Title

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

Authors

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[1,7]*

Affiliations

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
7. Lead contact

Correspondence

*v.torraca@westminster.ac.uk (V.T.)
*serge.mostowy@lshtm.ac.uk (S.M.)

Highlights

- *Shigella* can establish persistent infection in zebrafish
- Establishment of persistent infection *in vivo* depends on *Shigella* serotype
- The same serotypes establishing persistent infection in zebrafish are prevalent in a patient group in whom persistent infection is observed
- *Shigella* serotypes which establish persistent infection fail to promote macrophage cell death *in vivo*

In Brief

In high-income countries, sexual transmission among men who have sex with men (MSM) is the dominant cause of domestic *Shigella* transmission. Torraca *et al.* discover that three serotypes of *Shigella* associated with MSM transmission also persist in zebrafish. These results indicate that *Shigella* serotypes associated with MSM transmission can establish carriage in the host, and highlight the use of zebrafish infection to study *Shigella* persistent infection in humans.

Summary

Shigella represents a paraphyletic group of human-adapted *Escherichia coli* lineages that converged towards the same enteropathogenic pathovar. In low-income countries, *Shigella* is typically transmitted by ingestion of contaminated water and food, while in high-income countries sexual transmission among men who have sex with men (MSM) is the dominant route of transmission. Although more than 40 different serotypes of *Shigella* have been reported globally, within the MSM community the

majority of reported cases are attributed to three *Shigella* serotypes: the *Shigella sonnei* unique serotype (Ss) and the *Shigella flexneri* serotypes 2a and 3a. Here, using a zebrafish infection model, we demonstrate that *Shigella* can establish persistent infection *in vivo*. In this case, bacteria are not cleared by the host immune system and become tolerant to therapeutic doses of antibiotics. We show that persistence is dependent on the *Shigella* O-Antigen, a key constituent of the bacterial cell surface and determinant of serotype classification. Representative isolates of three *Shigella* serotypes associated with global dissemination and MSM transmission (*S. sonnei* Ss, *S. flexneri* 2a and 3a) all persist in zebrafish, while a serotype not associated with MSM transmission (*S. flexneri* 5a) does not. Strikingly, *Shigella* serotypes which establish persistent infection fail to promote macrophage cell death *in vivo*. We conclude that zebrafish can be a valuable platform to illuminate host and pathogen factors underlying the establishment of persistent infection with *Shigella* in humans.

Introduction

Shigellosis is a diarrhoeal disease caused mainly by *Shigella sonnei* and *Shigella flexneri*[1]. 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 (up to 176 days apart). In a subset of patients, a small single nucleotide polymorphism (SNP) distance (<8 nucleotides apart) between serial isolates was determined[2], which is suggestive of persistent carriage. Symptomatic and asymptomatic *Shigella* carriage, as well as prolonged bacterial shedding, have been indicated by culture and molecular testing methods in different populations (Latin America, Sub-Saharan Africa, South Asia and Oceania) and age categories[3–5]. Pathogens recognised to establish persistent infection represent a major public health burden; prominent examples include *Helicobacter pylori*, *Salmonella enterica* Typhimurium and *Mycobacterium tuberculosis*[6–9]. Persistent infections fail to be cleared by the immune system and are also notoriously recalcitrant to antimicrobial therapy[10]. These infections can become asymptomatic, and together with reactivation at later times (due to a variety of factors, including ageing, immunosuppression and superinfections), can promote future dissemination of the disease.

The zebrafish (*Danio rerio*) is a widely adopted vertebrate model, and valuable to investigate infection by a variety of human bacterial pathogens including enterobacteria, such as *Escherichia coli*, *Salmonella* and *Shigella*[11,12]. From the very early developmental stages, zebrafish larvae have an innate immune system highly homologous to that of humans, and transparent larvae enable high-resolution intravital imaging of fluorescently-tagged immune cells and bacteria *in vivo*[11,13]. 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 (>144 hours). 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 currently circulating in the MSM community[14–16].

Results and Discussion

Shigella can establish persistent infection in zebrafish

To test if *Shigella* can establish persistent infection in zebrafish, we injected a low dose (~1000 CFU, characteristically eliciting ~20% host death by 72 hpi) of mCherry-labelled *S. sonnei* 53G in the hindbrain ventricle (HBV) of zebrafish larvae at 3 days post-fertilisation (dpf) and assessed bacterial burden and host survival daily up to 144 hours post-injection (hpi) (**Figure 1A-C, Figure S1A-C**). Quantification of bacterial burden by CFU enumeration (**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 clearance phase characterised by a significant decrease of bacteria at a constant rate (24-96 hpi), and a persistence phase where few bacteria (i.e., <5% of the initial bacterial load) are maintained over an extended period of time (96-144 hpi).

Treatment of infection with a therapeutic dose of the antibiotic Nalidixic acid prevented host death in response to both low dose (**Figure 1B,C)** and high dose (~8000 CFU, characteristically eliciting ~80% host death by 72 hpi, **Figure S1D,E**) infections. Treatment with Nalidixic acid also inhibited bacterial growth (**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 clearance phase (**Figure 1B, Figure S1A-C**). In the presence of Nalidixic acid, 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 persistence phase had developed antibiotic tolerance (irrespective of previously being treated with an antibiotic), we treated larvae exhibiting acute and persistent *Shigella* infection with Nalidixic acid for 24h and found that treatment significantly decreased bacterial burden (~100 fold) during acute infection (**Figure 1D**). In contrast, the 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 Nalidixic acid 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 1E-F**). 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*, which is characterised by poor host clearance over an extended period, antibiotic tolerance and clonality.

Shigella **O-Antigen variants circulating in the MSM population promote 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 2A-B**, **Figure S2A-B**). Considering that the O-Ag of *S. sonnei* is encoded by the pSS plasmid (i.e., -pSS strains are also O-Ag-deficient), we conclude that the O-Ag (and not the T3SS) is essential for *Shigella* persistence.

More than 40 different serotypes of *Shigella* have been reported, and the O-Ag structure is a major variable in determining these different serotypes[17]. Only three serotypes have been directly associated with carriage in the MSM population to-date[2,18]. To test whether strain clusters associated with carriage in humans establish persistence in zebrafish, we constructed phylogenetic trees for *S. sonnei* and *S. flexneri* using high quality complete genomes deposited in the Bacterial and Viral Bioinformatics Resource Centre (BV-BRC, https://www.bv-brc.org/), sequencing data (short-read assemblies) collected from cases of likely carriage[18], as well as newly collected sequencing data (long-read assemblies) from *S. sonnei* and *S. flexneri* clinical isolates selected for further testing in zebrafish (**Figure S2G-H**). For the construction of trees, we used the Codon Tree method and the Phylogenetic tree building service made available by BV-BRC[19]. 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[18]) 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[20–22], however only one *S. sonnei* serotype exists (**Figure S2G**), which has comparable genetic diversity to individual serotypes of *S. flexneri* (e.g. *S. flexneri* 2a)[23,24]. 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 2C-D**, **Figure S2C-D**). In the case of *S. flexneri*, multiple serotypes have been identified[25], although only 2 serotypes (2a and 3a) have been so far associated with persistent carriage[2] (**Figure S2H**). Strikingly, representative strains of these serotypes also persist in zebrafish (**Figure 2E-F**, **Figure S2E-F**). In contrast, *S. flexneri* M90T (an isolate belonging to the 5a serotype and a persistence outgroup) is unable to persist in zebrafish (**Figure 2E-F**, **Figure S2E-F**). Overall, these results indicate a central role for the O-Ag in 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[2].

Shigella **can establish persistent infection in macrophages** ***in vivo***

Macrophages are widely 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*[8,9,26]. 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 3A-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 over a 12 h period of observation) can be detected as early as 24 hpi (**Figure 3C,Figure S3C**, **Video S2**) and lasted for several consecutive days (**Figure S3A-B**). These data were surprising, considering that macrophage cell death is widely recognised as a hallmark of *Shigella* infection[12,27,28].

Previous work indicated that *S. sonnei* may be less efficient in inducing macrophage death than *S. flexneri* M90T (5a) *in vitro* using THP1 cells[27,29]. 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[29,30]) 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 3D**). Infection with *S. flexneri* 5a M90T led to a significantly higher level of caspase-1 activity compared to infection with *S. sonnei* 53G, supporting the conclusion that persistent *Shigella* serotypes fail to induce macrophage death *in vivo*.

Conclusion

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[31–34]. 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. Our results show that O-Ag variants promote 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[23] 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*, strongly suggesting that macrophages represent a preferred niche for *Shigella* persistence. Strikingly, *S. sonnei* (which is highly associated with persistence) induces significantly less macrophage death *in vivo*, as compared to the *S. flexneri* serotype 5a (which has not been 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.

STAR Methods

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should contact Serge Mostowy (serge.mostowy@lshtm.ac.uk).

Materials availability

Shigella GM strains generated in this study are available upon request.

Data and code availability
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**.

EXPERIMENTAL MODEL AND SUBJECT DETAIL

Zebrafish model
Animal experiments were approved by the Home Office (project licenses Project license: PPL P4E664E3C) and performed following the Animals (Scientific Procedures) Act 1986. Wild-type AB zebrafish were used for all survival assays and CFU quantification experiments. For experiments involving imaging of zebrafish macrophages, these cells were labelled by crossing lines carrying the following the *Tg(mpeg1::Gal4-FF)*^{gl25} and *Tg(UAS:LIFEACT-GFP)*^{mu271} alleles. Eggs were obtained by naturally spawning, and both control and infected larvae were maintained at 28.5°C in embryo medium (0.5x E2 medium supplemented with 0.5 parts per million (ppm) of methylene blue. For injections and live microscopy, larvae were anaesthetized with 200 µg/ml tricaine (Sigma-Aldrich) in embryo medium. Injections were performed in larvae at 2 or 3 days post-fertilisation (dpf) and infected larvae were monitored up to 6 days post-infection. Sex was not determined, as all experiments were concluded before the stage of zebrafish sexual development.

METHODS DETAILS

Bacterial strains
Wild-type and genetically modified bacterial strains used in this work are further detailed in **Table S1 and Table S2**, respectively. 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> 8,000 µg/ml) and could not be transformed. Bacterial mutant strains were previously described in Torraca et al., 2019[33].

Zebrafish infection
Bacteria were cultured at 37°C overnight in Trypticase Soy Broth (TSB) supplemented (when appropriate) with 100 ug/ml Carbenicillin, diluted 50x in fresh medium, and grown to Log phase. For inoculum preparation, bacteria were then spun down, washed in 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 in the figure legend, 1 nl (corresponding to approximately 1000 CFU) of bacterial suspension was microinjected in the hindbrain ventricle (HBV) of 3 dpf zebrafish larvae. Exceptions include data in **Figure 3C**, **Figure S3C**, **Video S2** and **Figure 3D**. In these cases, 1000 CFU (**Figure 2C**, **Figure S3C**, **Video S2**) were delivered systemically in 2 dpf larvae or 10000 CFU (**Figure 3D**) were delivered systemically in 3 dpf larvae via the caudal vein.

Survival assays and bacterial burden
Larvae failing to produce a heartbeat during a 30-second observation period were considered nonviable. For larvae beyond 5 dpf, clinical scoring criteria were applied to

identify humane endpoints. Briefly, larvae were discontinued when the bacterial load (as assessed by live microscopy) was predictive of death in the following 24 hours (i.e., the hindbrain was filled with bacteria and/or infection was disseminated throughout the body) or when larvae resulted irresponsive to touch. These larvae were withdrawn from the experiment, euthanised and considered to be nonviable at the following timepoint.

For enumeration of live bacteria by CFU plating, larvae were washed in PBS, anesthetised, and mechanically homogenised in 200 µl (non-persistent stages, i.e., < 72 hpi) or 40 µl (persistent stages, i.e., ≥72 hpi) of PBS. Homogenates were serially diluted and plated onto Trypticase Soy Agar (TSA) containing 0.01% Congo red. Only viable larvae were used for CFU analysis.

Clonality assay
Larvae were infected as previously specified, but with an inoculum composed of GFP-labelled *S. sonnei* and mCherry-labelled *S. sonnei* in a 1:1 ratio. The ratio between the two strains was determined by CFU plating at 0 hpi and 144 hpi. At 0 hpi, appropriate dilutions of the samples were made before plating, to correct for any differences in the occurrence of stochastic bottlenecks, due to different sample sizes.

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 ImageJ/FIJI software.

Drug treatments
Nalidixic acid 1 µg/ml was directly added to the embryo medium. In pilot experiments, a range of concentrations were tested (0.5-4 µg/ml) and 1 µg/ml was the lowest concentration tested that fully prevented bacterial growth and death of zebrafish 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 10000 CFU of mCherry-labelled *S. sonnei* 53G, *S. flexneri* M90T or PBS containing mock solution. At 4 hpi larvae were dissociated by treatment with 1 ml of 4% trypsin for 15 minutes at 28ºC. Single-cell dissociation was facilitated by mechanical disruption using a P1000 pipette after trypsin treatment. Dissociated cells were harvested by centrifugation (5 min, 800xg at room temperature), washed in calcium-free PBS, separated by passage on a 4 µm cell strainer and suspended in 500 µl staining solution for active caspase-1 (FLICA® 660 Caspase-1 assay, probe: 660-YVAD-FMK, Immunochemistry technologies, #9122) prepared as per the manufacturer's guidelines. Upon staining, cells were pelleted, washed in PBS and finally fixed in 4% paraformaldehyde overnight. For flow cytometry, cells were washed with 2 ml PBS twice and resuspended in 300 µl PBS. Active caspase-1 staining of single macrophages was measured on a LSRII (BD Biosciences) and data was analysed with the software FlowJo version 10.7.1.

DNA extraction and sequencing

Genomic DNA was extracted from overnight bacterial cultures using the MasterPure Complete DNA Purification Kit (Lucigen Corporations, Wisconsin, US) according to manufacturer's instructions. DNA concentration, purity and quantity was assessed using Nanodrop spectrophotometry (DeNovix, Thermofisher, UK) and Qubit fluorometer (Invitrogen, US) according to manufacturers' instructions. Sequencing libraries were prepared using the ligation sequencing kit- SQK-LSK109 kit according to manufacturer's instructions (Oxford Nanopore Technologies, Ltd, Oxford, UK) with slight modifications to the input DNA amounts (i.e., the 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 version R9.4.1 (Oxford Nanopore Technologies Ltd, Oxford, UK).

Assembly and annotation

The fast5 read files generated from the MinION instrument were base-called and demultiplexed with Guppy v5.0 (Oxford Nanopore Technologies Ltd. Oxford, UK). Processed read files were filtered using Filtlong v0.2.0, then assembled using Flye v2.9[35]. Racon v1.5.0[36] was used to polish contigs with the 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 then used to polish with Illumina reads, where they were available[37]. The quality and statistics of each assembly was evaluated with QUAST v4.4.0 without a reference genome[38]. Genomes were annotated using the Prokaryotic Genome Annotation Pipeline (PGAP). The complete genome sequence data have been submitted to the National Centre for Biotechnology (NCBI) and have been deposited at GenBank under the BioProject number PRJNA869897.

Phylogenetic tree construction

Annotations of new genomes for the tree construction were performed using the Genome Annotation service made available by the Bacterial and Viral Bioinformatics Resource Centre (BV-BRC, https://www.bv-brc.org/)[19]. For the genome annotation, parameters were set as Annotation recipe: Bacteria/Archaea; Taxonomy name: *Shigella*. Phylogenetic trees were constructed using the Codon Tree method via the Phylogenetic tree building service made available by BV-BRC. For the tree building, parameters were set as Number of genes: 1000; Max allowed deletions: 1; Max allowed duplications: 1. Trees were constructed separately for *S. sonnei* and *S. flexneri* isolates. Data included the complete genomes of *S. sonnei* and *S. flexneri* isolates of known serotypes (publicly available via the BV-BRC database as for November 2022), as well as the draft genomes of all the isolates sequenced/resequenced and tested in this study (and reported in **Table S1**). The trees also include the draft genomes for isolates that had previously been reported to represent likely cases of carriage in humans with a distance between the collected isolates of at least 10 days[18]. Since this dataset represented pairs of isolates with <8 SNP distance, only sequences from the second isolates were used for tree construction. Data were visualised and annotated using FigTree v1.4.4. (http://tree.bio.ed.ac.uk/software/figtree/) and microreact (https://microreact.org/).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical tests were performed using Prism software (GraphPad Software, Inc.) or Microsoft Excel. The statistical significance of survival curves was determined using the log-rank Mantel-Cox test. In all other cases, statistical significance was determined

using an unpaired two-tailed t-test or one-way ANOVA with Sidak correction, as further specified in the figure legends. Analyses were performed on Log10-transformed valued for CFU counts. Data are represented as mean ± standard errors of the mean (SEM).

KEY RESOURCES TABLE

Acknowledgements

We thank the Wiebke Herzog lab for the Tg(UAS:LIFEACT-GFP)mu271 zebrafish line and all members of the Mostowy, Baker and Holt labs for discussion and experimental help. V.T. was supported by an LSHTM/Wellcome Institutional Strategic Support Fund (ISSF) Fellowship (204928/Z/16/Z). D.B. is supported by the Deutsche Forschungsgemeinschaft (DFG) Walter Benjamin Programme (BR 6637/1-1). Research in K.B. laboratory is supported by the Academy of Medical Sciences Springboard award (SBF002\1114), a Medical Research Council New Investigator award (MR/R020787/1), and the BBSRC (BB/V009184/1). Work in the K.E.H laboratory is supported by the Bill and Melinda Gates Foundation, Seattle (OPP1175797), KlebNet Project. Research in S.M. laboratory is supported by a European Research Council Consolidator Grant (772853 - ENTRAPMENT), Wellcome Trust Senior Research Fellowship (206444/Z/17/Z) and the Lister Institute of Preventive Medicine. K.B. and C.J. are affiliated with the National Institute for Health Research Health Protection Research Unit (NIHR HPRU) in Gastrointestinal Infections at the University of Liverpool in partnership with the United Kingdom Health Security Agency (UKHSA), in collaboration with the University of Warwick. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, the Department of Health and Social Care or UKHSA.

Author contributions

This project was conceived by V.T. and S.M. V.T. 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. 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 the authors.

FIGURE LEGENDS

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*. 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 hours post-infection (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 the 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). (**E-F**) mCherry and GFP labelled-*S. sonnei* were co-injected at a 1:1 ratio. At 144 hpi one of the two strains could be isolated at a much higher frequency than the other strain, indicating establishment of clonality. 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) $p \geq 0.05$; *$p<0.05$; ***$p<0.001$.

Figure 2. *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 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 (ΔOAg) 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. (**E-F**) Percentage of larvae carrying persistent infection at 144 hpi (E) and average bacterial Log_{10}CFU (F) 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. Statistics: one-way ANOVA with Sidak's correction on percentage data (A,C,E) or Log10-transformed data (B, D, F); ns (non-significant) $p \geq 0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

Figure 3. *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 (with macrophages in green) were injected at 3 dpf in the hindbrain ventricle with 1000 CFU mCherry-labelled *S. sonnei* 53G (red). An individual infected macrophage harbouring persistent bacteria is magnified. Scalebars: 100 µm (left); 10 µm (right, inset). (**B**) 40% of bacterial fluorescence colocalises with *mpeg1*$^{+}$ macrophages at persistent infection stages. (**C**) Individual macrophage harbouring mCherry-*S. sonnei*, followed over time for 12 h (from 24 to 36 hpi). For data in (C), 1000 CFU bacteria were delivered systemically in 2 dpf larvae. Scalebar: 10 µm. (**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, 10000 CFU bacteria were delivered systemically in 3 dpf larvae. Statistics: one-way ANOVA with Sidak's correction; ***$p<0.001$; ****$p<0.0001$.

Supplementary Figure legends

Figure S1 (related to Fig. 1). *Shigella* establishes persistent infection in zebrafish
(**A-C**) Linear regression analysis of the change of Log_{10}(CFU) of *S. sonnei* as a function of time for untreated larvae (A) or larvae treated with the antibiotic Nalidixic Acid (B), and relevant summary statistics (C). In the absence of antibiotics, *Shigella*

infection progresses in three distinct phases (acute, clearing and persistent), that can be described by individual linear equations. Treatment with therapeutic doses of Nalidixic acid prevents infection from undergoing acute expansion, but it does not promote complete eradication of the infection. The presence or absence of the antibiotic does not significantly impact the trend (slope) of Log_{10}(CFU) over time, neither during the clearing phase nor during the persistence phase. (**D-F**) Survival curves (D), Log_{10}(CFU) (E) and phenotypic analysis (F) of zebrafish larvae infected with a lethal dose (i.e. 8000 CFU) of *S. sonnei* and treated with different concentrations of Nalidixic acid. 1 µg/ml Nalidixic acid was sufficient to completely rescue host survival (D) and prevent significant bacterial growth *in vivo* during the acute infection phase (E). Toxicity at this antibiotic concentration is negligible (F) and <10% of treated larvae displayed aberrant phenotypes (i.e., cardiac oedemas). (**G**) *Shigella* persistent infection is associated with antibiotic tolerance. At 72 hpi (i.e., end of the clearing phase) only 0.53% of the tested larvae carried bacteria that developed an inheritable decreased susceptibility to the antibiotic treatment (i.e. 4x MIC).

Figure S2 (related to Fig. 2). *Shigella* O-Antigen serotype associated with MSM transmission enables persistent infection

(**A-B**) Log_{10}CFU of individual larvae (A) and survival curves (B), for WT *S. sonnei* 53G and several isogenic mutants. (**C-D**) Log_{10}CFU of individual larvae (C) and survival curves (D) for several *S. sonnei* isolates from Lineage II and Lineage III. (**E-F**) Log_{10}CFU of individual larvae (E) and survival curves (F) for several *S. flexneri* isolates from serotypes 2a, 3a and 5a. (**G-H**) Phylogenetic trees (codon trees) for *S. sonnei* (G) and *S. flexneri* (H) 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 (available from Bengtsson et al.[18]) 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, while black labels indicate strains that did not establish persistent infection in zebrafish. 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. Statistics: one-way ANOVA with Sidak's correction on Log10-transformed data (A, C, E); Log-rank Mantel-Cox test (B, D, F); ns (non-significant) $p \geq 0.05$; *$p<0.05$; **$p<0.01$; ***$p<0.001$; ****$p<0.0001$.

Figure S3 (related to Fig. 3). *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 72 hpi. *Tg(mpeg1::Gal4-FF)*^{gl25}/*Tg(UAS:LIFEACT-GFP)*^{mu271} larvae (with macrophages in green) were injected at 3 dpf in the hindbrain ventricle with 1000 CFU mCherry-labelled *S. sonnei* 53G (red). A cluster of infected macrophages harbouring persistent bacteria is magnified. Scalebars: 100 µm (left); 10 µm (right, inset). (**B**) Longitudinal imaging of the head region from a representative *S. sonnei* infected zebrafish larvae at 0, 24, 72 and 144 hpi. The same region of interest is magnified in the bottom left corner. Scalebar: 100 µm. (**C**) Longitudinal quantification of bacterial fluorescence in an individual macrophage harbouring mCherry-*S. sonnei,* followed over time for 12 h (from 24 to 36 hpi). The bacterial fluorescence remains

overall constant or only slightly increases (i.e., bacterial fluorescence shows only a 1.21 ± 0.21 fold change over a 12 h period of observation). For data in C, 1000 CFU bacteria were delivered systemically in 2 dpf larvae. Scalebar: 10 µm.

Table S1. Bacterial strains used in this study

Table S2. Details of newly sequenced/re-sequenced bacterial strains used in this study

Video S1 (related to Figure 3). Confocal z-stack of a macrophage carrying *S. sonnei* infection at 144 hpi.
Tg(mpeg1::Gal4-FF)^{gl25}/*Tg(UAS:LIFEACT-GFP)*^{mu271} larvae (with macrophages labelled in green) were injected at 3 dpf in the hindbrain ventricle with 1000 CFU mCherry-labelled *S. sonnei* 53G (red). Images were taken by confocal microscopy at 144 hpi.

Video S2 (related to Figure 3). Confocal time lapse of a macrophage carrying *S. sonnei* infection for 12 h.
Tg(mpeg1::Gal4-FF)^{gl25}/*Tg(UAS:LIFEACT-GFP)*^{mu271} larvae (with macrophages labelled in green) were injected at 2 dpf via the caudal vein with 1000 CFU mCherry-labelled *S. sonnei* 53G (red). Images were taken by confocal microscopy at 6.5 minute intervals for 12 h, from 24 to 36 hpi.

Reference List

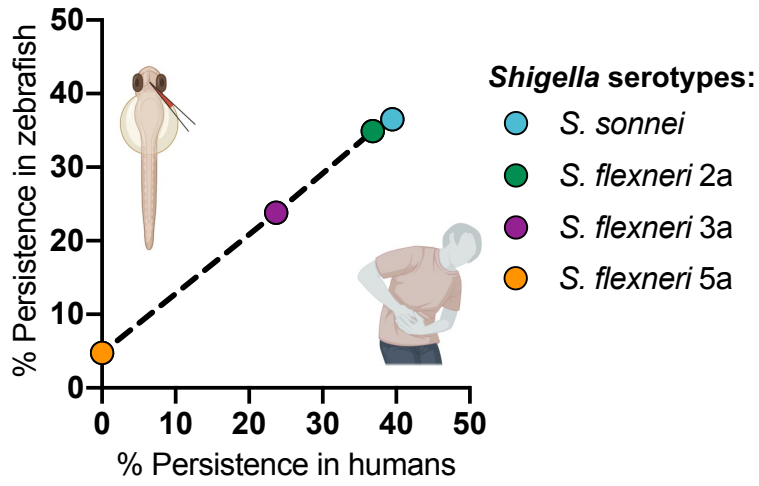
1. Thompson, C.N., Duy, P.T., and Baker, S. (2015). The rising dominance of *Shigella sonnei*: An intercontinental shift in the etiology of bacillary dysentery. PLoS Negl Trop Dis *9*, e0003708. 10.1371/journal.pntd.0003708.
2. Allen, H., Mitchell, H.D., Simms, I., Baker, K.S., Foster, K., Hughes, G., 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. Clinical Microbiology and Infection *27*, 126.e7-126.e13. 10.1016/j.cmi.2020.03.036.
3. 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. (2020). Duration of postdiarrheal enteric pathogen carriage in young children in low-resource settings. Clinical Infectious Diseases *0*, 0. 10.1093/cid/ciaa1528.
4. 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. American Journal of Tropical Medicine and Hygiene *97*, 1872–1875. 10.4269/ajtmh.17-0282.
5. Kotloff, K.L., Winickoff, J.P., Ivanoff, B., Clemens, J.D., Swerdlow, D.L., Sansonetti, P.J., Adak, G.K., and Levine, M.M. (1999). Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. Bull World Health Organ *77*, 651–666.
6. 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.

7. 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.
8. 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. *27*, 54-67.e5. 10.1016/J.CHOM.2019.11.011.
9. 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. Nature Communications 2019 10:1 *10*, 1–12. 10.1038/s41467-019-12224-3.
10. 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.
11. 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.
12. Duggan, G.M., and Mostowy, S. (2018). Use of zebrafish to study *Shigella* infection. Dis Model Mech *11*, dmm032151. 10.1242/dmm.032151.
13. 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.
14. 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.
15. 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.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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*. Nature Communications 2021 12:1 *12*, 1–12. 10.1038/s41467-021-22700-4.
21. 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.
22. Torraca, V., Holt, K., and Mostowy, S. (2020). *Shigella sonnei*. Trends Microbiol *28*, 696–697. 10.1016/j.tim.2020.02.011.
23. 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.
24. 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. Nature Microbiology 2022 7:2 *7*, 251–261. 10.1038/s41564-021-01054-z.
25. 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.
26. 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.
27. 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.
28. 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.
29. 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.
30. 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.
31. 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*, e1006467. 10.1371/journal.ppat.1006467.
32. 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.
33. 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.
34. 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.
35. Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P.A. (2019). Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology 2019 37:5 *37*, 540–546. 10.1038/s41587-019-0072-8.
36. 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.
37. 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.
38. 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.

Graphical Abstract



Graphical abstract. Zebrafish can be used to predict the likelihood of persistence of different serotypes of *Shigella*.
Values for the % Persistence in humans were calculated from data extracted from Bengtsson et al[18]. For each serotype/species the % of persistent cases over the total of persistent cases was calculated. Values for the % Persistence in zebrafish were extracted from Fig. 2C and 2E. For each serotype/species, the % of persistent cases in zebrafish over the total of persistent cases was calculated.

Figure 1. *Shigella* establishes persistent infection in zebrafish

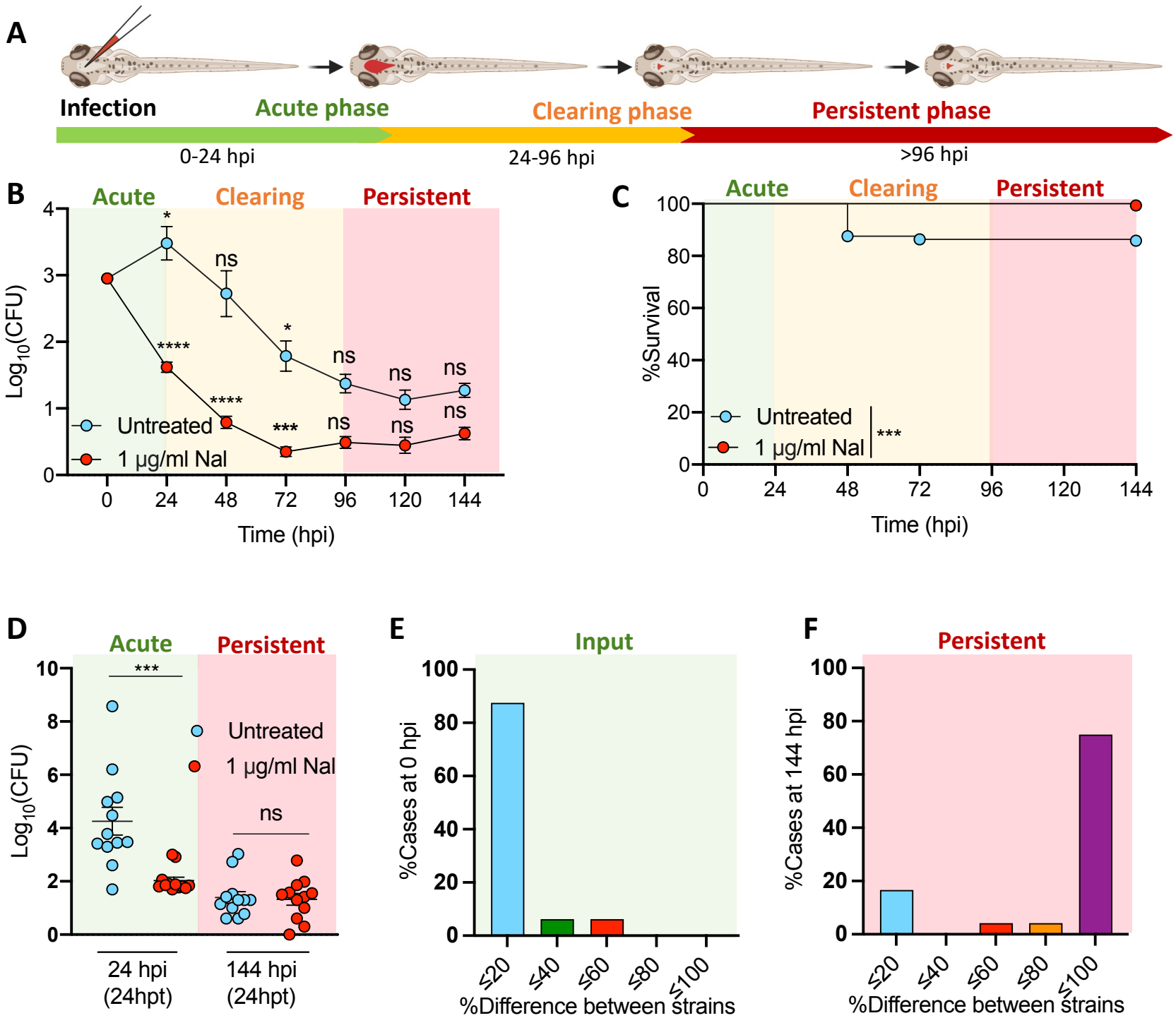


Figure 2. *Shigella* O-Antigen serotypes associated with MSM transmission enable persistent infection

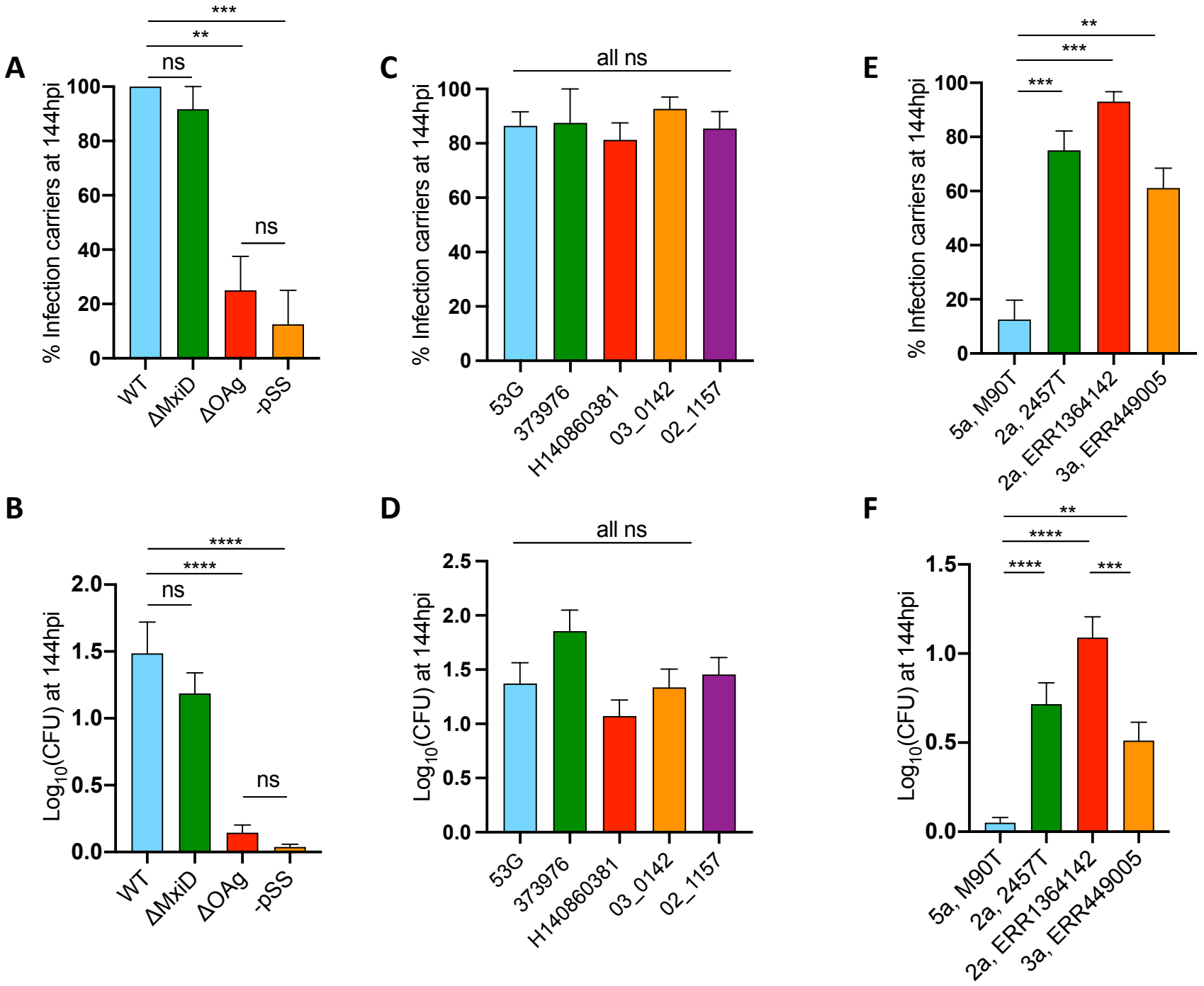
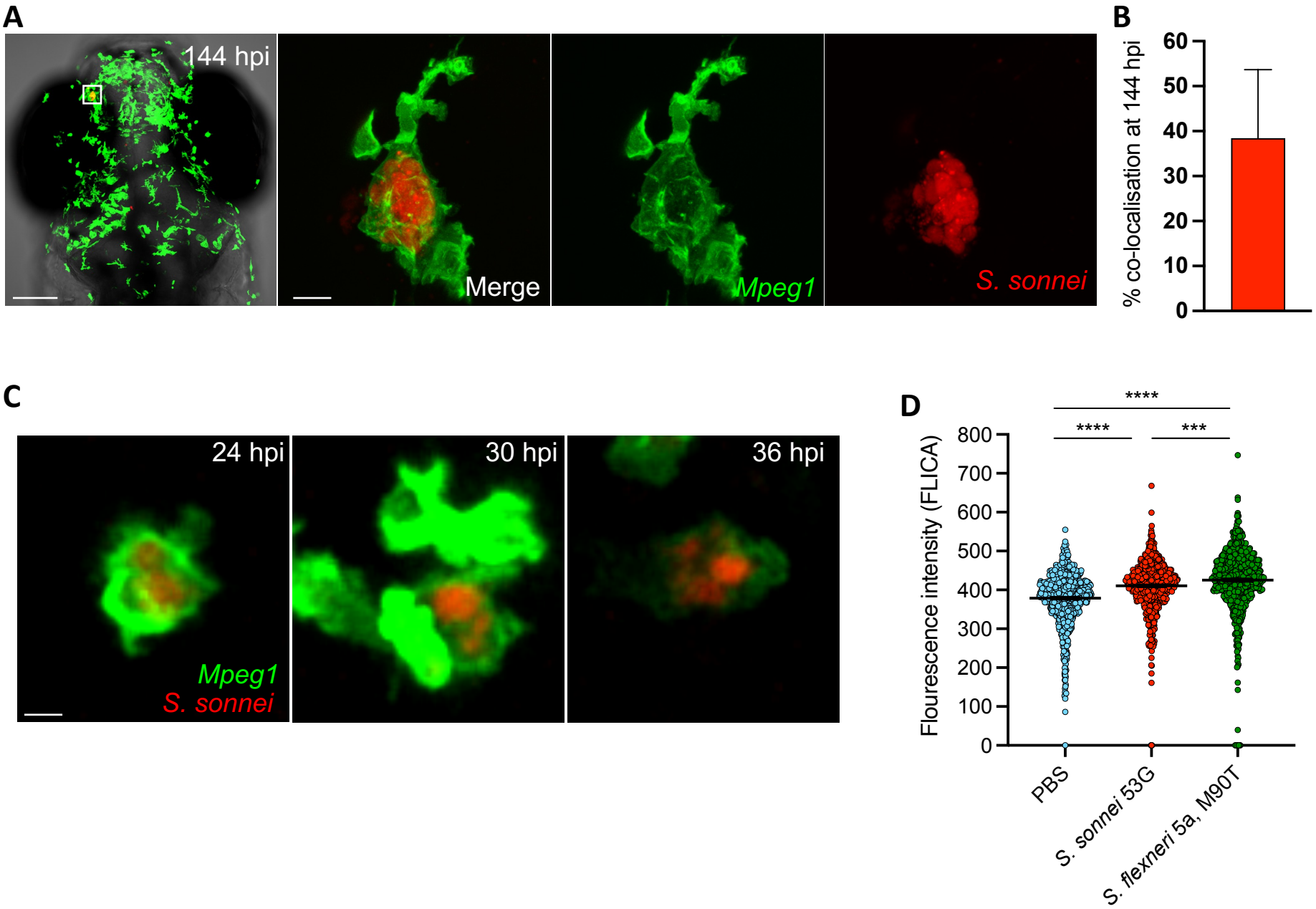



Figure 3. *Shigella* can establish persistent infection of macrophages *in vivo*