The Journal of Infectious Diseases

MAJOR ARTICLE

Shigella serotypes associated with carriage in humans establish persistent infection in zebrafish

Vincenzo Torraca[1,2,*], Dominik Brokatzky[1], Sydney L. Miles[1], Charlotte E. Chong[3], P. Malaka De Silva[3], Stephen Baker[4], Claire Jenkins[5], Kathryn E. Holt[1,6], Kate S. Baker[3], Serge Mostowy[1,]**

[1]Department of Infection Biology, London School of Hygiene and Tropical Medicine, London, United Kingdom; [2]School of Life Sciences, University of Westminster, London, United Kingdom; [3]Clinical Infection, Microbiology and Immunity, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, United Kingdom; [4]Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom; [5]National Infection Service, Public Health England, Colindale, London, United Kingdom; [6]Department of Infectious Diseases, Central Clinical School, Monash University, Melbourne, Victoria, Australia

Shigella represents a paraphyletic group of enteroinvasive *Escherichia coli*. More than 40 *Shigella* serotypes have been reported. However, most cases within the MSM (men who have sex with men) community are attributed to three serotypes: *Shigella sonnei* unique serotype and *Shigella flexneri* 2a and 3a serotypes. Using the zebrafish model, we demonstrate that *Shigella* can establish persistent infection *in vivo*. Bacteria are not cleared by the immune system and become antibiotic-tolerant. Persistence depends on O-Antigen, a key constituent of the bacterial surface and serotype determinant. Representative isolates associated with MSM transmission persist in zebrafish, while representative isolates of a serotype not associated with MSM transmission do not. Isolates of a *Shigella* serotype establishing persistent infections elicited significantly less macrophage death *in*

_____

*Alternate Corresponding Author: Vincenzo Torraca, Department of Infection Biology, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, United Kingdom. Email: vincenzo.torraca@lshtm.ac.uk

**Corresponding Author: Serge Mostowy, Department of Infection Biology, London School of Hygiene and Tropical Medicine, Keppel St, London WC1E 7HT, United Kingdom. Email: serge.mostowy@lshtm.ac.uk

vivo than isolates of a serotype unable to establish persistence. We conclude that zebrafish are a valuable platform to illuminate factors underlying establishment of *Shigella* persistent infection in humans.

Keywords: O-Antigen; Persistent infection; *Shigella flexneri; Shigella sonnei*; Serotype; Zebrafish

BACKGROUND

Shigellosis is a diarrhoeal disease caused mainly by *Shigella sonnei* and *Shigella flexneri*[1]. Shigellosis is normally managed with rehydration therapy and antibiotics (fluoroquinolones, extended spectrum beta-lactams, azithromycin[2]). However, antimicrobial resistance (AMR) is widespread and *Shigella* is considered an AMR priority pathogen[3,4].

Infection from *Shigella* is viewed to be self-limiting and short-lived. In high-income countries, sexual transmission of shigellosis in men who have sex with men (MSM) is the major route of domestic dissemination. In a recent retrospective cohort study performed in the United Kingdom, *Shigella* was longitudinally sampled from patients. In a subset of cases, a small single nucleotide polymorphism (SNP) distance (<8 nucleotides) between serial isolates was determined[5], which is suggestive of persistent carriage. The MSM population where serial detection has been frequently identified is also at increased risk of HIV and reinfections from sexual partners. However, symptomatic and asymptomatic carriage of *Shigella* has also been reported in children and the general adult population in different settings[6–10]. Pathogens recognised to establish persistent infection represent a major public health burden; prominent examples include *Helicobacter pylori*, *Salmonella enterica* and *Mycobacterium tuberculosis*[11–14]. Persistent infections are not completely cleared by the immune system and are recalcitrant to antimicrobial therapy[15]. These infections can become asymptomatic and promote dissemination upon disease reactivation.

The zebrafish is a widely adopted model, and valuable to investigate infection by a variety of human bacterial pathogens including enterobacteria (*Escherichia coli*, *Salmonella* and *Shigella*)[16,17]. Zebrafish larvae have an innate immune system highly homologous to that of humans, and their transparency enables high-resolution intravital imaging of fluorescently-tagged immune cells and bacteria *in vivo*[16,18]. Here, using a *Shigella*-zebrafish infection model, we discover that clinical isolates of *Shigella* can establish a persistent and antibiotic-tolerant infection in zebrafish for at least 6 days. Our results demonstrate a new role for *Shigella* O-Antigen (O-Ag) variants in enabling persistence and highlight zebrafish as an animal model that can be used to investigate the surge of three dominant *Shigella* serotypes circulating in the MSM community[19–21].

2

METHODS

Bacterial strains

Wild-type and genetically modified bacterial strains are detailed in **Table S1** and **Table S2**. Parental strains were made fluorescent by electroporation of pFPV25.1 (GFP-labelled) or pFPV-mCherry (mCherry-labelled), conferring Carbenicillin resistance. Exception was made for *S. sonnei* 373976, which was intrinsically resistant to high concentrations of Carbenicillin (MIC>8,000μg/ml). Bacterial mutant strains were previously described[22].

Zebrafish model and animal experimentation guidelines

Animal experiments were approved by the Home Office (PPL P4E664E3C) and performed following the Animals (Scientific Procedures) Act 1986. Wild-type AB zebrafish were used for survival assays and CFU quantification experiments. For experiments involving macrophage imaging, the transgenic line *Tg(mpeg1::Gal4-FF)*[gl25]/*Tg(UAS::LIFEACT-GFP)*[mu271] was used. Eggs were obtained by naturally spawning and larvae were maintained at 28.5°C in embryo medium (0.5x E2 medium supplemented with 0.5ppm of methylene blue). For injections and live microscopy, larvae were anaesthetized with 200μg/ml tricaine (Sigma-Aldrich) in embryo medium. Infected larvae were monitored up to 6 days post-infection. Sex was not determined, as experiments were concluded before zebrafish sexual development.

Zebrafish infection

Bacteria were cultured at 37°C overnight in Trypticase Soy Broth supplemented (when appropriate) with 100μg/ml Carbenicillin, diluted 50x in fresh medium, and grown to Log phase. For inoculum preparation, bacteria were spun down, washed in Phosphate Buffer Saline (PBS) and resuspended to an $OD_{600}=2$ in an injection buffer containing 2% polyvinylpyrrolidone (Sigma-Aldrich) and 0.5% phenol red (Sigma-Aldrich) in PBS. Unless otherwise specified, 1nl (corresponding to ~1,000 CFU) of bacterial suspension was microinjected in the hindbrain ventricle (HBV) of zebrafish larvae at 3 days post-fertilisation (dpf).

Survival assays and bacterial burden

Larvae failing to produce a heartbeat during a 30-second observation were considered nonviable. For larvae beyond 5 dpf, clinical scoring criteria were applied to identify humane endpoints. Briefly, larvae were discontinued when bacterial load was predictive of death in the following 24h (i.e., hindbrain filled with bacteria and/or systemic dissemination) or when larvae were irresponsive to touch. These larvae were withdrawn from the experiment, euthanised and considered nonviable at following time points.

For CFU enumeration, larvae were washed in PBS, anesthetised, and mechanically homogenised in 200μl (non-persistent stages) or 40μl (persistent stages) of PBS. Homogenates were serially

3

diluted and plated onto Trypticase Soy Agar containing 0.01% Congo red (Sigma-Aldrich). Only viable larvae were used for CFU analysis.

Clonality assay

Larvae were infected as previously specified but with an inoculum of GFP-*S. sonnei* and mCherry-*S. sonnei* in a 1:1 ratio. The ratio between the two strains was determined by CFU plating at 0 hours post-infection (hpi) and 144 hpi. At 0 hpi, dilutions of samples were made to correct for differences in the occurrence of stochastic bottlenecks.

Microscopy

For *in vivo* time-lapse imaging, larvae were immobilised in 1% low-melting-point agarose. For high-resolution confocal microscopy, larvae were positioned in 35-mm-diameter glass-bottom MatTek dishes and imaging was performed using a Zeiss LSM 880 with fast Airyscan and 20× or 40× water immersion objectives. Image files were processed using Fiji ImageJ. Co-localisation % between macrophages and bacteria was determined from fluorescent images. Z-stacks encompassing bacterial signal were projected in a bi-dimensional image. The mCherry channel (*S. sonnei*) and the GFP channel (macrophages) were extracted and converted to binary masks. The % of mCherry area overlapping with the GFP area was calculated.

For imaging fixed samples, larvae were fixed in 2% paraformaldehyde in PBS at 4°C overnight, followed by 2 times washing with PBS and 15min permeabilisation with 1x Proteinase K (Component H from Click-iT Plus TUNEL Assay C10619, ThermoFisher Scientific) at room temperature (RT). Larvae were then washed in PBS, stained with Hoechst 33342 1:500 in water for 30min RT, washed and transferred to 60%(v/v) glycerol via gradient washes.

Drug treatments

Nalidixic acid (Nal) was added to embryo medium to reach the desired concentration. In pilot experiments, a range of concentrations was tested (0.5-4.0µg/ml) and 1µg/ml was the lowest concentration tested that fully prevented bacterial growth and death of larvae upon infection with a lethal dose of *S. sonnei* 53G.

Flow cytometry and Caspase-1 activity quantification

A pool of 3 dpf *Tg(mpeg1::Gal4-FF)*[gl25]/*Tg(UAS::LIFEACT-GFP)*[mu271] larvae were infected systemically with 10,000 CFU of mCherry-*S. sonnei* 53G, mCherry-*S. flexneri* M90T or PBS containing mock solution. At 4 hpi larvae were dissociated by treatment with 1ml of 4% trypsin for 15min at 28°C. Single-cell dissociation was facilitated by mechanical disruption using a P1000 pipette after trypsin treatment. Dissociated cells were harvested by centrifugation (5min, 800xg, RT), washed in calcium-free PBS, separated by passage on a 4µm cell strainer and suspended in 500µl staining solution for active caspase-1 (FLICA® 660 Caspase-1 assay, probe: 660-YVAD-FMK, Immunochemistry technologies, #9122) prepared as per manufacturer's guidelines. Upon

staining, cells were pelleted, washed in PBS and fixed in 4% paraformaldehyde overnight. For flow cytometry, cells were washed with 2ml PBS twice and resuspended in 300μl PBS. Active caspase-1 staining of single macrophages was measured on a LSRII (BD Biosciences) and data analysed using FlowJo v10.7.1.

DNA extraction and sequencing

Genomic DNA was extracted from overnight bacterial cultures using MasterPure Complete DNA Purification Kit (Lucigen Corporation) according to manufacturer's instructions. DNA concentration, purity and quantity were assessed using Nanodrop spectrophotometry (DeNovix) and Qubit fluorometer (Invitrogen) according to manufacturers' instructions. Sequencing libraries were prepared using the ligation sequencing kit SQK-LSK109 kit according to manufacturer's instructions (Oxford Nanopore Technologies) with modifications to input DNA (amounts were increased at least two-fold at the initial step). DNA libraries were sequenced using a MinION sequencer and FLO-MIN106 Flow Cell R9.4.1 (Oxford Nanopore Technologies).

Assembly and annotation

The fast5 read files generated from the MinION instrument were base-called and demultiplexed with Guppy v5.0 (Oxford Nanopore Technologies). Processed read files were filtered using Filtlong v0.2.0 and assembled using Flye v2.9[23]. Racon v1.5.0[24] was used to polish contigs with nanopore reads and Medaka v1.6.1 (https://github.com/nanoporetech/medaka) was used to polish Racon polished contigs, with nanopore reads specifying the model r941_min_sup_g507. Polypolish v0.5.0 was used to polish with Illumina reads, where available[25]. The quality and statistics of each assembly were evaluated with QUAST v4.4.0 without a reference genome[26]. Genomes were annotated using Prokaryotic Genome Annotation Pipeline (PGAP). The complete genome sequence data have been submitted to NCBI and deposited at GenBank under BioProject number PRJNA869897.

Phylogenetic tree construction

Annotations of genomes and tree constructions were performed using Genome Annotation services made available by the Bacterial and Viral Bioinformatics Resource Centre (BV-BRC, https://www.bv-brc.org/)[27]. For genome annotation, parameters were set as Annotation recipe: Bacteria/Archaea; Taxonomy name: *Shigella*. Phylogenetic trees were constructed using the Codon Tree building method and parameters were set as Number of genes: 1,000; Max allowed deletions: 1; Max allowed duplications: 1. Data included complete genomes of *S. sonnei* and *S. flexneri* isolates (publicly available via BV-BRC database as of November 2022), and draft genomes of all isolates sequenced/re-sequenced and tested in this study (Table S1). In cases where *S. flexneri* serotype was unknown, this was predicted using ShigEiFinder[28]. Trees also include draft genomes for isolates reported to represent likely cases of carriage in humans with a distance between collected isolates of at least 10 days[29]. Since this dataset represented pairs of isolates with <8 SNP distance, only sequences from second isolates were used. Data were annotated, edited and

5

visualised using FigTree v1.4.4. (http://tree.bio.ed.ac.uk/software/figtree/), iTOL (https://itol.embl.de/) and Inkscape 1.2.2 (https://inkscape.org/).

Quantification and statistical analysis

Statistical tests were performed using Prism (GraphPad Software) or Microsoft Excel. Statistical significance of survival curves was determined using log-rank Mantel-Cox test. The difference in distribution at different time points in the clonality assay was determined by chi-squared test. In all other cases, statistical significance was determined using an unpaired two-tailed t-test or one-way ANOVA with Sidak correction. Analyses were performed on Log10-transformed valued for CFU counts. Data in bar plots are represented as mean ± standard errors of the mean (SEM).

RESULTS

Shigella can establish persistent infection in zebrafish

To test if *Shigella* can establish persistent infection in zebrafish, we injected a low dose (~1,000 CFU, characteristically eliciting ~20% host death by 72 hpi) of mCherry-*S. sonnei* 53G in the hindbrain ventricle (HBV) of zebrafish larvae at 3 dpf and assessed bacterial burden and host survival daily up to 144 hpi (**Figure 1A-C**, **Figure S1A-C**, **Dataset S1**). Quantification of bacterial burden (**Figure 1B**) and linear regression analysis (**Figure S1A**) indicate that infection progresses in three distinct phases: an acute phase characterised by bacterial replication (0-24 hpi), a clearing phase characterised by a significant decrease of bacteria at a constant rate (24-96 hpi), and a persistent phase where few bacteria (<5% of the initial bacterial load) are maintained over an extended period of time (96-144 hpi).

Persistent infections are often recalcitrant to antibiotic treatment. To test if the persistent bacterial population is tolerant to antibiotics, we established a therapeutic dose to treat *S. sonnei* infection of zebrafish larvae with Nal. 1μg/ml Nal was the lowest tested dose that prevented the death of all larvae exposed to a lethal dose of *S. sonnei* (~8,000 CFU, characteristically eliciting ~80% host death by 72 hpi) as well as bacterial proliferation within the host (**Figure S1D,E**). 0.5μg/ml Nal was also tested but this was ineffective and unable to control host death or bacterial proliferation within the host (**Figure S1D,E**). Higher doses of 2 and 4μg/ml were able to prevent bacterial proliferation, but treatments led to more phenotypic aberrations in exposed larvae compared to larvae treated with 1μg/ml (both infected and uninfected, **Figure S1F**), and led to death due to toxicity (**Figure S1D**) rather than bacterial burden (**Figure S1E**). We therefore adopted the minimal therapeutic dose of 1μg/ml Nal for all other experiments.

Treatment of infection with 1μg/ml Nal prevented host death in response to low dose (**Figure 1B,C)** and high dose (**Figure S1D,E**) infections. It also inhibited bacterial growth in both cases (**Figure 1B, Figure S1E)** and therefore larvae did not experience an acute phase of infection (as compared to untreated larvae); instead, they immediately progressed towards a clearing phase

(**Figure 1B, Figure S1A-C**). In the presence of 1μg/ml Nal, the rate of bacterial clearance is not significantly different from that of untreated larvae (**Figure S1A-C**), where infection is not fully cleared and has progressed towards a persistence phase. To investigate if bacteria infecting zebrafish in the persistent phase had developed antibiotic tolerance (irrespective of previously being treated with an antibiotic), we treated larvae exhibiting acute and persistent *Shigella* infection with 1μg/ml Nal for 24h and found that treatment significantly decreased bacterial burden (~100 fold) during acute infection (**Figure 1D**). In contrast, bacterial burden in larvae in the persistence phase did not significantly decrease with antibiotic treatment (**Figure 1D**). We observed that spontaneous evolution of antibiotic resistance occurred in 0.53% of larvae (at 72 hpi) undergoing antibiotic treatment, while it was not observed in untreated larvae (**Figure S1G**). This indicates that failure of 1μg/ml Nal to clear persistent bacteria in most larvae (99.47%) results from antibiotic tolerance.

To study whether persistent bacteria represent a clonally expanded population or derive from a stochastic reduction of the bacterial load, we co-injected *S. sonnei* 53G labelled with either GFP or mCherry (and otherwise isogenic) at a 1:1 ratio (**Figure 2A,B**). As expected, at 0 hpi both strains could be recovered from most larvae (~90%) at similar loads (difference in the recovery of the two strains ≤20%). Strikingly, at 144 hpi, most larvae (~80%) dominantly carried only one of the two strains (difference in the recovery of the two strains >80%). These results indicate that persistent *Shigella* infections represent a clonal population.

Together, these data show that *Shigella* can establish persistent infection *in vivo*, characterised by poor host clearance over an extended period, antibiotic tolerance and clonality.

Shigella O-Antigen variants circulating in the MSM population are associated with persistence *in vivo*

To investigate bacterial factors required for persistence *in vivo*, we performed infections using various *S. sonnei* mutants including a Type III Secretion System (T3SS) deficient strain (ΔMxiD), an O-Ag deficient strain (ΔO-Ag) and a strain having lost the virulence plasmid (-pSS). We observed that the ΔMxiD strain can establish persistent infection *in vivo*, however, the ΔO-Ag and -pSS strains cannot (**Figure 3A,B**, **Figure S2A,B**). Considering that the O-Ag of *S. sonnei* is encoded by the pSS plasmid (-pSS strains are also O-Ag-deficient), we conclude that O-Ag (and not T3SS) is essential for *Shigella* persistence. In agreement, we tested other isolates of *S. sonnei* (all iso-serotypic), including recently collected clinical isolates. While host survival rates vary upon infection with different isolates, all strains established persistent infections at similar levels (**Figure 3C,D**, **Figure S2C,D**).

More than 40 different serotypes of *Shigella* have been reported, and O-Ag structure/composition is a major variable in determining these different serotypes[30]. Only three serotypes have been directly associated with carriage in the MSM population[5,29]. To test whether strain clusters associated with carriage in humans establish persistence in zebrafish, we constructed phylogenetic

7

trees for *S. sonnei* and *S. flexneri* using high-quality complete genomes deposited in the BV-BRC database, sequencing data collected from cases of likely carriage[29], and newly collected sequencing data from *S. sonnei* and *S. flexneri* clinical isolates selected for further testing in zebrafish (**Figure 4A,B**). For the construction of trees, we used the Codon Tree method and the Phylogenetic tree-building service made available by BV-BRC[27]. Based on this phylogenetic analysis, we selected *Shigella* isolates that represented either a persistence[20] ingroup (i.e., falling in a genetic cluster that encompasses isolates previously associated with carriage in humans[29]) or a persistence outgroup (i.e., not directly clustering with isolates previously associated with carriage in humans) for zebrafish infection.

For *S. sonnei*, different sub-lineages have been identified[2,31,32], however only one *S. sonnei* serotype exists (**Figure 4A**), which has comparable genetic diversity to individual serotypes of *S. flexneri* (e.g., *S. flexneri* 2a) (**Figure 4B**)[33,34]. Both representatives of the persistence ingroup and the persistence outgroup of *S. sonnei* persist in zebrafish *in vivo*. This finding is not surprising, considering that *S. sonnei* isolates have short genetic distances and all share an identical O-Ag that defines the sole *S. sonnei* serotype (**Figure 3C,D**, **Figure S2C,D**). In the case of *S. flexneri*, multiple serotypes have been identified[35], although only 2 serotypes (2a and 3a) have been so far associated with persistent carriage[5] (**Figure 4B**). Strikingly, representative strains of these serotypes also persist in zebrafish (**Figure 5A-D**, **Figure S3A,B**). In contrast, 2 isolates belonging to the 5a serotype (*S. flexneri* M90T and SRR12769770, representing a persistence outgroup) are unable to persist in zebrafish (**Figure 5A-D**, **Figure S3A,B**). Overall, these results indicate that O-Ag type is associated with establishment of persistence. Our findings are consistent with epidemiological evidence showing that transmission in the MSM community is associated with *S. sonnei*, *S. flexneri* 2a,* and *S. flexneri* 3a serotypes, but not *S. flexneri* 5a[5].

Shigella can establish persistent infection in macrophages *in vivo*

Macrophages are viewed as a first line of host defence against *Shigella* infection and have been reported to act as a long-term reservoir for intracellular bacterial pathogens, such as *Salmonella enterica* and *M. tuberculosis*[13,14,36]. Having established a zebrafish model of *Shigella* persistent infection, we sought to determine whether this was localised to macrophages. Using the *Tg(mpeg1::Gal4-FF)*[gl25]/*Tg(UAS::LIFEACT-GFP)*[mu271] transgenic zebrafish line (labelling macrophages) and mCherry-labelled *S. sonnei* 53G, we identified ~40% of bacterial fluorescence co-localising with macrophages at 144 hpi (**Figure 6A,B, Video S1**). Longitudinal studies using high-resolution confocal microscopy showed that macrophages harbouring a stable bacterial load (i.e., bacterial fluorescence shows only a 1.21±0.21 fold change in 12h) can be detected as early as 24 hpi (**Figure 6C, Video S2**) and lasted for several consecutive days (**Figure S4A-C**). These data were surprising, considering that macrophage cell death is widely recognised as a hallmark of *Shigella* infection[17,37,38]. *Shigella* establishing persistent infection inside macrophages did not maintain the classical rod shape which can be observed before phagocytosis and at early time points following macrophage invasion (e.g., at 4 hpi)[22] (**Figure 6A**, **Figure S4D,E**).

Previous work indicated that *S. sonnei* may be less efficient in inducing macrophage death than *S. flexneri* M90T (5a) *in vitro* using THP1 cells[37,39]. We hypothesised that this reduced ability to induce macrophage death may explain the ability to establish persistent infection of macrophages. To test this *in vivo*, we infected zebrafish larvae systemically (an infection route that promotes *Shigella*-macrophage interactions[39,40]) with *S. sonnei* 53G or *S. flexneri* 5a M90T. At 4 hpi, we dissociated infected larvae and PBS injection control larvae and quantified the level of caspase-1 activity (a marker of macrophage pyroptotic cell death) by flow cytometry (**Figure 6D**). Considering that this assay measures differences in caspase-1 activity at the whole macrophage population level, differences between groups appear small because only a fraction of the macrophage population is infected. Despite this limitation, infection with *S. flexneri* 5a M90T led to significantly higher levels of caspase-1 activity compared to infection with *S. sonnei* 53G, consistent with the possibility that persistent *Shigella* serotypes elicit significantly less macrophage death *in vivo* than serotypes unable to establish persistence.

DISCUSSION

The zebrafish infection model has been instrumental to study host-pathogen interactions in response to a wide variety of human pathogens, including different *Shigella* species[22,41–43]. Here, we show that different *Shigella* serotypes (*S. sonnei*, *S. flexneri* 2a and *S. flexneri* 3a) that are highly prevalent in the MSM community and associated with direct host-to-host transmission, also establish persistent infection in zebrafish. In contrast, the *S. flexneri* 5a serotype, which is less prevalent in the MSM community, is unable to establish persistent infection *in vivo*. There are likely multiple factors that may explain why the *S. flexneri* 5a serotype is less prevalent in the MSM community and worldwide. Based on our evidence, we propose that the inability of the *S. flexneri* 5a serotype to persist could be a significant factor underlying its replacement with other more prevalent serotypes able to establish persistence.

Our results show that O-Ag variants promote the establishment of persistent infection, but it is not yet known how the O-Ag precisely contributes. *S. sonnei* O-Ag is unique among enterobacterial pathogens and does not share obvious structural homology with the O-Ag expressed by *S. flexneri* 2a or 3a. However, *Shigella* species are a remarkable example of convergent evolution[33] and the persistence phenotype mediated by different O-Ag serotypes may represent an example of pathoadaptive convergence due to occupation of the same niche. During the persistence phase, bacteria frequently co-localise with macrophages *in vivo*, suggesting that macrophages represent a preferred niche for *Shigella* persistence. *Shigella* establishing persistent infection often assumed pleomorphic shapes and lost their rod shape following macrophage invasion. This is in agreement with previous work demonstrating that several bacterial species transition to L-forms within macrophages (for example in response to lysozyme[44]) and during zebrafish infection[45].

Isolates of *S. sonnei* serotype (which is associated with persistence) induce significantly less macrophage death *in vivo* as compared to isolates of *S. flexneri* serotype 5a (which has not been

9

linked to persistence and MSM transmission). Considering that persistent infection of zebrafish recapitulates the epidemiological trend of persistence in humans, we propose that zebrafish can be used to discover host and pathogen factors underlying *Shigella* persistent infection in humans.

All authors declare that they have no conflicts of interest.

Author contributions:

This project was conceived by V.T. and S.M. V.T. and S.L.M. performed all zebrafish experiments and analysed all datasets. V.T., S.L.M., C.E.C. and M.D.S. prepared DNA samples and analysed sequencing results. V.T. constructed the phylogenetic trees with inputs from K.E.H., S.L.M. and K.B. V.T. and D.B. performed experiments to detect active caspase-1 in zebrafish. S.B. and C.J. provided valuable strains for the work. V.T. and S.M. wrote the manuscript with input from all authors.

FOOTNOTES

Data availability: *Shigella* GM strains generated in this study are available upon request. Sequencing data generated in this study have been deposited in the Sequencing Read Archive at BioProject number PRJNA869897 and are publicly available as of the date of publication. Accession numbers are listed in the key resources table and against individual strains in **Table S1**.

References

1. Thompson, C.N., Duy, P.T., and Baker, S. (2015). The Rising Dominance of *Shigella sonnei*: An Intercontinental Shift in the Etiology of Bacillary Dysentery. PLoS Negl Trop Dis *9*, e0003708. 10.1371/journal.pntd.0003708.

2. Torraca, V., Holt, K., and Mostowy, S. (2020). *Shigella sonnei*. Trends Microbiol *28*, 696–697. 10.1016/j.tim.2020.02.011.

3. World Health Organization (2017). Global Priority List of Antibiotic-Resistant Bacteria to Guide Research, Discovery, and Development of New Antibiotics (WHO Press).

4. Baker, S., and Scott, T.A. (2023). Antimicrobial-Resistant *Shigella*: Where Do We Go Next? Nat Rev Microbiol. 10.1038/S41579-023-00906-1.

5. Allen, H., Mitchell, H.D., Simms, I., Baker, K.S., Foster, K., Hughes, G., Dallman, T.J., and Jenkins, C. (2021). Evidence for Re-Infection and Persistent Carriage of *Shigella* Species in Adult Males Reporting Domestically Acquired Infection in England. Clin Microbiol Infect *27*, 126.e7-126.e13. 10.1016/j.cmi.2020.03.036.

6. McMurry, T.L., Mcquade, E.T.R., Liu, J., Kang, G., Kosek, M.N., Lima, A.A.M., Bessong, P.O., Samie, A., Haque, R., Mduma, E.R., et al. (2021). Duration of Postdiarrheal Enteric Pathogen Carriage in Young Children in Low-Resource Settings. Clin Infect Dis *72*, E806–E814. 10.1093/CID/CIAA1528.

7. Collard, J.M., Andrianonimiadana, L., Habib, A., Rakotondrainipiana, M., Andriantsalama, P., Randriamparany, R., Rabenandrasana, M.A.N., Weill, F.X., Sauvonnet, N., Randremanana, R.V., et al. (2022). High Prevalence of Small Intestine Bacteria Overgrowth and Asymptomatic Carriage of Enteric Pathogens in Stunted Children in Antananarivo, Madagascar. PLoS Negl Trop Dis *16*. 10.1371/JOURNAL.PNTD.0009849.

8. Horwood, P.F., Soli, K.W., Maure, T., Naito, Y.I., Morita, A., Natsuhara, K., Tadokoro, K., Baba, J., Odani, S., Tomitsuka, E., et al. (2017). A High Burden of Asymptomatic Gastrointestinal Infections in Traditional Communities in Papua New Guinea. Am J Trop Med Hyg *97*, 1872–1875. 10.4269/AJTMH.17-0282.

9. Kotloff, K.L., Nataro, J.P., Blackwelder, W.C., Nasrin, D., Farag, T.H., Panchalingam, S., Wu, Y., Sow, S.O., Sur, D., Breiman, R.F., et al. (2013). Burden and Aetiology of Diarrhoeal Disease in Infants and Young Children in Developing Countries (the Global Enteric Multicenter Study, GEMS): A Prospective, Case-Control Study. Lancet *382*, 209–222. 10.1016/S0140-6736(13)60844-2.

10. Bliss, J., Bouhenia, M., Hale, P., Couturier, B.A., Iyer, A.S., Rumunu, J., Martin, S., Wamala, J.F., Abubakar, A., Sack, D.A., et al. (2018). High Prevalence of *Shigella* or Enteroinvasive *Escherichia coli* Carriage among Residents of an Internally Displaced Persons Camp in South Sudan. Am J Trop Med Hyg *98*, 595–597. 10.4269/AJTMH.17-0339.

11. Danielli, A., Amore, G., and Scarlato, V. (2010). Built Shallow to Maintain Homeostasis and Persistent Infection: Insight into the Transcriptional Regulatory Network of the Gastric Human Pathogen *Helicobacter pylori*. PLoS Pathog *6*, e1000938. 10.1371/JOURNAL.PPAT.1000938.

12. Luk, C.H., Valenzuela, C., Gil, M., Swistak, L., Bomme, P., Chang, Y.-Y., Mallet, A., and Enninga, J. (2021). *Salmonella* Enters a Dormant State Within Human Epithelial Cells for Persistent infection. PLoS Pathog *17*. 10.1371/JOURNAL.PPAT.1009550.

13. Pham, T.H.M., Brewer, S.M., Thurston, T., Massis, L.M., Honeycutt, J., Lugo, K., Jacobson, A.R., Vilches-Moure, J.G., Hamblin, M., Helaine, S., et al. (2020). *Salmonella*-Driven Polarization of Granuloma Macrophages Antagonizes TNF-Mediated Pathogen Restriction during Persistent Infection. Cell Host Microbe *27*, 54-67.e5. 10.1016/J.CHOM.2019.11.011.

14. Hasenoehrl, E.J., Rae Sajorda, D., Berney-Meyer, L., Johnson, S., Tufariello, J.M., Fuhrer, T., Cook, G.M., Jacobs, W.R., and Berney, M. (2019). Derailing the Aspartate Pathway of *Mycobacterium tuberculosis* to Eradicate Persistent Infection. Nat Commun 2019 10:1 *10*, 1–12. 10.1038/s41467-019-12224-3.

15. Grant, S.S., and Hung, D.T. (2013). Persistent Bacterial Infections, Antibiotic Tolerance, and the Oxidative Stress Response. Virulence *4*, 273. 10.4161/VIRU.23987.

16. Torraca, V., and Mostowy, S. (2018). Zebrafish Infection: from Pathogenesis to Cell Biology. Trends Cell Biol *28*, 143–156. 10.1016/J.TCB.2017.10.002.

17. Duggan, G.M., and Mostowy, S. (2018). Use of Zebrafish to Study *Shigella* Infection. Dis Model Mech *11*, dmm032151. 10.1242/dmm.032151.

18. Gomes, M.C., and Mostowy, S. (2020). The Case for Modeling Human Infection in Zebrafish. Trends Microbiol *28*, 10–18. 10.1016/J.TIM.2019.08.005.

19. Dallman, T.J., Charles, H., Prochazka, M., Sinka, K., Hughes, G., Godbole, G., and Jenkins, C. (2021). Emergence of Novel Strains of *Shigella flexneri* Associated with Sexual Transmission in Adult Men in England, 2019–2020. J Med Microbiol *70*, 1437. 10.1099/jmm.0.001437.

20. Baker, K.S., Dallman, T.J., Ashton, P.M., Day, M., Hughes, G., Crook, P.D., Gilbart, V.L., Zittermann, S., Allen, V.G., Howden, B.P., et al. (2015). Intercontinental Dissemination of Azithromycin-Resistant Shigellosis through Sexual Transmission: A Cross-Sectional Study. Lancet Infect Dis *15*, 913–921.

21. Baker, K.S., Dallman, T.J., Field, N., Childs, T., Mitchell, H., Day, M., Weill, F.X., Lefèvre, S., Tourdjman, M., Hughes, G., et al. (2018). Genomic Epidemiology of *Shigella* in the United Kingdom Shows Transmission of Pathogen Sublineages and Determinants of Antimicrobial Resistance. Sci Rep *8*. 10.1038/s41598-018-25764-3.

22. Torraca, V., Kaforou, M., Watson, J., Duggan, G.M., Guerrero-Gutierrez, H., Krokowski, S., Hollinshead, M., Clarke, T.B., Mostowy, R.J., Tomlinson, G.S., et al. (2019). *Shigella sonnei* Infection of Zebrafish Reveals that O-Antigen Mediates Neutrophil Tolerance and Dysentery Incidence. PLoS Pathog *15*. 10.1371/journal.ppat.1008006.

23. Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019). Assembly of Long, Error-Prone Reads Using Repeat Graphs. Nat Biotechnol *37*, 540–546. 10.1038/s41587-019-0072-8.

24. Vaser, R., Sović, I., Nagarajan, N., and Šikić, M. (2017). Fast and Accurate *de novo* Genome Assembly From Long Uncorrected Reads. Genome Res *27*, 737–746. 10.1101/GR.214270.116/-/DC1.

25. Wick, R.R., and Holt, K.E. (2022). Polypolish: Short-Read Polishing of Long-Read Bacterial Genome Assemblies. PLoS Comput Biol *18*. 10.1371/JOURNAL.PCBI.1009802.

26.   Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: Quality Assessment Tool for Genome Assemblies. Bioinformatics *29*, 1072–1075. 10.1093/BIOINFORMATICS/BTT086.

27.   Davis, J.J., Gerdes, S., Olsen, G.J., Olson, R., Pusch, G.D., Shukla, M., Vonstein, V., Wattam, A.R., and Yoo, H. (2016). PATtyFams: Protein Families for the Microbial Genomes in the PATRIC Database. Front Microbiol *7*. 10.3389/FMICB.2016.00118.

28.   Zhang, X., Payne, M., Nguyen, T., Kaur, S., and Lan, R. (2021). Cluster-Specific Gene Markers Enhance *Shigella* and Enteroinvasive *Escherichia coli in silico* Serotyping. Microb Genom *7*. 10.1099/MGEN.0.000704.

29.   Bengtsson, R.J., Dallman, T.J., Allen, H., Malaka De Silva, P., Stenhouse, G., Pulford, C. V., Bennett, R.J., Jenkins, C., and Baker, K.S. (2021). Accessory Genome Dynamics and Structural Variation of *Shigella* from Persistent Infections. mBio *12*. 10.1128/MBIO.00254-21.

30.   Liu, B., Knirel, Y.A., Feng, L., Perepelov, A. V., Senchenkova, S.N., Wang, Q., Reeves, P.R., and Wang, L. (2008). Structure and Genetics of *Shigella* O-Antigens. FEMS Microbiol Rev *32*, 627–653. 10.1111/j.1574-6976.2008.00114.x.

31.   Holt, K.E., Baker, S., Weill, F.-X., Holmes, E.C., Kitchen, A., Yu, J., Sangal, V., Brown, D.J., Coia, J.E., Kim, D.W., et al. (2012). *Shigella sonnei* Genome Sequencing and Phylogenetic Analysis Indicate Recent Global Dissemination from Europe. Nat Genet *44*, 1056–1059. 10.1038/ng.2369.

32.   Hawkey, J., Paranagama, K., Baker, K.S., Bengtsson, R.J., Weill, F.X., Thomson, N.R., Baker, S., Cerdeira, L., Iqbal, Z., Hunt, M., et al. (2021). Global Population Structure and Genotyping Framework for Genomic Surveillance of the Major Dysentery Pathogen, *Shigella sonnei*. Nat Commun *12*, 1–12. 10.1038/s41467-021-22700-4.

33.   Hawkey, J., Monk, J.M., Billman-Jacobe, H., Palsson, B., and Holt, K.E. (2020). Impact of Insertion Sequences on Convergent Evolution of *Shigella* Species. PLoS Genet *16*, e1008931. 10.1371/journal.pgen.1008931.

34.   Bengtsson, R.J., Simpkin, A.J., Pulford, C. V., Low, R., Rasko, D.A., Rigden, D.J., Hall, N., Barry, E.M., Tennant, S.M., and Baker, K.S. (2022). Pathogenomic Analyses of *Shigella* Isolates Inform Factors Limiting Shigellosis Prevention and Control across LMICs. Nat Microbiol *7*, 251–261. 10.1038/s41564-021-01054-z.

35.   Sethuvel, D.P.M., Ragupathi, N.K.D., Anandan, S., and Veeraraghavan, B. (2017). Update on: *Shigella* new Serogroups/Serotypes and their Antimicrobial Resistance. Lett Appl Microbiol *64*, 8–18. 10.1111/LAM.12690.

36.   Nix, R.N., Altschuler, S.E., Henson, P.M., and Detweiler, C.S. (2007). Hemophagocytic Macrophages Harbor *Salmonella enterica* During Persistent Infection. PLoS Pathog *3*, 1982–1992. 10.1371/journal.ppat.0030193.

37.   Zychlinsky, A., Prevost, M.C., and Sansonetti, P.J. (1992). *Shigella flexneri* Induces Apoptosis in Infected Macrophages. Nature 1992 358:6382 *358*, 167–169. 10.1038/358167a0.

38.   Watson, J.L., Sanchez-Garrido, J., Goddard, P.J., Torraca, V., Mostowy, S., Shenoy, A.R., and Clements, A. (2019). *Shigella sonnei* O-Antigen Inhibits Internalization, Vacuole Escape, and Inflammasome Activation. mBio *10*. 10.1128/mBio.02654-19.

39. Mostowy, S., Boucontet, L., Mazon Moya, M.J., Sirianni, A., Boudinot, P., Hollinshead, M., Cossart, P., Herbomel, P., Levraud, J.-P., and Colucci-Guyon, E. (2013). The Zebrafish as a New Model for the *in vivo* Study of *Shigella flexneri* Interaction with Phagocytes and Bacterial Autophagy. PLoS Pathog *9*, e1003588. 10.1371/journal.ppat.1003588.

40. Colucci-Guyon, E., Tinevez, J.Y., Renshaw, S.A., and Herbomel, P. (2011). Strategies of Professional Phagocytes *in vivo*: Unlike Macrophages, Neutrophils Engulf only Surface-Associated Microbes. J Cell Sci *124*, 3053–3059. 10.1242/JCS.082792.

41. Torraca, V., Masud, S., Spaink, H.P., and Meijer, A.H. (2014). Macrophage-Pathogen Interactions in Infectious Diseases: New Therapeutic Insights from the Zebrafish Host Model. Dis Model Mech *7*, 785–797. 10.1242/DMM.015594.

42. Willis, A.R., Torraca, V., Gomes, M.C., Shelley, J., Mazon-Moya, M., Filloux, A., Lo Celso, C., and Mostowy, S. (2018). *Shigella*-Induced Emergency Granulopoiesis Protects Zebrafish Larvae from Secondary Infection. mBio *9*, e00933-18. 10.1128/mBio.00933-18.

43. Mazon-Moya, M.J., Willis, A.R., Torraca, V., Boucontet, L., Shenoy, A.R., Colucci-Guyon, E., and Mostowy, S. (2017). Septins Restrict Inflammation and Protect Zebrafish Larvae from *Shigella* Infection. PLoS Pathog *13*. 10.1371/JOURNAL.PPAT.1006467.

44. Kawai, Y., Mickiewicz, K., and Errington, J. (2018). Lysozyme Counteracts β-Lactam Antibiotics by Promoting the Emergence of L-Form Bacteria. Cell *172*, 1038-1049.e10. 10.1016/J.CELL.2018.01.021.

45. Mickiewicz, K.M., Kawai, Y., Drage, L., Gomes, M.C., Davison, F., Pickard, R., Hall, J., Mostowy, S., Aldridge, P.D., and Errington, J. (2019). Possible Role of L-Form Switching in Recurrent Urinary Tract Infection. Nat Commun *10*. 10.1038/S41467-019-12359-3.

Figure 1. *Shigella* establishes persistent infection in zebrafish

(**A**) Experimental diagram of *S. sonnei*-zebrafish infection. *S. sonnei* infection undergoes three phases, an acute phase (i.e., increasing bacterial load) a clearing phase (i.e., a steady decrease of bacterial load) and a persistent phase (i.e., bacterial load does not further decrease). (**B,C**) CFU count and survival analysis from zebrafish larvae infected with *S. sonnei* 53G. Larvae were treated with the antibiotic Nalidixic acid (Nal) or were left untreated. *S. sonnei* establishes a persistent infection in zebrafish, even during treatment with a therapeutic antibiotic dose. From 96 hpi, the bacterial load in larvae remains constant over consecutive days. (**D**) CFU count from zebrafish larvae treated with Nal (or left untreated), during acute and persistent infection stages. During the acute infection stage (0 to 24 hpi) *S. sonnei* is sensitive to antibiotic treatment, while during the persistent infection stages (120 to 144 hpi) *S. sonnei* becomes insensitive to antibiotic treatment. Data were analysed at 24 hours post-treatment (hpt). Statistics: unpaired t-test on Log10 transformed data against the previous timepoint (B); Log-rank Mantel-Cox test (C); one-way ANOVA with Sidak's correction (D); ns (non-significant) $p \geq 0.05$; *$p < 0.05$; ***$p < 0.001$.

14

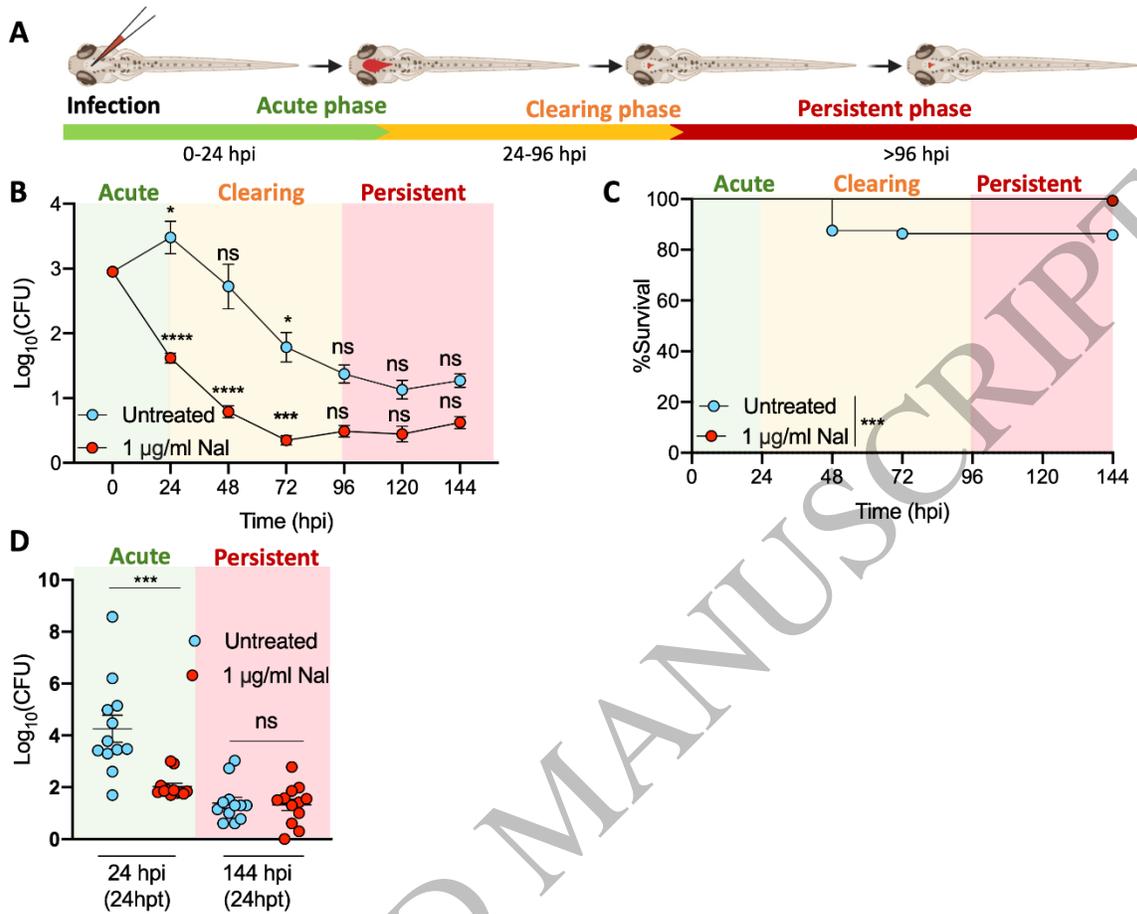Figure 2. Persistent *Shigella* represents a clonally-expanded population

(**A,B**) mCherry and GFP labelled S. *sonnei* 53G were co-injected at a 1:1 ratio. At 0 hpi (A) most larvae (~90%) displayed an expected 1:1 ratio of the co-injected strains (i.e., % difference between strains <20%). At 144 hpi (B) most larvae (~80%) were infected by one of the two strains at a much higher frequency than the other strain (i.e., % difference between strains >80%), indicating the establishment of clonality. Statistics: Differences in distribution between data at 0 and 144 hpi were calculated by chi-squared test (****p<0.0001).
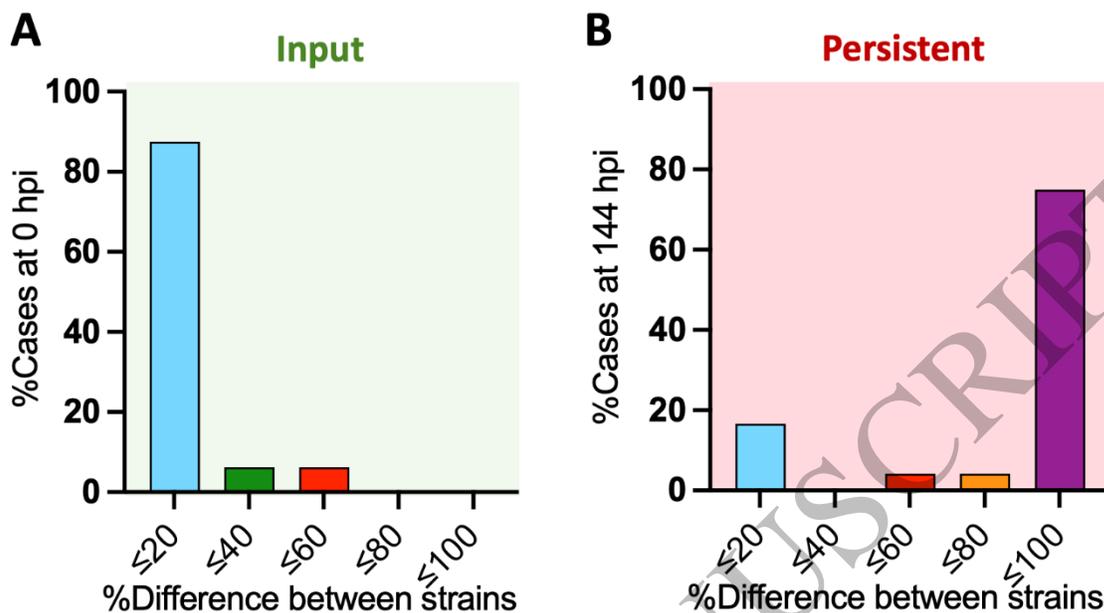
Figure 3. *Shigella* O-Antigen is essential to establish persistent infection

(**A,B**) Percentage of larvae carrying persistent infection at 144 hpi (A) and average bacterial Log_{10}CFU (B), for WT *S. sonnei* 53G and several isogenic mutants. A T3SS mutant (ΔMxiD) can establish persistent infection at similar levels as WT. However, the -pSS mutant (depleted of the virulence plasmid) or an O-Antigen mutant (ΔO-Ag) are significantly reduced in their capability to establish persistent infections. (**C,D**) Percentage of larvae carrying persistent infection at 144 hpi (C) and average bacterial Log_{10}CFU (D) for several *S. sonnei* isolates from Lineage II and Lineage III. Both Lineage II and III *S. sonnei* isolates are capable of establishing persistent infections. Although belonging to different lineages, all *S. sonnei* isolates share an identical O-Antigen. Statistics: one-way ANOVA with Sidak's correction on percentage data (A,C) or Log10-transformed data (B,D); ns (non-significant) $p \geq 0.05$; **$p < 0.01$; ****$p < 0.0001$.
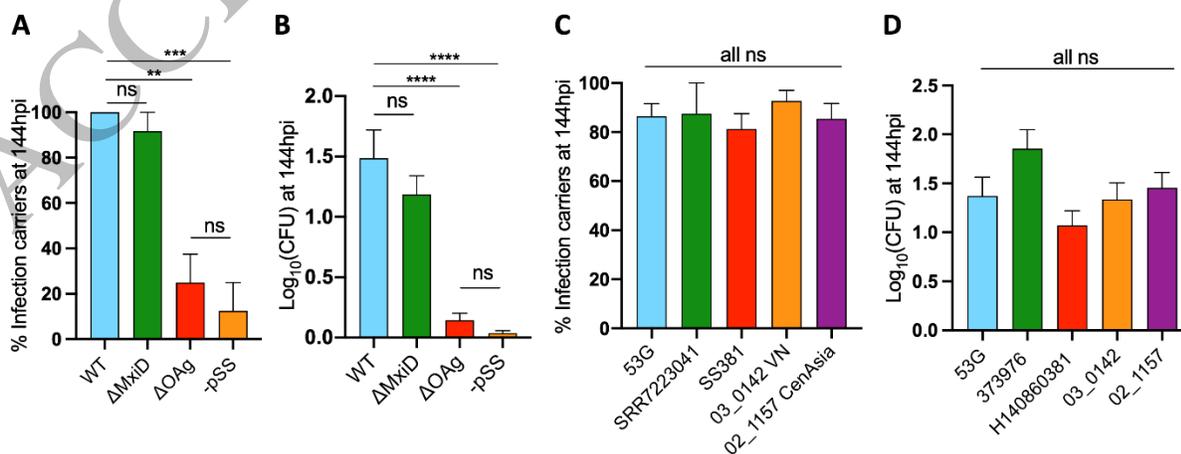

DOI: 10.1093/infdis/jiad326

16

Figure 4. Phylogenetic distribution of persistent *Shigella*

(**A,B**) Phylogenetic trees (codon trees) for *S. sonnei* (A) and *S. flexneri* (B) constructed by using complete reference genomes of known serotypes (available in BV-BRC, https://www.bv-brc.org/), draft genome sequences of isolates associated with persistence carriage in humans[29] and the isolates used in this study (sequenced with nanopore technology). Red labels indicate strains associated with persistent carriage in humans; orange labels indicate strains that established persistent infection in zebrafish; black labels indicate strains that did not establish persistent infection in zebrafish; blue labels indicate other strains available in the BV-BRC database and utilised to reconstruct the phylogenetic trees. For the *S. sonnei* tree, different lineages (II and III) and sub-lineages (3.6. and 3.7) are highlighted with different colours. For the *S. flexneri* tree, the three main clusters encompassing the serotypes investigated in this study (2a, 3, and 5) are highlighted with different colours and the serotype of all the strains is reported in brackets after the strain name.
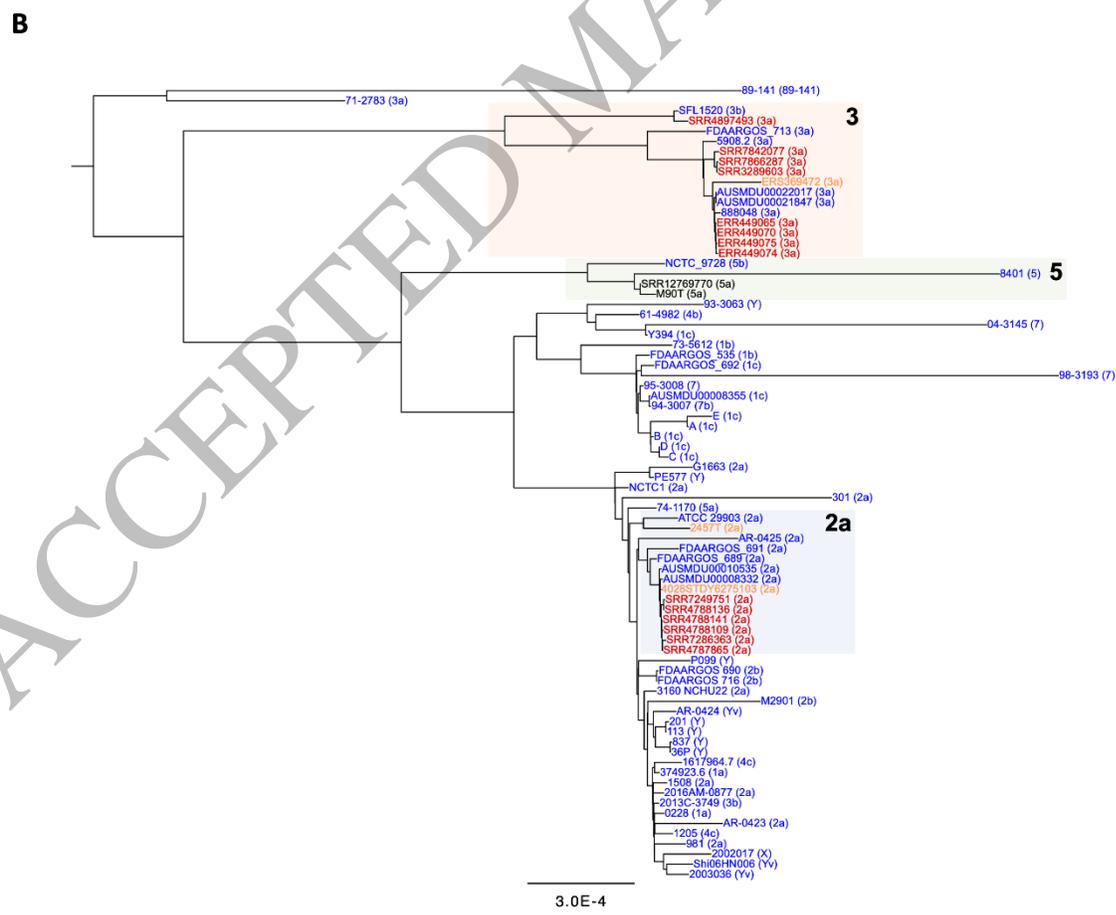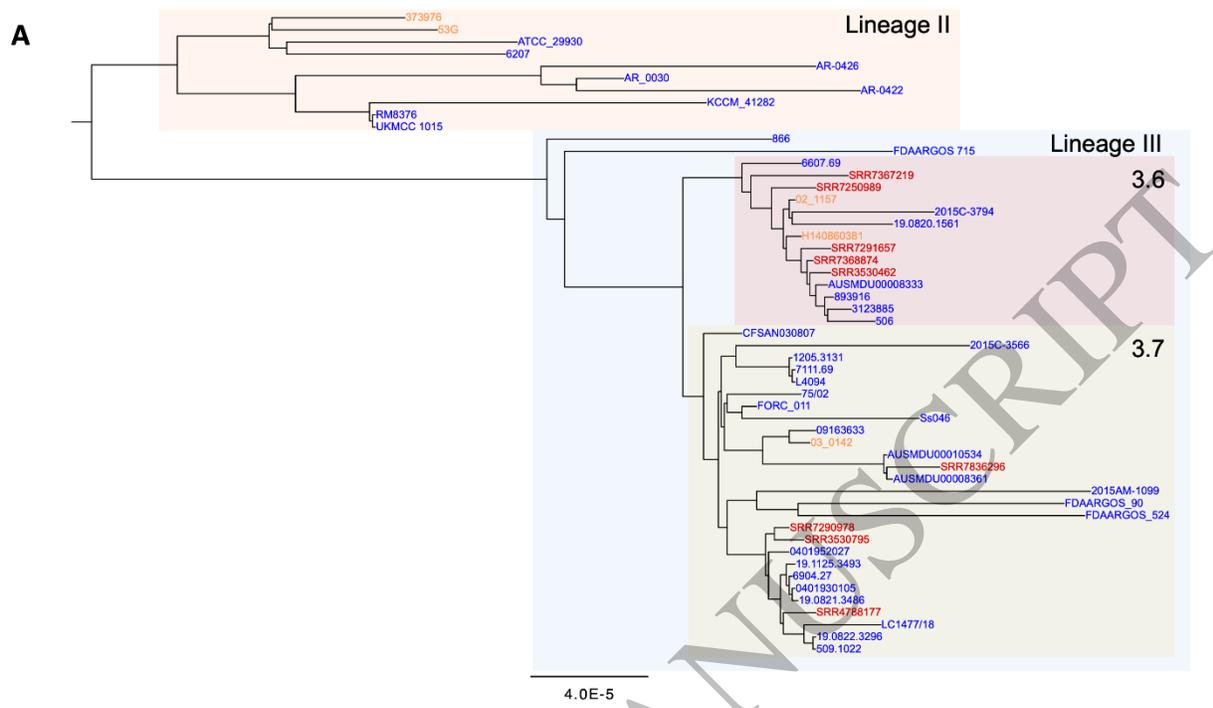
17

Figure 5. Different *Shigella* O-Antigen serotypes associated with MSM transmission enable persistent infection

(**A,B**) Percentage of larvae carrying persistent infection at 144 hpi (A) and average bacterial $Log_{10}CFU$ (B) for several *S. flexneri* isolates from serotypes 2a, 3a and 5a. *S. flexneri* isolates of serotypes 2a and 3a are both capable of establishing persistent infection, to a much greater extent than *S. flexneri* of serotype 5a. (**C,D**) Percentage of larvae carrying persistent infection at 144 hpi (A) and average bacterial $Log_{10}CFU$ (B) for *S. flexneri* M90T (serotype 5a), *S. sonnei* 53G and *S. flexneri* SRR12769770 (a 2020 clinical isolate of *S. flexneri* from serotype 5a). SRR12769770 does not show a significant difference in the establishment of persistence when compared to *S. flexneri* M90T. Statistics: one-way ANOVA with Sidak's correction on percentage data (A,C) or Log10-transformed data (B,D); ns (non-significant) $p \geq 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$.
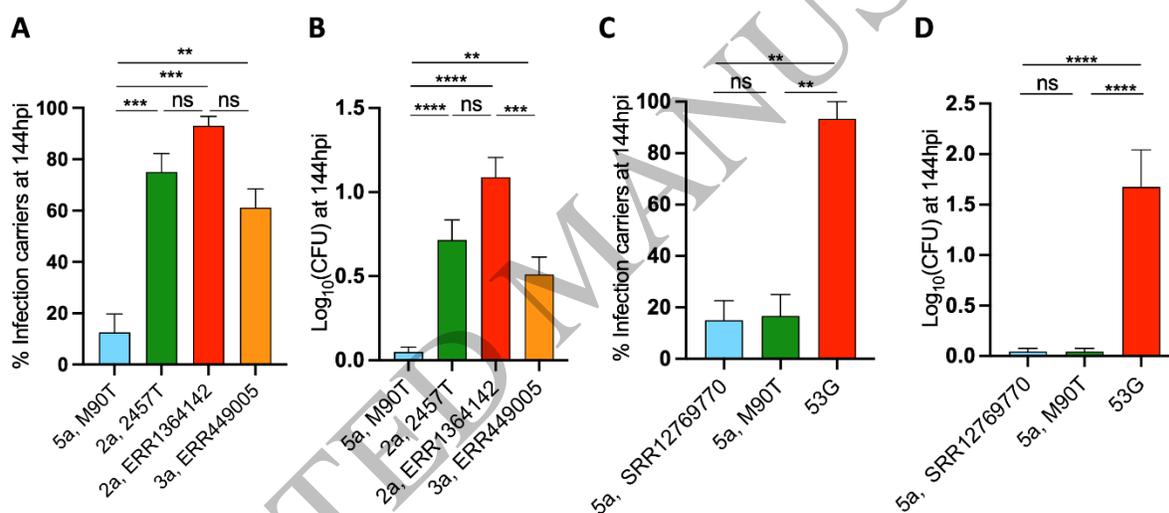
Figure 6. *Shigella* can establish persistent infection of macrophages *in vivo*

(**A**) Head region and individual macrophage detail from a representative *S. sonnei*-infected zebrafish larvae at 144 hpi. *Tg(mpeg1::Gal4-FF)*[gl25]/*Tg(UAS::LIFEACT-GFP)*[mu271] larvae (Mpeg1, with macrophages in green) were injected at 3 dpf in the hindbrain ventricle with 1,000 CFU mCherry-labelled *S. sonnei* 53G (red). An individual infected macrophage harbouring persistent bacteria is magnified. Scalebars: 100μm (left); 10μm (right, inset). (**B**) 40% of bacterial fluorescence co-localises with *mpeg1*⁺ macrophages at persistent infection stages. (**C**) Individual *Tg(mpeg1::Gal4-FF)*[gl25]/*Tg(UAS::LIFEACT-GFP)*[mu271] macrophage (Mpeg1, green) harbouring mCherry-*S. sonnei* (red) followed over time for 12h (from 24 to 36hpi). For images in (C), 1,000 CFU bacteria were delivered systemically in 2 dpf larvae. Scalebar: 10μm. (**D**) *S. sonnei* 53G induces a reduced level of caspase-1 activation *in vivo* in the zebrafish model, as compared to *S. flexneri* M90T. For data in D, 10,000 CFU bacteria were delivered systemically in 3 dpf larvae. Statistics: one-way ANOVA with Sidak's correction; ***p<0.001; ****p<0.0001.