environmental *Salmonella* isolates from non-human sources, including 4 domestic animal-associated isolates, matched the disease-causing types. Our findings are consistent with the reservoir of Typhimurium iNTS infections being the human gastrointestinal tract and transmission occurring within households.
Introduction

Salmonella infections are a substantial cause of morbidity and mortality in many low- and middle-income countries and are an important cause of enteric infections leading to invasive bloodstream infections (BSI) [1,2]. Both typhoidal and Non-Typhoidal Salmonella (NTS) contribute extensively to the burden of BSI. Globally in 2017, invasive NTS (iNTS) disease was estimated to cause 535,000 illnesses and 77,500 deaths [2]. Sub-Saharan Africa accounts for 85.8% of global iNTS deaths [2]. There were an estimated 14.3 million cases of typhoid and paratyphoid fever in 2017, resulting in 135,900 deaths, 15.8% of which were in Sub-Saharan Africa [1].

Salmonella Typhi has emerged as a major cause of BSI in southern and eastern Africa since 2012 [3–7], and Malawi has a very high incidence of 444/100000 person-years of observation [8]. The pathways by which people are exposed to S. Typhi are uncertain [9]. Improved understanding of reservoirs and transmission routes would help with the design of effective interventions, but delineating the relative importance of long (i.e. waterborne or environmentally mediated) and short cycle (i.e. within household) transmission routes has remained difficult for a variety of faecal oral pathogens [9,10].

In High Income Countries (HICs), NTS are predominantly associated with self-limiting enterocolitis. In Sub-Saharan Africa (SSA), however, the epidemiological picture is dominated by iNTS disease, diagnosed by positive blood-culture, which carries a high case-fatality [11–15]. iNTS in Africa is associated with host susceptibilities, including HIV in adults and malaria infections, malnutrition or HIV in young children [13,15–17]. Two Salmonella serotypes, Typhimurium and Enteritidis, have been reported to be the commonest causes of iNTS in Malawi, South Africa, Kenya, Mozambique and Mali, with S. Typhimurium Sequence Type

...
ST) 313 (ST313) and S. Enteritidis ST11 being the most frequently identified [17–24]. In the city of Blantyre, Malawi, there have been two epidemics of iNTS disease, the first caused by S. Enteritidis between 1999-2002, and the second caused by S. Typhimurium between 2002-2008, with S. Typhimurium continuing to cause a significant burden of disease [17]. Both epidemics were associated with multi-drug resistance [17].

While NTS infection in HICs is generally associated with zoonotic transmission, the reservoirs and transmission route of the pathogens responsible for iNTS in SSA remains unknown. Seasonal patterns of iNTS, and the close domestic association of humans and animals in parts of SSA suggest a possible role for environmental factors [25,26]. Serotypes associated with iNTS have been isolated from animals in South Africa, although no genetic characterisation was carried out on these samples [27,28]. In contrast, microbiological investigations in Kenya, based either on serotype and Pulsed Field Gel Electrophoresis or more recently WGS, found isolates in the stool of household members which were closely related to index cases, and only rarely found linked serotypes in environmental samples such as water from nearby rivers, household and market vendor food, or household animals [29,30]. The Global Enteric Multicentre Study found iNTS associated sub-types of Salmonella in the stool of 42 children with diarrhoea and 17 healthy children, highlighting the potential role of the human gut as a source of iNTS infections [31]. In The Gambia, there was no overlap observed between STs observed in iNTS cases and those observed in rectal swabs from domestic animals from the households of cases [32], while in Tanzania and Kenya, extensive sampling of “meat pathway” isolates and human disease isolates only isolated ST313 from human samples [33]. ST313 was isolated from pigs in Nairobi, but was shown to group with ST313 from the UK, which are not associated with invasive disease [34,35]. Genomic evidence suggests that S. Typhimurium ST313, one of the commonest causes of iNTS, is under-going genome degradation [22,23], a
phenomenon associated with narrower host range (i.e. becoming human restricted) and increased invasiveness in a range of bacteria, including S. Typhi and S. Paratyphi A [36,37]. Transcriptomic and phenotypic evidence also points towards specific niche adaptations to human hosts among African isolates of NTS causing invasive disease [38,39].

To address the knowledge gap around sources for strains of Salmonella causing invasive Salmonella infections, we conducted extensive human, animal, and household environmental sampling of iNTS and typhoid index-case households and geographically matched control households in Blantyre, Malawi. This study is the first to combine extensive bootsock environmental sampling alongside sampling of humans and animals, and to use WGS of the resulting isolates, aiming primarily to investigate overlap of invasive index strains with household sources, and also to comprehensively determine broader household patterns of bacterial relatedness in relation to invasive Salmonella diseases.
Methods

Setting

Queen Elizabeth Central Hospital (QECH) provides free healthcare to the approximately 1.3 million inhabitants of Blantyre District, Malawi, and is the tertiary referral hospital for southern Malawi. Since 1998, the Malawi-Liverpool-Wellcome Trust Clinical Research Programme has provided a quality controlled diagnostic blood culture service for febrile adult and paediatric medical patients admitted to QECH since 1998. This service is provided for admitted adults (>16 years old) with axillary temperature over 37.5°C or clinical suspicion of sepsis, and for children (<16 years old) who were malaria slide negative, or positive and critically ill, or with clinical suspicion of sepsis.

Case and control recruitment

Sequential blood culture confirmed adult or paediatric cases with invasive Salmonella infections (either iNTS or typhoid fever) presenting at QECH between March 2015 and October 2016 were approached as index cases. We sought written informed consent from adults, and guardian consent for minors. Patients living outside the Blantyre district and those with recurring iNTS disease were excluded. Index case households were sampled within a maximum of 14 days following initial presentation to QECH. Following recruitment, the field team visited the index cases in their households, where GPS co-ordinates were taken and a household socio-demographic questionnaire was completed. Control households were then selected by random bottle-spin and pacing 100 m from the index case household, recruited with informed consent, and GPS and questionnaire data were also collected. Exclusion criteria for control households were current treatment for invasive Salmonella disease for any family member, or declining consent. In this event, the next-nearest house in the same direction was selected.
We carried out a microbiological survey of the index-case household and the control household, comprising stool from household members, stool or rectal swabs from domestic and household-associated animals and systematic boot-sock sampling of the living environment (latrine, rubbish area, bedroom, cooking area and the house perimeter).

**Stool (humans):** Samples (5 to 10 g) were collected from all available family members in the index case and control households. Sample containers were provided 24-48 hours before the main sample-collection visit, and household members were requested to collect up to three consecutive stools from each individual. The three samples belonging to each person were pooled, to maximise recovery of *Salmonella*.

**Stool (animals):** Fresh stool or rectal swab samples were collected from domestic or household-associated animals (cows, chickens, goats, pigs, dogs, cats, rats, doves, guinea pigs or gecko lizards).

**Environmental:** The investigator’s shoe was first covered with a waterproof sterile bag to prevent cross-contamination, then a sterile bootsock was put over this, moistened with sterile water, and single bootsocks were used to collect samples from each of the follow sites: pit-latrine, outdoor rubbish area, cooking and bedroom areas, and from the outside perimeter of the household, using a standardised protocol for location and number of paces (10 paces), and a step and twist motion. Bootsocks were then removed and placed into a sterile plastic bag for transport.

All samples were transferred from the field to the laboratory in coolboxes, and processed on the same day.
Sample Enrichment and culture

Supplementary Figure 1 shows a schematic of sample processing. A portion (~1g) of stool was emulsified in 9ml of Selenite F selective enrichment broth (Oxoid, UK) and incubated at 37°C for 18-24 hours. For samples other than stool (i.e. bootsocks, swabs), we used a pre-enrichment step; these samples were incubated in Buffered Peptone Water (BPW) at 37°C for 18-24 hours. One ml of the BPW sample was then transferred into 9 ml Selenite broth and incubated for a further 18-24 hours at 37°C.

After incubation, 100µl of bacterial culture taken from the surface of the Selenite broth was sub-cultured onto Xylose Lysine Deoxycholate (XLD) agar, incubated at 37°C for 18-24 hours and examined for suspected colonies of Salmonella (red/pink colonies with black centres). Pure candidate colonies were biochemically confirmed as Salmonella species using API 20E (BioMerieux). Serotyping was performed by the slide agglutination technique. Putative Salmonella isolates were sub-cultured from beads onto Nutrient agar (Oxoid, UK) and incubated for 24 hours at 37°C. Following growth, a drop of 0.85% saline was placed onto a microscope slide and a single colony from the Nutrient agar was emulsified into the saline. Positive serotyping was confirmed if agglutination occurred in antigens O4, O9, Vi, Hd, Hg, Hi, Hm. If an organism was positive for Vi antigen but negative for O9 antigen, a dense suspension of the organism was prepared in 0.85% saline, autoclaved at 121°C for 15 minutes and the agglutination repeated. S. Typhi was differentiated from NTS isolates if agglutination occurred for O9, Vi and Hd antigens. All API 20E confirmed Salmonella isolates which were not S. Typhi were sent to the University of Liverpool, UK, for WGS.

Whole Genome Sequencing (WGS)

Bacterial genomic DNA was extracted from overnight cultures in LB (1% tryptone, 0.5% yeast extract, 0.5% NaCl; pH 7.0) using the DNeasy Blood and Tissue kit (Qiagen), as per the Gram-
negative bacteria protocol for the Quick-DNA™ Universal Kit (Zymo Research), Biological Fluids & Cells protocol. Sequencing libraries were constructed using the TruSeq DNA PCR-free library preparation kit (Illumina) or the TruSeq Nano DNA HT library preparation kit (Illumina) using a target insert size of 550bp. Libraries were sequenced on an Illumina MiSeq instrument using a 2x250bp paired end protocol at the Centre for Genomic Research, University of Liverpool, UK. Raw sequencing reads were trimmed to remove Illumina adapter sequences using Cutadapt [40] and further trimmed to remove poor quality sequence using Sickle [41] version 1.2 with a minimum window quality score of 20, discarding reads which were less than 10 bp after trimming. Trimmed reads were submitted to Enterobase (https://enterobase.warwick.ac.uk/) for draft genome assembly, MLST assignment, *in silico* serotype prediction and phylogenetic analysis [42]. SNP identification within Enterobase was carried out via a mapping and SNP calling pipeline, and phylogenetic tree construction done with RAxML [43–46] Determinants of antimicrobial resistance genes were identified using the amr-finder-plus software version 3.9.8, using database version 2020-12-17.1 [47].

Data availability

All sequencing data is available from ENA in the BioProject X. Per sample BioSample accessions are listed in Supplementary Table 1.
Results

Sixty-seven index-case households (35 typhoid and 32 iNTS) were eligible for enrolment between March 2015 and October 2016. Seven index-case households declined to participate, leaving 32 typhoid index case households, with 32 controls, and 28 iNTS index case households, with 28 controls recruited for the study, giving a total of 120 households. All households were in high-density low-income housing locations within Blantyre.

Isolation of Salmonella by culture: Table 1 shows sampling and positivity rates for Salmonella by sample-type and by household. We collected a total of 1203 samples from the 120 index-case and control households (mean 10 samples per household), yielding 43 isolates (all NTS) which were confirmed as Salmonella by WGS, giving an overall sample positivity of 3.6% (Table 1). Overall, sample positivity was very similar across human, animal and bootsock environment samples (human stool 16/491, 3.3%; animals 4/110, 3.6%; bootsocks 23/620, 3.8%), and across different household categories (Table 1). Animal sampling was from 110 animals in 71/120 (59%) households, and included both domesticated and non-domesticated species (cows, chickens, pigs, dogs, cats, rats, doves, guinea pigs and geckos).

Salmonella was found in 25/120 households (overall household positivity of 21%). There was a significantly higher household positivity in typhoid households (28%) and their geographically-matched neighbours, compared to NTS cases and control households (13%) (Chi-squared P-value = 0.036), particularly in the environmental bootsock category of samples.

No S. Typhi were isolated from any category of household samples, including from people living in the same house as index cases.
Table 1: Isolation of Salmonella from different sample types in iNTS and typhoid case and paired control households

<table>
<thead>
<tr>
<th>Sample Type</th>
<th>Total samples</th>
<th>No. samples with Salmonella</th>
<th>% positive samples</th>
<th>No. HH tested</th>
<th>No. HH with Salmonella</th>
<th>% Positive HH</th>
<th>Total samples</th>
<th>No. samples with Salmonella</th>
<th>% positive samples</th>
<th>No. HH tested</th>
<th>No. HH with Salmonella</th>
<th>% Positive HH</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>iNTS Index Case Households (HH) (n=28)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human Stool</td>
<td>116</td>
<td>9</td>
<td>7.8</td>
<td>28</td>
<td>4</td>
<td>14.3</td>
<td>97</td>
<td>0</td>
<td>0.0</td>
<td>28</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Animal stool/swab</td>
<td>27</td>
<td>0</td>
<td>0.0</td>
<td>20</td>
<td>0</td>
<td>0.0</td>
<td>38</td>
<td>0</td>
<td>0.0</td>
<td>17</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>HH environment</td>
<td>137</td>
<td>0</td>
<td>0.0</td>
<td>28</td>
<td>0</td>
<td>0.0</td>
<td>131</td>
<td>0</td>
<td>0.0</td>
<td>3</td>
<td>10.7</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>280</strong></td>
<td><strong>9</strong></td>
<td><strong>3.2</strong></td>
<td><strong>28.0</strong></td>
<td><strong>4</strong></td>
<td><strong>14.3</strong></td>
<td><strong>266</strong></td>
<td><strong>6</strong></td>
<td><strong>2.3</strong></td>
<td><strong>28</strong></td>
<td><strong>3</strong></td>
<td><strong>10.7</strong></td>
</tr>
<tr>
<td><strong>Typhoid Index Case househods (HH) (n=32)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human Stool</td>
<td>157</td>
<td>4</td>
<td>2.5</td>
<td>32</td>
<td>2</td>
<td>6.3</td>
<td>121</td>
<td>3</td>
<td>2.5</td>
<td>32</td>
<td>3</td>
<td>9.4</td>
</tr>
<tr>
<td>Animal stool/swab</td>
<td>22</td>
<td>2</td>
<td>9.1</td>
<td>18</td>
<td>2</td>
<td>11.1</td>
<td>23</td>
<td>2</td>
<td>8.7</td>
<td>16</td>
<td>2</td>
<td>12.5</td>
</tr>
<tr>
<td>HH environment</td>
<td>168</td>
<td>5</td>
<td>3.0</td>
<td>32</td>
<td>4</td>
<td>12.5</td>
<td>166</td>
<td>12</td>
<td>7.2</td>
<td>32</td>
<td>9</td>
<td>28.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>347</strong></td>
<td><strong>11</strong></td>
<td><strong>3.2</strong></td>
<td><strong>32</strong></td>
<td><strong>7</strong></td>
<td><strong>21.9</strong></td>
<td><strong>310</strong></td>
<td><strong>17</strong></td>
<td><strong>5.5</strong></td>
<td><strong>32</strong></td>
<td><strong>11</strong></td>
<td><strong>34.4</strong></td>
</tr>
</tbody>
</table>
Relatedness of Sequence Types from different sample categories: Figure 1 demonstrates the wide diversity and relationships of Sequence Types (STs) isolated from different sample categories (invasive disease, healthy human stool, animals, and the household environment). Table 2 shows the individual household origins of each isolate, and the relationship between ST and imputed serovar for each isolate. Among invasive NTS disease cases, 8 of 27 index case NTS isolates could not be resuscitated from the archive for sequencing. The remaining 19 typed index NTS isolates were all S. Typhimurium from just 3 STs; 13 were ST313, 5 were ST3257 and 1 was ST19. These latter STs were each only one locus variant away from ST313. One representative typhoid index strain (ST2) is also included for reference. Among healthy household members there were 16 isolates from 8 STs; among animals there were 4 isolates from 3 STs; and from environmental samples there were 23 isolates from 9 different STs.

Importantly, in two instances, there was ST overlap between invasive index-case isolates and healthy human household samples; these were 2 cases of invasive disease caused by S. Typhimurium ST313 and ST3257, that were linked to isolates from 2 healthy household members in their respective households, who also had also ST313 and ST3257 isolated from their stools (Figure 1). Interestingly, although stool isolates from household members were found equally among adults (n=7) and children (n=8), in both cases the household member carrying a matching isolate was an adult. The genetic relatedness between index and human household strains was explored at higher resolution by WGS phylogenetic analysis (Figure 2 and below).

Also importantly, there was no overlap in STs found between isolates causing invasive human disease and any animal or environmental isolates, despite a good rate of isolation from
all 3 sample-types. There was, however, non-co-localised overlap in some STs between healthy humans and household-associated animals. A monophasic ST3262 was found in a cat, a gecko, and a human, but these were all in different households. Similarly, ST3261 was found in 2 healthy humans and a chicken, but these 3 events were also not from the same households (Table 1). Significantly, ST3261 nor ST3262 are not known to have caused human invasive disease in Blantyre. ST325 was found only in an animal (gecko).

As might be expected in this low-income setting, where household sanitation is limited and faecal contamination of the household may occur, there were STs isolated from the environment that overlapped with those from healthy humans, and household animals. Of 9 STs found in the environment, 3 (ST3262, ST3263, ST2152) showed some overlap with healthy human or animal isolates. However, a further 6 STs were found only in the household environment, revealing the diversity of Salmonella strains in this setting.
Figure 1: Minimum spanning tree of all isolates. Each circle represents a sequence type, the size of the circle is proportional to the number of isolates of that ST. Red text indicates the number of locus variants between two STs. Two invasive disease index isolates had a matched isolate from a healthy human sample of the same ST from the same household. These isolates are indicated by black and red outlines.
Table 2: Index case and household isolates described in this study

<table>
<thead>
<tr>
<th>Household (HH) ID</th>
<th>Index case NTS or typhoid?</th>
<th>Household case or control?</th>
<th>Household sample type</th>
<th>Sequence Type (ST)</th>
<th>Imputed serovar</th>
<th>Match to Index case blood culture?</th>
<th>Househoold (HH) ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Animal stool (chicken)</td>
<td>3261</td>
<td>Agoueve</td>
<td>Cubana</td>
<td>1C</td>
</tr>
<tr>
<td>2S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>3257</td>
<td>Typhimurium</td>
<td></td>
<td>2S</td>
</tr>
<tr>
<td>3C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Family member stool</td>
<td>3261</td>
<td>Agoueve</td>
<td>Cubana</td>
<td>3C</td>
</tr>
<tr>
<td>4S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td>4S</td>
</tr>
<tr>
<td>6S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Family member stool</td>
<td>3261</td>
<td>Agoueve</td>
<td>Cubana</td>
<td>6S</td>
</tr>
<tr>
<td>6C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock perimeter</td>
<td>3263</td>
<td>Havana</td>
<td>II</td>
<td>6C</td>
</tr>
<tr>
<td>6C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Family member stool</td>
<td>3263</td>
<td>Havana</td>
<td>II</td>
<td>6C</td>
</tr>
<tr>
<td>8C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock bedroom</td>
<td>3263</td>
<td>Havana</td>
<td>II</td>
<td>8C</td>
</tr>
<tr>
<td>9S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td>9S</td>
</tr>
<tr>
<td>9S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Family member stool</td>
<td>3263</td>
<td>Havana</td>
<td>II</td>
<td>9S</td>
</tr>
<tr>
<td>9S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Family member stool</td>
<td>3263</td>
<td>Havana</td>
<td>II</td>
<td>9S</td>
</tr>
<tr>
<td>9S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Family member stool</td>
<td>3263</td>
<td>Havana</td>
<td>II</td>
<td>9S</td>
</tr>
<tr>
<td>9S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Family member stool</td>
<td>3263</td>
<td>Havana</td>
<td>II</td>
<td>9S</td>
</tr>
<tr>
<td>12S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>3257</td>
<td>Typhimurium</td>
<td></td>
<td>12S</td>
</tr>
<tr>
<td>14S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td>14S</td>
</tr>
<tr>
<td>15S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Animal stool (gecko)</td>
<td>3262</td>
<td>II 42:r:-</td>
<td>IIIb</td>
<td>15S</td>
</tr>
<tr>
<td>15S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Boot sock perimeter</td>
<td>3262</td>
<td>II 42:r:-</td>
<td>IIIb</td>
<td>15S</td>
</tr>
<tr>
<td>16S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>19</td>
<td>Typhimurium</td>
<td></td>
<td>16S</td>
</tr>
<tr>
<td>16S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Family member stool</td>
<td>3262</td>
<td>II 42:r:-</td>
<td>IIIb</td>
<td>16S</td>
</tr>
<tr>
<td>16C</td>
<td>iNTS</td>
<td>Control</td>
<td>Boot sock latrine</td>
<td>2347</td>
<td>Mgulani</td>
<td></td>
<td>16C</td>
</tr>
<tr>
<td>16C</td>
<td>iNTS</td>
<td>Control</td>
<td>Boot sock cooking area</td>
<td>2347</td>
<td>Mgulani</td>
<td></td>
<td>16C</td>
</tr>
<tr>
<td>17S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Boot sock latrine</td>
<td>3262</td>
<td>II 42:r:-</td>
<td>IIIb</td>
<td>17S</td>
</tr>
<tr>
<td>18S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>3257</td>
<td>Typhimurium</td>
<td></td>
<td>18S</td>
</tr>
<tr>
<td>22C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock latrine</td>
<td>3609</td>
<td>Djama</td>
<td>II 42:229:-</td>
<td>22C</td>
</tr>
<tr>
<td>23S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Animal stool (gecko)</td>
<td>316</td>
<td>Montevideo</td>
<td></td>
<td>23S</td>
</tr>
<tr>
<td>25S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Family member stool</td>
<td>313</td>
<td>Typhimurium</td>
<td>match1</td>
<td>25S</td>
</tr>
<tr>
<td>25S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td>match1</td>
<td>25S</td>
</tr>
<tr>
<td>Group</td>
<td>Type</td>
<td>Sample</td>
<td>Description</td>
<td>ID</td>
<td>Serology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>--------</td>
<td>-------------</td>
<td>----</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25C</td>
<td>iNTS</td>
<td>Control</td>
<td>Boot sock latrine</td>
<td>473</td>
<td>Hadar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25C</td>
<td>iNTS</td>
<td>Control</td>
<td>Boot sock bedroom</td>
<td>473</td>
<td>Hadar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25C</td>
<td>iNTS</td>
<td>Control</td>
<td>Boot sock cooking area</td>
<td>473</td>
<td>Hadar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27C</td>
<td>iNTS</td>
<td>Control</td>
<td>Boot sock bedroom</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>28C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Animal rectal swab</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>28C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock latrine 1</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>28C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock latrine 1</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>29C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock latrine</td>
<td>2152</td>
<td>Gaminara</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Family member stool</td>
<td>3608</td>
<td>Ogbete</td>
<td>II</td>
<td></td>
</tr>
<tr>
<td>31S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Index case blood</td>
<td>2</td>
<td>Typhi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>3257</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock latrine</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>42S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>44S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Family member stool</td>
<td>3257</td>
<td>Typhimurium</td>
<td>match2</td>
<td></td>
</tr>
<tr>
<td>44S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>3257</td>
<td>Typhimurium</td>
<td>match2</td>
<td></td>
</tr>
<tr>
<td>46S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>48S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Boot sock bedroom</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>50S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Boot sock bedroom</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>50S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Boot sock cooking area</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>50C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock perimeter</td>
<td>3611</td>
<td>II 40:b:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>53C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock bedroom</td>
<td>3262</td>
<td>II 42:r:</td>
<td>IIIb</td>
<td></td>
</tr>
<tr>
<td>56C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock bedroom</td>
<td>293</td>
<td>Amager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock perimeter</td>
<td>293</td>
<td>Amager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56C</td>
<td>Typhoid</td>
<td>Control</td>
<td>Boot sock cooking area</td>
<td>3610</td>
<td>Aberdeen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>57S</td>
<td>iNTS</td>
<td>Case household</td>
<td>Index case blood</td>
<td>313</td>
<td>Typhimurium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>59S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Family member</td>
<td>14</td>
<td>Senftenberg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>59S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Family member stool</td>
<td>2152</td>
<td>Gaminara</td>
<td></td>
<td></td>
</tr>
<tr>
<td>59S</td>
<td>Typhoid</td>
<td>Case household</td>
<td>Family member stool</td>
<td>2152</td>
<td>Gaminara</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sequence types and WGS-inferred serovars from different sample categories:

The relationship between STs and WGS-inferred serovars for all isolates in the study is documented in Table 2 (genome database accession numbers are available in Supplementary Table 1). Of 19 index iNTS case isolates from blood culture, 19/19 were S. Typhimurium, of 3 different STs; ST313 (n=13), ST3257 (a Single Locus Variant of ST313, n=5) and ST19 (n = 1) (Supplementary Table 1, Figure 1).

Serovars that were associated with healthy humans were Havana (7), Typhimurium (2), Gaminara (2), Agoueve/Cubana (2), Seftenberg (1) and Ogbete (1). Serovars associated with animals were II42:r:-|IIIb 42:r:-:z50 (1 gecko, 1 cat), Montevideo (1 gecko) and Agoueve/Cubana (1 chicken). The wide range of serovars found in the household environment included II42:r:-|IIIb 42:r:-:z50 (10), Hadar (3), Havana (3), Amager (2), Mгулani (2) Djамa (1) and Aberdeen (1).

Reflecting the broad diversity of isolates, many STs/serovars were represented by just a single isolate. In addition, where there were multiple isolations of the same serovar, a large proportion were localised from within the same household. Striking examples of this are: of 9 isolates of S. Havana, 6 arose from one household (9S), from 6 out of 7 children sampled; of 3 isolates of S Hadar, all arose from the same household (25C, latrine, bedroom and cooking area); of 12 isolates of II42:r:-|IIIb 42:r:-:z50 there were multiple positive samples in three households, including from animal and environmental samples from the same households. Similarly, 2 isolates each of S. Mгулani and S. Amager isolates were linked to the same respective households (16C and 56C). This indicates a very marked geographical diversity between households, and a degree of homogeneity within households among NTS that are not linked to invasive disease.
**Whole genome phylogeny of index/household pairs**: We employed whole genome-derived phylogeny to infer at high resolution the relationship between the human invasive disease cases of *S. Typhimurium* ST313 & ST3257 and the stool carriage isolates of their respective index household members. In both cases, the isolate that caused the invasive disease formed a monophyletic group with the carriage isolate from the same index household (Figure 2). The SNP distance was 2 SNPs in one case, and 3 SNPs in the other (Figure 2).

**Figure 2**: Whole genome maximum likelihood phylogenetic tree of *S. Typhimurium* which shows the close relationship between the index case and family member isolates from households 25 and 44. Sequencing read sets were compared against the reference genome D23580.
Antimicrobial resistance genes: We had WGS data for 63 isolates from index cases of human invasive disease and household sampling. Antimicrobial resistance gene results are shown in Supplementary Table 1. There were no antimicrobial resistance genes associated with 52 of them, including 68% of index blood-stream *S. Typhimurium* isolates (9 ST313s, 3 ST3257 and 1 ST19), and 89% of the animal and bootsock isolates. Two bootsock (both *S. Mbulani*) and one animal stool (*S. Montevideo*) isolate(s) carried *fosA7* genes associated with resistance to fosfomycin. Six index-case blood-stream *Typhimurium* isolates and two *Typhimurium* from the stool of household members encoded all or some of the *blaTEM-1, catA1, sul1, sul2 and dfrA1* resistance genes which are typically associated with invasive *S. Typhimurium* in Malawi.
Discussion

Our findings successfully address our primary question, and support the hypothesis that humans may be the primary source of the organisms responsible for iNTS in sub-Saharan Africa. All index cases were caused by 3 STs of S. Typhimurium, and matching S. Typhimurium STs were isolated from the stool of their respective household members. In both cases, the isolates from household members were of the same ST and within 2-3 SNPs distance of the isolate from the index case. On the other hand, despite successfully culturing 27 Salmonella isolates from the environment and animals, representing 13 different STs/serovars, none were iNTS-associated Salmonella serovars.

Our results and interpretation are consistent with and extend those of several previous reports. Firstly, in Kenya, iNTS infections caused by Salmonella Typhimurium or Enteritidis in an informal urban settlement were not found to be associated with the rearing of any domestic animals [48]. Secondly, in Burkina Faso the only isolates that matched the index cases were obtained from the stool of household members, despite extensive animal and water sampling in the households of iNTS patients [49]. Finally, analysis of cases of moderate to severe diarrhoea and healthy controls in Africa and Asia found ST313 in both sick and healthy children, supporting the finding that Salmonella are found in the stool of healthy children [31].

Our study also extends the understanding of the sources of iNTS infection in several additional ways. Most importantly, we sampled from a broad range of locations. Previous case-control studies have sampled adults and children in the same household, and household-associated animals, but this is the first study to also sample the household environment. This yielded a similar culture positivity-rate to other sample-types, but also yielded a much wider...
range of STs/serovars, indicating that *Salmonella* has a more diverse ecology in the household environment than has previously been appreciated. The previously unappreciated broad diversity of *Salmonella* in the household environment in this setting is in contrast to the narrow range of STs causing human invasive disease, and the lack of overlap between these categories is very striking.

Ours is also the most comprehensive study of its type, finding the greatest number of non-invasive disease isolates of any publication addressing this topic to date. We isolated a total of 27 animal and environmental *Salmonella*, almost double the number isolated from a comparable study [49]. Animal-ownership is not high in overcrowded urban areas, but we nevertheless sampled 110 household-associated animals from 59% of households, yielding 3 STs from 4 animals of 3 species (1 chicken, 1 cat and 2 geckos). Bootsock sampling provided 53% of the non-invasive disease isolates in our study. This excellent yield from bootsocks for detecting *Salmonella* in other contexts has been reported in other contexts [50], and we recommend that similar studies carried out in the future also use this method.

Our methodology was thus able to resolve microbiological connections between people and their household environments at both the individual household level and the community level. The presence of some STs in multiple sample types was reassuring, indicating that we were able to resolve a degree of connectivity between the different sample source types in this setting, which might be expected as a result of faecal contamination of the household in a setting with poor sanitation infrastructure. However, the dominant picture was a diversity of non-invasive disease *Salmonella*, and unique distribution of different STs by sample-type. 6/9 STs found in the environment were found in neither animals nor healthy humans. Three STs were only found in one sample type, and even when there were multiple isolates of a single
ST, these were often concentrated in the same household or a limited number of households. This suggests that some serovars may occupy a specific niche.

*S. Typhi* is human-restricted in its ability to cause disease, but is thought to inhabit the environment during long-cycle transmission in the community. While we failed to isolate *S. Typhi* from the environment of either case or control households, this is known to be technically challenging, and our results highlight the need for further methodological development in this area [51].

There was no definite broader discernible pattern of STs/serovars observed between different categories of households. There was, however, higher household-positivity rates in typhoid-linked case and control household pairs, compared to iNTS-linked case and control household pairs. Case and matched control houses were 100m apart, and this was chosen as a distance that meant households were geographically close, but would not be likely to share latrines or experience direct surface-water cross-contamination, since at that distance there would usually be several intervening plots in this setting. Poor sanitation is known to be a risk factor for typhoid because of the long-cycle transmission cycle, while the main epidemiological risks described for iNTS disease are found in the host; malaria, malnutrition, anaemia and HIV. It is possible that a higher rate of household contamination with NTS serovars in case/control household pairs in proximity to typhoid cases reflects worse general water and sanitation, and possibly higher poverty in that neighbourhood, compared to iNTS case/control household pairs, but this would require further investigation.

While antimicrobial resistance is a major concern in our setting, only 12% of our household isolates had one or more genes known to be associated with antimicrobial resistance, in
contrast to 39% of our human invasive disease isolates. The fact that 62% of the ST313
isolated here carried no antimicrobial resistance genes suggests that they could belong to the
recently described ST313 Lineage 3 [52]. We did not carry out AMR phenotyping for the
isolates obtained in this study, but AMR profile inference from genome sequence is highly
accurate for Salmonella [53].

A major limitation to our study is that both the index case and the ST313 colonised household
member could have been infected from an unsampled environmental reservoir, either outside
the household or within the household. We do not report on food or water sampling. While
we had hoped to sample food, we found that in practice most households had no refrigeration,
and there were very rarely “leftovers” as meals were eaten at one sitting, meaning this had
limited feasibility. An easier approach in this setting might be to sample from local markets
and food-supply networks. Even in investigations of food-borne outbreaks in HICs, food
testing results are difficult to interpret and have inherent limitations [54]. Epidemiological
analysis identified no vehicle of infection in 38% of putative foodborne outbreaks in the UK,
and microbiological confirmation would only occur in a minority of those 38% [55]. In
addition to a contaminated foodstuff, water sources, both drinking and washing/cleaning,
could be involved in the transmission of organisms causing iNTS, as is seen for S. Typhi
[9,56], and further methodological development would be required for these studies. If the
index-case were infected outside the home e.g. at work, school, or nursery, which is a
possible risk factor in our setting [9], then our study would not identify that microbiological
source.
Conclusions

In summary, we conducted comprehensive environmental and household contact sampling of the homes of patients with invasive Salmonella disease, and geographically matched control households. The only household isolates of serotypes associated with iNTS came from the stool of household members of iNTS index cases. This study contributes to accumulating evidence that the source of iNTS infections is the human gastrointestinal tract. Significantly, we found no strains causing human iNTS disease among any household-associated animals, despite extensive animal sampling and a comparable NTS positivity rate. We also uncovered a much wider range of NTS isolates that are not associated with invasive disease, across all sampling sites, than has previously been appreciated, indicating a diverse ecology for Salmonella in the household environment.

Acknowledgements

We are grateful to Alistair Darby, John Kenny and staff at the Centre for Genomic Research at the University of Liverpool for assistance with genome sequencing. Jay Hinton is grateful for funding by a Wellcome Trust Senior Investigator Award (106914/Z/15/Z). The Malawi-Liverpool-Wellcome Programme is core-funded by a grant from Wellcome (206545/Z/17/Z). Melita Gordon is supported by a Research Professorship from the UK Department of Health National Institute of Health Research (NIHR300039)


doi:10.1101/gr.091017.109

Intra-continental spread of human invasive Salmonella Typhimurium pathovariants in

Salmonella Enteritidis lineages associated with enterocolitis in high-income settings and
doi:10.1038/ng.3644

25. Morpeth SC, Ramadhani HO, Crump JA. Invasive non-Typhi Salmonella disease in

association between rainfall and non-typhoidal Salmonella bloodstream infections in
21617. doi:10.1038/s41598-021-01030-x

27. Gelaw AK, Nthaba P, Matle I. Detection of Salmonella from animal sources in South
2018;89. doi:10.4102/jsava.v89i0.1643

Prevalence of Campylobacter and Salmonella in African food animals and meat: A
systematic review and meta-analysis. International Journal of Food Microbiology.

multidrug-resistant non-typhoidal Salmonella genotypes among
patients and asymptomatic carriers in endemic informal settlements in Kenya. PLOS
Neglected Tropical Diseases. 2020;14: e0008440. doi:10.1371/journal.pntd.0008440

al. Characteristics of Salmonella recovered from stools of children enrolled in the
Global Enteric Multicenter Study. Clinical Infectious Diseases. 2021 [cited 10 Feb
2021]. doi:10.1093/cid/ciab051

Differences between Non-Typhoidal Salmonella (NTS) Recovered from Children and
Animals Living in Close Contact in The Gambia. PLOS Neglected Tropical Diseases.
2011;5: e1148. doi:10.1371/journal.pntd.0001148

Investigating the Meat Pathway as a Source of Human Nontyphoidal Salmonella


