

RESEARCH ARTICLE

Reversible excision of the *wzy* locus in *Salmonella* Typhimurium may aid recovery following phage predation

Oliver JD Charity^{1,2✉†}, Gaetan Thilliez^{1,3✉}, Haider Al-Khanaq¹, Luke Acton^{1,2}, Rafał Kolenda^{1,4}, Matt Bawn^{1,5,6}, Liljana Petrovska^{7a}, Robert A. Kingsley^{1,2*}

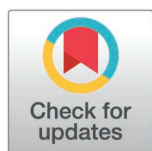
1 Quadram Institute Bioscience, Norwich, United Kingdom, **2** University of East Anglia, Norwich Research Park, Norwich, United Kingdom, **3** School of Biotechnology, Dublin City University, Dublin, Ireland, **4** Department of Biochemistry and Molecular Biology, Faculty of Veterinary Medicine, Wrocław University of Environmental and Life Sciences, Wrocław, Poland, **5** Earlham Institute, Norwich, United Kingdom, **6** School of Natural and Environmental Sciences, Newcastle University, Newcastle, United Kingdom, **7** Animal & Plant Health Agency (APHA), Weybridge, London, United Kingdom

† Deceased

✉ These authors contributed equally

✉ Current address: Gastrointestinal Infections and Food Safety (One Health) Division, Clinical and Public Health Group, United Kingdom Health Security Agency, London, United Kingdom

* rob.kingsley@quadram.ac.uk



OPEN ACCESS

Citation: Charity OJ, Thilliez G, Al-Khanaq H, Acton L, Kolenda R, Bawn M, et al. (2025) Reversible excision of the *wzy* locus in *Salmonella* Typhimurium may aid recovery following phage predation. PLoS Genet 21(5): e1011688. <https://doi.org/10.1371/journal.pgen.1011688>

Editor: Xavier Didelot, University of Warwick, UNITED KINGDOM OF GREAT BRITAIN AND NORTHERN IRELAND

Received: September 19, 2024

Accepted: April 11, 2025

Published: May 2, 2025

Copyright: © 2025 Charity et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: Whole genome sequence used in the analysis described in this manuscript are deposited in freely available in the Short Read Archive (SRA) of the National Center for Biotechnology Information (NCBI) in

Abstract

Bacteriophage (phage) are promising novel antimicrobials but a key challenge to their effective implementation is the rapid emergence of phage resistance. An improved understanding of phage-host interactions is therefore needed. The Anderson phage typing scheme differentiates closely related strains of *Salmonella enterica* serovar Typhimurium (*S. Typhimurium*) based on sensitivity to a panel of phage preparations. Switches in phage type are indicative of changes in phage sensitivity and inform on the dynamics of phage interaction with their host bacteria. We investigated the molecular basis of switches between the relatively phage sensitive *S. Typhimurium* DT8 and phage resistant DT30 strains that are present in the same phylogenetic clade. DT30 strains emerged from DT8 strains predominantly by deletion of a genomic region affecting the *wzy* locus encoding an O-antigen polymerase. The deletion site was flanked by two perfect direct repeats designated attL and attR. During broth culture in the presence of a typing phage that used O-antigen as primary receptor the Δwzy genotype increased in frequency compared with culture in the absence of phage and removal of attL prevented deletion of the *wzy* locus. Co-culture of *S. Typhimurium* DT8 with a strain lacking *wzy* resulted in reversion of the latter to wild type. We propose a model in which reversible deletion of the *wzy* locus enables recovery of *S. Typhimurium* DT8 following predation by phage that use O-antigen as their primary receptor. This was consistent with ancestral state reconstruction of DT8 and DT30 phylogeny that supported a model of reversible transition from DT8 to DT30 in natural populations.

Bioproject PRJNA514245 and PRJNA1227906 and as detailed in [S1 Table](#). All oligonucleotide primers used for genotyping or construction of recombinant bacterial strains are described in [S2 Table](#). Numerical data that underlies graphs is in S4 file.

Funding: RAK laboratory was supported by research grants BB/N007964/1 and BB/M025489/1, and by the BBSRC Institute Strategic Programme Microbes in the Food Chain BB/R012504/1 and its constituent project(s) BBS/E/F/000PR10348 and BBS/E/F/000PR10349 and the BBSRC Institute Strategic Programme Microbes and Food Safety BB/X011011/1 and its constituent projects, BBS/E/F/000PR13635 and BBS/E/F/000PR13636 (all of the above to RAK). OC was supported by a BBSRC DTP studentship (BB/M011216/1). LP was supported by Defra project RDOZ0347. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Salmonella Typhimurium is a major pathogen of livestock that adversely affects productivity and animal welfare and poses a risk of foodborne disease in the human population. Antibiotics are used to control *Salmonella* infections in livestock that contributes to the antimicrobial resistance global emergency. Viruses of bacteria (phage) are one alternative to antibiotics to control *Salmonella* in the food chain, but their successful implementation as antimicrobials is restricted by the rapid emergence of resistance to phage. A better understanding of the outcome of phage-bacteria interactions is needed to optimise the design and implementation of phage-based antimicrobials. This study identifies a genetic mechanism that confers resistance to phage that use O-antigen as a receptor on the surface of *Salmonella*. Resistance is likely to impart a fitness cost on the bacterium but importantly the mechanism has the potential to be revert to a fully fit state once phage predation ceases. A model for how the mechanism may contribute to survival and recovery following phage predation is proposed. This is important since it suggests that some *Salmonella* may be able to recover from phage predation and retain virulence.

Introduction

The bacterial family Enterobacteriaceae include important bacterial pathogens of animals and humans including *Salmonella enterica* [1,2] that are also a primary concern in the emergence of antimicrobial resistance (AMR). The extensive use of antibiotics in agriculture is thought to be a key factor in the emergence of AMR [3]. Phages are viruses that specifically infect and kill bacteria and as such have potential as alternatives or adjuncts to antibiotics to treat infections, prevent transmission between hosts and reduce persistence in the environment [4–11]. A key barrier to the implementation of phages as antimicrobials is that bacteria that are initially sensitive to phage predation readily become resistant through spontaneous mutations affecting the expression of the bacterial cell-surface receptor used by phages to initiate infection. In our previous study, emergence of resistance within 18 hours of inoculation ranged from 10–100% of *S. enterica* serotype Typhimurium (*S. Typhimurium*) co-cultured with the single phage that were due to mutations that effected expression of lipopolysaccharide (O-antigen) on the cell surface [12].

Bacteria and phages exist in a continuous co-evolutionary arms race in which phage predation drives the evolution of antiphage systems (reviewed in [13,14]) and phages evolve counter-measures to evade these systems [15–19]. Phages require receptors which are present on the cell surface of bacteria for successful attachment and entry to continue the lytic cycle or undergo lysogeny. Sensitivity to phage predation is consequently heavily influenced by the presence or absence of these receptors [20]. In laboratory experiments primarily investigating *Escherichia coli* interactions with various model phages, resistance rapidly emerged largely due to mutations affecting expression of the primary receptor [21]. However, prolonged co-evolution has been

observed, for example in batch cultures of *Pseudomonas fluorescens* involving multiple reciprocal selective sweeps of host and phage [22–24]. Frequently phages target surface receptors such as lipopolysaccharide molecules [25,26], whose mutation in many cases is accompanied by a high pleiotropic cost to fitness, limiting further co-evolution [27].

Evidence for the co-evolution and outcome of phage interactions with bacteria in nature is challenging and have therefore necessarily been largely limited to observational or correlational studies. This is in part due to the lack of large datasets linking phage sensitivity with bacterial genotypic diversity [24]. A unique opportunity to investigate the dynamics of host genotype and phage sensitivity arose from the availability of a combination of phage type data and whole genome sequence generated during surveillance of *S. Typhimurium* in the United Kingdom [28]. Prior to the year 2014 the Anderson phage typing scheme was used to differentiate strains (phage types) of *S. Typhimurium* based on their sensitivity to 37 different phage preparations [29,30]. These phage types are stratified into more than 300 patterns of sensitivity which include 209 definitive types (DTs: DT1 - DT209), and over 118 undefined types (Us: U210 – U327) that have patterns not included in the original scheme. Whole genome sequencing (WGS) for surveillance of *S. Typhimurium* is now routine [31] and in the years 2014 and 2015 both phage typing and whole genome sequencing was performed on over 1700 *S. Typhimurium* isolates for surveillance and outbreak detection [28]. The *Salmonella* Typhimurium typing phages (STMPs) group in three main clusters based on their sequence similarity. Most of the phage are P22-like phage and likely use O-antigen as primary receptor, STMP8 and 18 are ES18-like phage that likely target FhuA, while STMP12 and 13 are SETP3-like with an unknown receptor [32].

Phylogenomic analysis of *S. Typhimurium* sequence data revealed that major lineages were characterised by distinct patterns of sensitivity to phage indicated by the prevalence of specific phage types [28]. An observation that was also true for enterohemorrhagic *E. coli* where phage type correlated with bacterial lineages [33]. Furthermore, multiple lineages of *S. Typhimurium* containing multidrug resistant (MDR) strains that emerged and spread widely in livestock species over the past 50 years exhibited less sensitivity to predation by the panel of typing phages. There was also evidence for evolution to decreased sensitivity to the panel of phage during emergence and clonal expansion of the most recently emerged MDR monophasic (I,4,[5],12:i:-) clone of sequence type 34 (ST34) that emerged around 2005 in Europe [34,35]. The ST34 epidemic strains were predominantly DT193 in the year 2014 but ancestral state reconstruction indicated that the common ancestor of the ST34 epidemic clone was DT120 that had greater sensitivity to the panel of phage. Decreased sensitivity to the phage was associated with the multiple acquisitions of a prophage mTmII that resulted in clonal expansion of progeny perhaps due to the presence of an as yet unidentified antiphage mechanism [28].

Here we investigate the genetic basis of changes in sensitivity to phage predation of *S. Typhimurium*. We study a sub-population of *S. Typhimurium* that are adapted to circulation in populations of duck and associated with occasional epidemics of human salmonellosis in the UK and Ireland [36,37]. *S. Typhimurium* in ducks emerge in epidemic waves [38] and are predominantly DT8 with a lower proportion of DT30 [39]. DT30 is characterised by a decrease in sensitivity to phage predation but the phylogenetic relationship with DT8 has not been reported. DT8 and DT30 are both susceptible to STMP8, an ES18 like phage likely to target FhuA, an outer membrane protein. DT8 is also susceptible to STMP10, 11, 16, 20, 22, 23, 26 29 and 32, all of which are P22-like phages expected to target the O-antigen [32].

We found that DT30 are present sporadically within the *S. Typhimurium* DT8 subclade reported previously [36], significantly associated with precise excision of the *wzy* gene encoding an O-antigen polymerase. Excision was dependent on direct repeats flanking the locus and co-culture of *wzy* and Δwzy variants resulted in reconstitution of the *wzy* locus. We propose a model in which excision of *wzy* is a mechanism of phage resistance upon predation by phage that use O-antigen as receptor to initiate infection that has the potential to lead to recovery to wild type following the cessation of predation.

Materials and methods

Bacterial isolates and whole genome sequencing

163 DT8 and 34 DT30 strains were isolated during surveillance and obtained from the Animal and Plant Health protection Agency (APHA), and Public Health England (PHE) United Kingdom (now UKHSA). Whole genome sequence data for

these isolates was produced by Illumina Mi-Seq. Three DT8 clade long read reference genomes were (S04527-10 DT8, L01157-10 DT8, S03645-11 DT30) generated by Pacific Biosciences SMRT sequencing technology [40]. Information on the source, genome sequence accession and phage type for each strain described in this study are reported in [S1 Table](#).

Phylogenetic analysis

Phylogenetic trees from whole genome sequence data were constructed using SNPs identified in the whole genome sequences by aligning reads using BWA-MEM [41], variant calling with Freebayes [42] and SNP filtering using vcflib/vcftools [43], combined as a pipeline using Snippy v4.3.6 [44]. Maximum-likelihood trees were constructed using a general time-reversible substitution model with gamma correction for amongst-site rate variation with RAxML v8.0.20 with 1000 bootstraps using scripts and parameters ([S1 File](#)) [45].

Bacterial growth condition and recombinant strain construction

Bacterial strains were routinely cultured in Miller Lysogen Broth (LB) containing 10g NaCl, 10g tryptone and 5g yeast extract per litre (Formedium, Swaffham, UK) with shaking at 200 revolutions per minute in atmospheric conditions at 37°C. For culture on solid medium, LB was supplemented with 1.5% agar (Oxoid). Construction of recombinant strains was using the lambda red recombineering with primers GTGTAGGCTGGAGCTGCTTCG and CATATGAATATCCTCCTTAGT to amplify the *cat* gene or *aphII* gene from plasmids pKD13 or pKD14, respectively [46]. Variable non-priming sequence at the 5' end of primers was used to direct allelic replacement in the chromosome ([S1 Table](#)). This approach was used to knock out the *wzy* gene (replaced by *aphII*), insertion of the *cat* gene into the *wzy-thrS* intergenic region, the attL^{wzy} (replaced by *aphII* gene) region and to insert a kanamycin resistance cassette in the neutral intergenic region adjacent to the *iciA* gene in strain L01157-10 [47]. The Lambda red recombinase encoding plasmid, pSIM18 was used for the recombineering step [48]. This plasmid uses a temperature sensitive replicons and promotor to activate recombination machinery proteins *exo*, *beta*, and *gam* when exposed to 42°C, and cure the lambda red containing plasmids at 37°C. Knock in of a *wzy-tet* cassette, and *tet* cassette alone was carried out as described previously with minor modifications [49]. A modified pDOC-GG-*glmS* plasmid with ampicillin selection marker was used as a delivery vector. The plasmid encodes a λ -red recombinase which allows sequence specific homologous recombination. The constructs were designed to insert the cassettes in the *glmS* intergenic region which was previously reported as a neutral locus for *S. Typhimurium* fitness [49]. All the restriction enzyme digestion steps were using BsaI type II restriction enzyme (NEB). The Golden gate assembly [50] of the modified pDOC-GG was using primers in [S1 Table](#). The *tet* cassette was obtained by digestion of the muABbla tet 5' plasmid. The cassettes were inserted in different *S. Typhimurium* strains as described in the text.

Estimation of ancestral phenotype histories

We generated various models, and tested model fit by using the likelihood ratio test (LRT) in R. We compared scaled maximum likelihood generated ancestral probabilities at bifurcating nodes between models generated using an equal rate for transition (Q) or allowing different rates for Q. We also generated models using a Markov-Chain with Monte-Carlo (MCMC) sampling approach using equal rates for Q and another with a predefined distribution of Qs. The likelihood ratio distribution asymptotically converges toward a χ^2 distribution. We determined which model was significant using the LRT and identify where the result sits within a χ^2 probability table, using one degree of freedom for an extra parameter of different rates of Q ([S1 File](#)). The tested models were scaled ML with equal rate (Q=-143.6473) or different rate (Q=-90.9481) using ace from the R package ape [52] and MCMC [51] with equal rate (Q=-140.3482) or different rate (Q=-185.748) using the make.simmap function from phytools [55]. This last model (MCMC with different rate) had the greatest log-likelihood and was therefore accepted as the most likely model.

The ancestral histories across a reconstructed phylogeny of 165 DT8 isolates and 34 DT30 isolates were assessed to determine the probable phenotype of each ancestor at each node in the tree. This was then used to determine a likely

history of changes from the phage sensitive to phage resistant phenotype (DT8 to DT30). The MCMC approach was conducted with discrete character mapping using posterior sampled maps from SIMMAP [52]. One thousand sampled stochastic character maps were constructed after a burn-in period of 1000 iterations for Q , followed by 1,000,000 Markov-Chain steps and sampling for the posterior every 1000 generations using the pre-computed distribution of Q [53]. As a complementary approach, pastML [53] was also used to estimate switches from DT8 to DT30 at each node, using the MPPA (marginal posterior probabilities approximation) model and estimating the character variation from tips. Resulting data was interpreted and viewed using R package phytools [53], iTOL [54], and pastML [55]. Scripts and parameters are available (S1 File).

For the MCMC approach, a permutation test was used to assess if the estimated character state at a node was due to a skewed result from sample bias, with 20 permutations. Tip labels were randomly assigned to the tree through the 'sample' R command and a starting tree constructed using the maximum likelihood-based function ACE software, with a subsequent tree sampled every 100 iterations of MCMC, 100 times, resulting in a distribution of 100 trees with ancestral states for each permutation. To assess whether the permuted data was from the same distribution as that estimated from actual tip data, pairwise, Mann-Whitney-Wilcoxon tests were performed.

Bacterial genome wide association

Genome wide association was conducted to discern potential genetic polymorphisms associated with a trait of phage resistance for 34 DT30 isolates, or susceptibility to 10 phages for 164 DT8 isolates. This was undertaken through DNA of length K (k-mer) based analysis, where k-mers were identified from draft genome assemblies of each of the 196 isolate sequences using frequency based string mining algorithm FSM-lite [56]. The approach adjusts the probability of association of k-mers based on phylogenetic structure. The population structure was estimated using mash and converted to a three dimensional distance matrix [57]. Subsequently, a mixed linear model approaches were used for testing k-mer significance implemented with Sequence Element Enrichment (SEER) using scripts and parameters (S1 File) [58]. This was done initially with no significance filtering of k-mers and then the top 1% of k-mers in the range of p -values determined a likelihood ratio test (LRT) p -value cutoff of 1×10^{-3} . k-mers with a likelihood ratio test (LRT) p -value $< 1 \times 10^{-1}$ were plotted.

Determination of specific sequence by mapping and local assembly

For sequence read mapping to the *wzy* locus to determine read-depth, a genomic region of 6kb encompassing the *wzy* locus and flanking genes was extracted from reference sequence L01157-10 (RefSeq Accession no. GCF_902500305.1), and short-read Illumina WGS data for 196 DT8 complex isolates was mapped to this region using Bowtie2 without filtering secondary mapped reads to ensure that any reads mapping to the region would be included [59]. Bedtools-2.26.0 was used to extract the read depth per nucleotide [60]. These were split into 250bp bins and the mean average of each section used as raw data for heatmap using R package gheatmap of ggtree [61]. The read data was normalised by dividing the raw value for 250bp bins by the average read depth over the chromosome. This enabled visualisation of each *wzy* region in the context of each sequencing run. SNPs were identified in *wzy* regions using snippy v4.3.6 as previously mentioned with DT8 isolates S04527-10 and L01157-10 as references for variant calling [40]. For detection of the *wzy* locus in representative *S. Typhimurium* genomes, ARIBA was used along with a custom database containing the sequence of the *att*, *wzy*, *nucA* and *thrS* loci from the strain L01157-10 whole genome sequence.

Determination of phage sensitivity and phage type

Phage typing was carried out as described in Public Health England's phage typing protocol for *S. Typhimurium* using typing phages. STMP8, 10, 18, 20, 29, and 32 were a kind gift from UKHSA. Briefly, a single colony of the bacteria to be tested was incubated in 4 ml of nutrient broth with static, atmospheric conditions at 37°C for 2 hours. Subsequently a nutrient agar plate was flooded with the culture before drying. 10 μ L of each phage suspension at recommended titre dilution

was spotted onto the plate and incubated for 16 hours. Phage typing was used to assess that strains of DT8 and DT30 were susceptible/ resistant to the corresponding phage. Constructed mutants were also subject to challenge with the typing phage to assess the phenotype. A colony picked from the clearance zone in an experiment in which strain L01157-10 was exposed to phage STMP10 was used in reversion experiments (L01157-10 Δ wzy).

Determination of genotype and wzy circularisation by PCR amplification

Presence of the wzy region was routinely tested using primer pair TGCGACTATCAGGTTACCGT and GTTAGCGTGCGGTCAAGATC that anneal in the *nucA* and *thrS* genes. qPCR was utilized to quantify the wzy⁺ genotype in clonal populations with and without phage selection. Specific amplification of the circularised wzy locus was using primers AAGCCGAGACTCAGAGTGAC and CTCCGCCCTAATCCACATCT, and the same primers were used for the determination of the nucleotide sequence of the amplicon using TubeSeq (Eurofins). The consensus sequence of the PCR product and a circularised version of the wzy region (S2 File) were aligned using Muscle [62].

Quantification of wzy genotype

The relative and absolute quantification of wzy and Δ wzy genotypes was determined by qPCR using an Applied Biosystems StepOnePlus. The strain to be tested was grown for 16–18 hours at 37°C in LB broth with shaking atmospheric conditions. This was sub-cultured in 10 mL of LB broth with 10⁴ cells per mL using a guide of an OD₆₀₀ of 3.5 equalling 10⁹ cells per mL before challenge with 3.34x10⁴ or 3.34x10⁵ pfu/mL with STMP10. Then at 2-hour intervals for 8 hours, and a sample 24 hours post inoculum, 400 μ L of culture was taken and stored at -20°C. Once all samples had been frozen genomes were extracted using Promega Maxwell. L01157-10 was used as a positive control for growth, L01157-10 wzy⁺ for control of growth of a resistant mutant under phage pressure, and a pure lysate with 3.44x10⁵ phage to control for any effect of phage DNA increasing the Ct values. qPCR was undertaken using extracted genomic DNA with *rpoD* as a housekeeping control gene, and wzy was assessed using primer pairs wzy_qPCR_absent_F, wzy_qPCR_absent_R and wzy_qPCR_present_F, wzy_qPCR_present_R. A limit of detection for wzy⁺ was established to the nearest log₁₀–1 in 10 dilution of culture.

Transfer of the wzy locus during co-culture

Co-culture and PCR amplification of the wzy locus from suspected revertant colonies was used to assess the reversion of wzy. A donor strain of *S. Typhimurium* L01157-10 wzy⁺ was constructed by the insertion of a *cat* gene conferring resistance to chloramphenicol into the intergenic region between wzy and attR^{wzy}. A recipient strain of *S. Typhimurium* L01157-10 Δ wzy picked from the otherwise cleared plaque formed by phage STMP32 in an agar overlay assay and determined to have the deletion by analysis of whole genome sequence was further engineered by inserting an *aphII* gene conferring resistance to kanamycin in the 5' intergenic region of *iciA* and subsequent selection of a nalidixic acid resistant variant by culture on LB agar containing 100 mg/l nalidixic acid. Two genetically unlinked selectable markers were used in the recipient to discount the possibility of transfer from the recipient strain to the donor. We reasoned that potential transfer of the *aphII* gene by transduction and selection for nalidixic acid-resistant variants of L01157-10 wzy⁺ are both rare events and therefore the frequency of both events occurring together is extremely unlikely. Donor and recipient strains were each cultured for 18 hours in LB and diluted to OD_{600nm} of 0.05 and 2.5 ml of each mixed and incubated for 24 hours with shaking. The CFU of donor and recipients were enumerated by culture of serial dilutions on LB agar plates supplemented with 0.025 mg/ml chloramphenicol (donor), 0.05 mg/ml kanamycin and 0.03 mg/ml nalidixic acid (recipient) and all three antibiotics to enumerate revertants. The presence of the wzy region was tested using primer pair TGCGACTATCAGGTTACCGT and GTTAGCGTGCGGTCAAGATC that anneal in the *nucA* and *thrS* genes, presence of wzy was tested using primers GCCTGAAGATTTTGCGCAT, TGC-GCTGACTTTGTTTCCTG that both anneal within the wzy gene and presence of the *cat* cassette with ACAAACGGCATGATGAACCT and GCACAAGTTTTATCCGGCCT that both anneal within the *cat* gene. Six revertants were subject to WGS with Illumina HiSeq 2500 to check the genotypes of the mutants and reversion of wzy within the same chromosomal location.

Results

The *S. Typhimurium* DT30 strains are sporadically distributed within the DT8 clade

Isolates typed as DT8 and DT30 are predominantly isolated from ducks and form a distinct clade within the *S. Typhimurium* population structure [28,36,40]. Of 2213 *S. Typhimurium* strains isolated from ducks between 1992 and 2016, 68% (n = 1509) were DT8 and 23% (n = 352) were DT30. To further investigate the phylogenetic relationship of DT8 and DT30 strains we established a convenience collection of 176 whole genome sequences of *S. Typhimurium* isolated between the years 1993 and 2013 during surveillance by the Animal and Plant Health Agency (APHA), in the UK (S1 Table). These whole genome sequences were supplemented with 41 isolates from human infection and two genomes described previously [63]. A maximum likelihood phylogenetic tree based on sequence variation in the core genome of 200 isolates (Fig 1A) indicated that strains of DT30 were distributed widely on the phylogenetic tree of the clade among DT8 with little evidence of clonal expansion.

The *wzy* locus is polymorphic in the DT8-DT30 clade and is flanked by two direct perfect repeats.

To test the hypothesis that genome sequence polymorphisms account for decreased sensitivity to phage predation of DT30 compared to DT8 strains, a genome-wide association study (GWAS) was undertaken to identify k-mers associated with these two phagetypes. 2,617 k-mers that had a significant (likelihood ratio test, LRT $p < 0.1$) association with the DT8 or DT30 trait after adjusting for the phylogenetic signal, mapped to four loci in the DT8 L01157-10 chromosome. A minority of k-mers (n = 602) had an LRT p-value ranging from 0.1 and 0.01 and mapped to *yjbR* (n = 200) encoding an uncharacterised DNA-binding protein encoding gene, *fucO* (n = 201) encoding lactaldehyde reductase or *gss* (n = 201) encoding a bifunctional glutathionylspermidine synthetase/amidase. The majority of k-mers, (n = 2,015) mapped to *wzy* gene which encodes an O-antigen polymerase and had a likelihood ratio of association p-value < 0.0001 (Fig 1B). Quantification of whole genome sequence reads of the *S. Typhimurium* DT8 and DT30 strains that aligned to sequence within and flanking the *wzy* locus of *S. Typhimurium* DT8 strain L01157-10 whole genome sequence indicated a region of zero mapped reads in 14 of 34 DT30 isolates, consistent with a deletion of this region (Fig 1A). Like other *S. Typhimurium* strains, the *wzy* gene was present in the *rfa* locus in the strain L01157-10 whole genome sequence and the *wzy* gene in the *rfa* locus was absent [64]. A further 20 strains that were typed as DT30 had sequence coverage across the *wzy* locus. Of these, three DT30 strains with sequence coverage had mutations resulting in a premature nonsense mutation in *wzy* and likely to abrogate function of the o-antigen polymerase, two SNPs and a small indel. Another strain had a non-synonymous SNP predicted to result in a lysine to asparagine substitution with unknown impact on function (Fig 1A).

The assembled sequence of *S. Typhimurium* DT8 strain L01157-10 had 2816bp between the stop codon of *nucA* and the start codon of *thrS* compared to 259bp in the assembled whole genome sequence of all 14 strains with the deletion at the *wzy* locus. Therefore, the deletion was 2558bp. The genome sequence flanking the *wzy* gene contained two 113bp identical direct repeats that coincided with the deletion boundaries observed in the DT30 strains in which the deletion at the *wzy* locus contained a single copy of the repeat sequence (S1 Fig). The direct repeats flanking the *wzy* gene were therefore designated attL^{wzy} (*nucA* - *wzy* intergenic) and attR^{wzy} (*wzy* - *thrS* intergenic) and attB^{wzy} (Δ *wzy*) due to their resemblance to att repeat sequences that flank prophage in the bacterial chromosome and that mediate excision [65]. The att sequence was present in all 134 *S. Typhimurium* genome short-read sequence of representative strains (S3 Table), described previously [40]. We did not investigate the distribution in other serotypes of *S. enterica*. Together these observations were consistent with a site-specific mechanism for deletion of this region that is likely to be widely present in *S. Typhimurium* (Fig 1C).

Deletion of the *wzy* gene results in resistance to typing phage

To investigate the role of *wzy* in the sensitivity to predation by typing phages, a series of genetically engineered variants of DT8 strain L01157-10 in which the *wzy* gene was either deleted or replaced in an alternative genomic

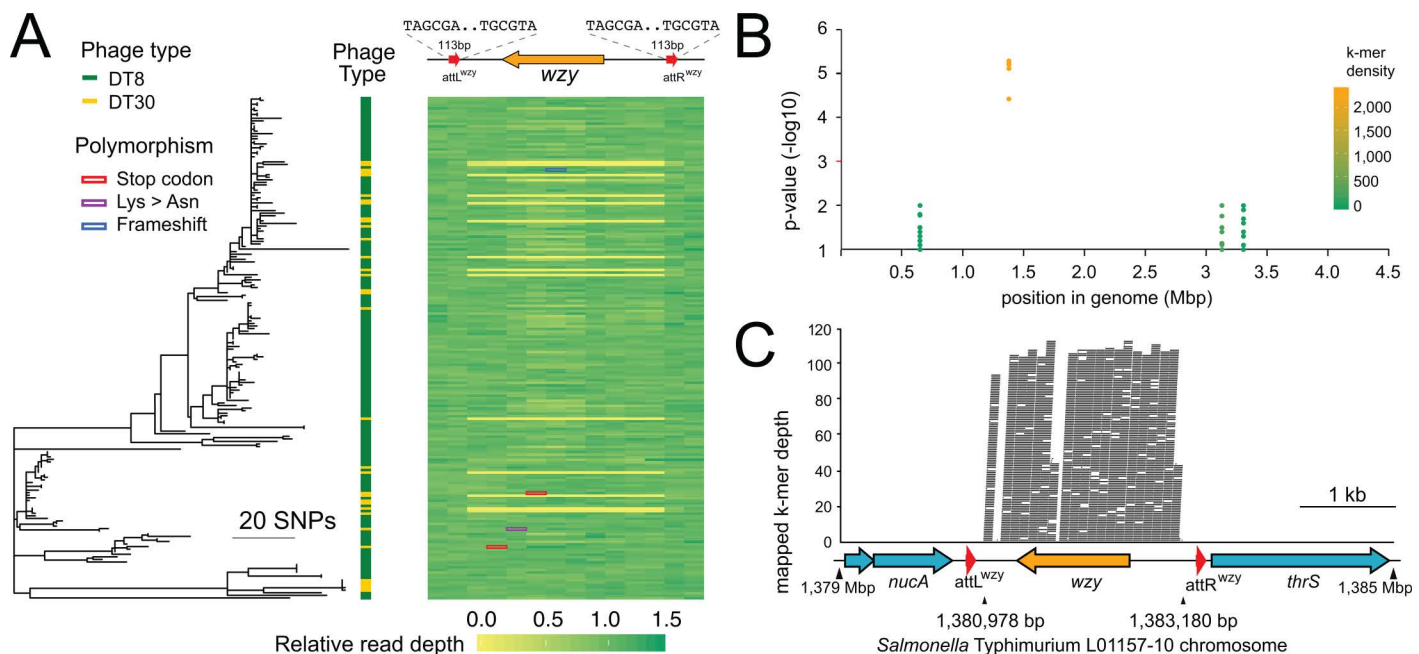


Fig 1. Phylogenetic relationship and association of sequence polymorphisms of *S. Typhimurium* DT8 and DT30 strains. (A) Maximum-likelihood phylogenetic tree constructed using 2,297 core genome variable sites from 162 DT8 strains and 34 DT30 strains. The heatmap indicates the relative level of reads mapped to the *wzy* locus of *S. Typhimurium* DT8 reference strain L01157-10 in 250 bp bins for each strain. Coloured outlines denote the approximate location of polymorphisms that may disrupt Wzy function. (B) Bacterial genome wide association (GWAS) to identify sequence polymorphisms associated with phage resistance phenotypes of DT8 (susceptible) and DT30 (resistant). Manhattan plot of all significant k-mers position and likelihood ratio test (LRT) probability values of significant k-mers mapped to in *S. Typhimurium* DT8 reference strain L01157-10 chromosomal. The density of k-mers at a particular genomic coordinate is indicated by colour (inset key). (C) Location and depth of k-mers mapping to the *wzy* locus. Horizontal lines are k-mers with significant association with phage sensitivity defining DT8 and DT30.

<https://doi.org/10.1371/journal.pgen.1011688.g001>

location for complementation were constructed. Strain L01157-10 $\Delta wzy::aphII$ in which *wzy* was deleted and replaced by the *aph* gene was resistant to phage STMP32 tested DT8-lysing phage (Fig 2). Reintroduction of the *wzy* gene into the 3' intergenic region of the *glmS* gene, linked to a *tet* gene to aid selection, resulted in return of sensitivity to phage STMP32 while the *tet* gene alone did not result in restored sensitivity (Fig 2). The phage sensitivity phenotype of the L01157-10 $\Delta wzy::aph$ strain was therefore the same as DT30 strains. However, it is known that mutation other than those in *wzy* can result in changes in LPS expression that can also affect sensitivity to phage predation. We therefore tested if insertion of the *wzy* gene linked to a *tet* gene could also complement the phage sensitivity phenotype of strain S03645-11 DT30 that was resistant to typing phage STMP32 (Fig 2). Consistent with mutation of *wzy* in this strain being responsible for the loss of sensitivity to typing phage and the DT30 phenotype strain S03645-11 *wzy*, *tet* was sensitive to the typing phage. But introduction of *tet* alone resulted in no change to sensitivity (Fig 2). Since a large proportion of DT30 strains had this specific deletion of the *wzy* gene and flanking sequence, we therefore concluded that this was a common mechanism for the switch from DT8 to DT30 phage sensitivity phenotype.

The Δwzy genotype variants are present in cultures of *S. Typhimurium* DT8 that increase in frequency during phage predation.

Since the *wzy* locus was flanked by attL^{wzy} and attR^{wzy} direct repeats was consistent with a site-specific mechanism of deletion we tested the hypothesis that excision of the locus occurs spontaneously during culture by PCR amplification

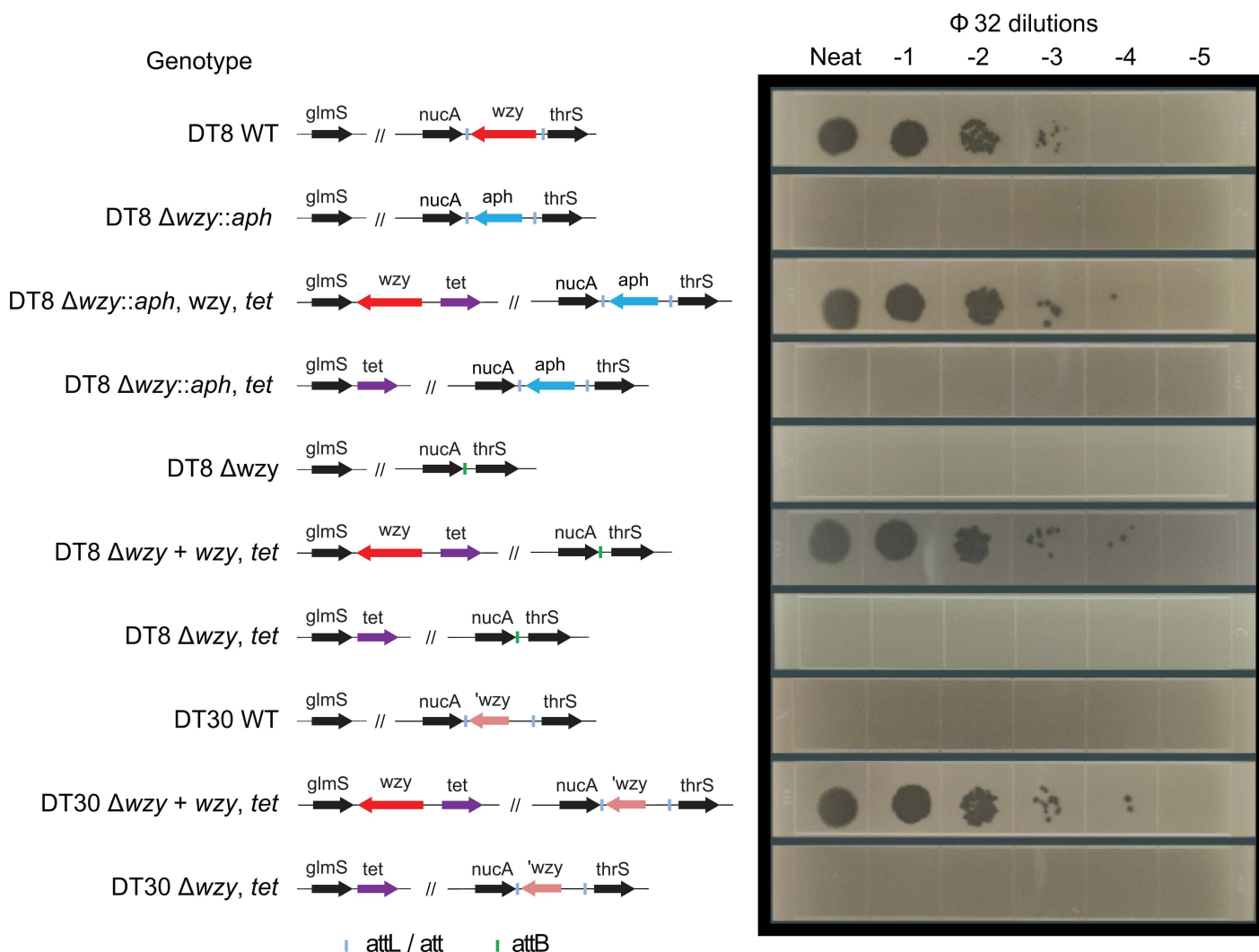


Fig 2. Sensitivity of *S. Typhimurium* DT8, DT30 and recombinant derivative strains to typing phage 32. The genotype of each strain is indicated with shaded arrows indicating genes as labelled (right) and images of serial dilutions of STMP32 overlaid on agar containing each strain.

<https://doi.org/10.1371/journal.pgen.1011688.g002>

across the *wzy* locus. PCR of genomic DNA prepared from stationary phase cultures of two DT8 strains (L01157-10 and S04527-10) using oligonucleotide primers that flanked the *wzy* locus including the attL^{wzy} and attR^{wzy} repeats, resulted in two amplicons of 2.8 kb and 150 bp consistent with a mixture of the wild type *wzy* locus and the deletion variant. In contrast, a naturally occurring variant of L01157-10 in which the *wzy* locus had been deleted, picked from resistant bacteria in the zone of clearance due to lysis by phage STMP10 placed over a lawn of L01157-10 (L01157-10 Δwzy), resulted in only the 150 bp amplicon consistent with deletion of the *wzy* locus (Fig 3A). Using qPCR with the same oligonucleotide primers we estimated that in stationary phase LB broth culture of L01157-10, the ratio of *wzy*: Δwzy was approximately 500:1. To investigate the impact of phage predation on the ratio of *wzy*: Δwzy , we determined the change in frequency of each genotype during culture of strain L01157-10 in the presence and absence of phage STMP10. In the absence of phage STMP10, the frequency of each genotype was stable during 24 hours of culture in LB (Fig 3B). In the presence of phage STMP10, the frequency of the wild type *wzy* locus decreased and the frequency

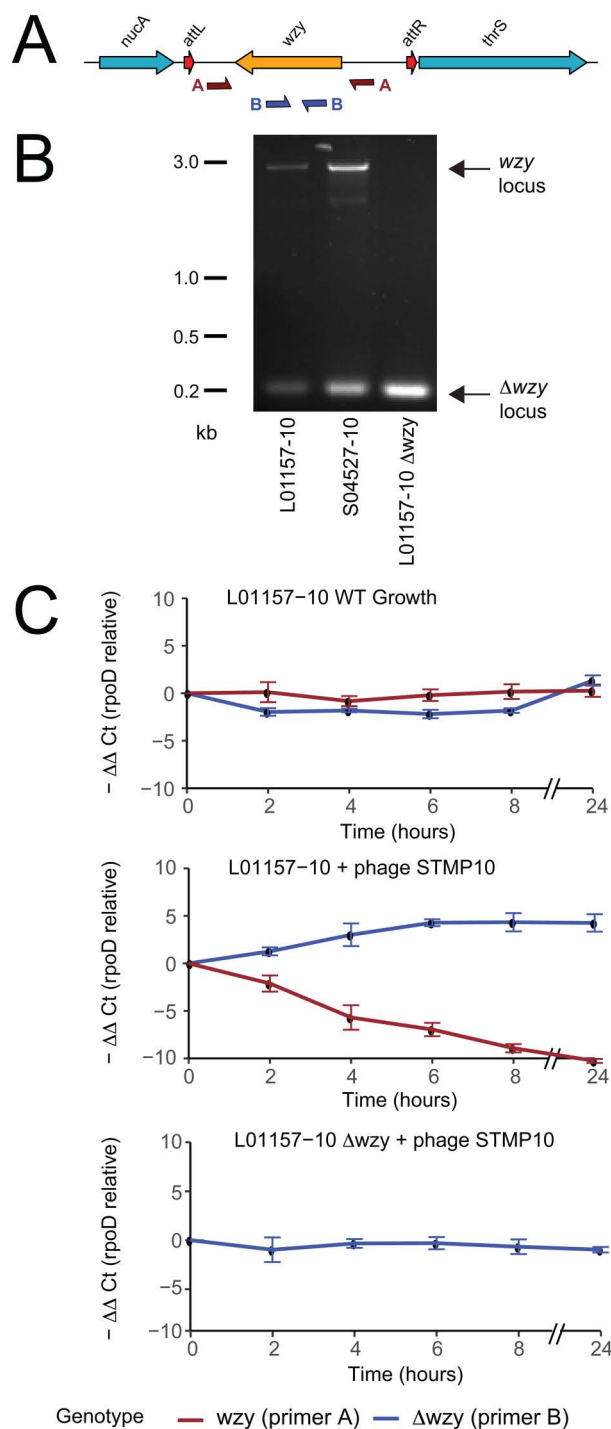


Fig 3. Determination of *wzy* locus genotype during culture and selection in the presence of phage predation. (A) Genetic map of the *wzy* locus indicating the position of oligonucleotide primers used for genotyping using PCR and qPCR. Genes are drawn to scale but arrows indicating the position of primers is not drawn to scale. (B) PCR amplification of the *wzy* locus of two *S. Typhimurium* DT8 strains L01157-10 and S04527-10, and a spontaneous phage resistant variant of L01157-10 using primer pair A with a long extension time to detect Δwzy and wild type *wzy* genotypes. (C) Quantitative PCR of *wzy* genotypes of *S. Typhimurium* DT8 with and without pressure from STMP10 using primer pair A with a short extension time to detect Δwzy genotype and primer pair B to detect the wild type *wzy* genotype.

<https://doi.org/10.1371/journal.pgen.1011688.g003>

of the Δwzy genotype increased, consistent with selection of the deletion variant in response to phage predation. The genotype of a strain in which *wzy* was deleted and replaced by the *aph* gene remained constant in the presence of phage STMP10. Together these data suggest that the *wzy* locus is spontaneously excised during culture and the deletion variant is maintained at a constant frequency. Further, upon predation by a phage that uses O-antigen as a receptor, the frequency of the deletion variant increases relative to the wild type resulting in a change in ratio of *wzy*: Δwzy to approximately 30:1. These data are consistent with deletion being a mechanism of evasion of predation by phage that use O-antigen as a receptor.

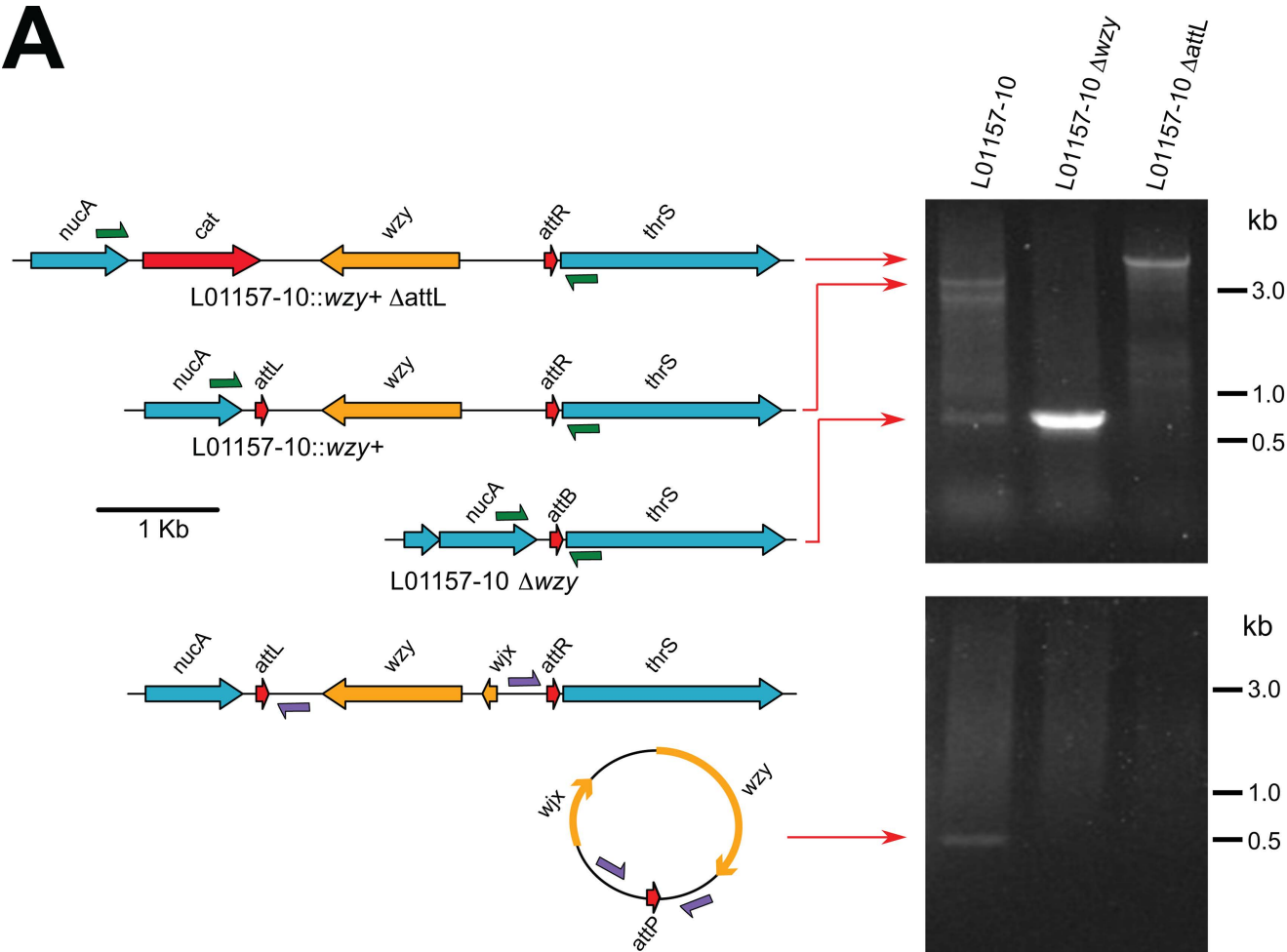
Deletion of the *wzy* locus is through an attL-dependent excision mechanism

To investigate whether the direct repeats attL^{wzy} and attR^{wzy} were required for excision of *wzy* locus, strain L01157-10 recombinants in which either attL^{wzy} or attR^{wzy} were replaced by the *aphII* gene were constructed by allelic replacement. An L01157-10 DattR^{wzy} strain had a growth defect during culture in LB as indicated by small colonies on LB agar, potentially due to the deletion affecting the *thrS* promoter region and therefore was not investigated further. However, an L01157-10 DattL^{wzy} strain had a similar growth characteristic as the wild-type parent strain. PCR amplification across the *wzy* locus of strain L01157-10 resulted in a 2.8 kb band and a 150 bp amplicons but only the larger amplicon in L01157-10 DattL^{wzy} (Fig 4A). This was consistent with attL^{wzy} being essential for deletion of the *wzy* locus. Prophage excision that is also dependent on attL and attR repeats results in the formation of a circular molecular of the phage genome [65]. To investigate if the *wzy* locus also formed a circular molecule, we PCR amplified genomic DNA prepared from strain L01157-10 using oligonucleotide primers that annealed within the excised region of the *wzy* locus in an outward facing orientation. This reaction resulted in a 0.5 kb fragment consistent with a circularisation of the *wzy* locus. The nucleotide sequence of the amplicon aligned to the sequence that would be expected following att-mediated circularisation (S2 and S3 Files). We did not detect the 0.5 kb amplicon in L01157-10 Δwzy . Taken together, the observation suggested that loss of *wzy* is driven by excision of the region using the flanking attL repeat and likely the attR repeat and although a circular form is present long enough to be detected by PCR amplification it is was not retained in a strain in which the *wzy* locus had been excised in the past (Fig 4B).

Δwzy deletion reversion to *wzy*⁺ during co-culture with a wild-type strain.

For *Salmonella* Typhimurium to successfully evade predation by deleting *wzy* without suffering a fitness cost during its lifecycle, the *wzy* gene must be restored once predation ceases. We therefore investigated if Δwzy can revert to wild type *wzy*⁺ during co-culture of a Δwzy recipient strain with a donor strain encoding a wild type *wzy* gene. A donor L01157-10 strain was constructed in which the *cat* gene conferring resistance to chloramphenicol was inserted between the 5' end of *wzy* and attR^{wzy} (Fig 5A, L01157-10 *wzy*, *cat*) so that transfer of the *wzy* locus could be selected by culture on chloramphenicol. A recipient L01157-10 Δwzy strain that was spontaneously resistant to nalidixic acid due to a non-synonymous (D87Y) in the *gyrA* gene, and with an *aphII* gene conferring resistance to kanamycin inserted in the 5' region of the *iciA* gene by allelic exchange was used/constructed (L01157-10 Δwzy , *aphII*, *nal*^r, Fig 5B). A mixture of donor and recipient strain cultured in LB for 24 hours resulted in chloramphenicol, kanamycin and nalidixic acid triple-resistant L01157-10 at a frequency of 1×10^{-8} per CFU. The culture of the donor and recipient strain in LB supplemented with 0.5 mg/ml mitomycin C, which is known to induce prophage activation, increased the frequency by 100-fold to 1×10^{-6} per CFU (Fig 5B). No triple-resistant L01157-10 recipient strains emerged when this strain was cultured on its own in the presence or absence of mitomycin C supplementation. The assembled whole genome sequence of six arbitrarily selected candidate revertant strains was determined. All had the *aphII* gene, a non-synonymous (D87Y) in the *gyrA* gene and the *cat* gene inserted in the 5' intergenic region of *wzy*, expected for reversion (Fig 5C).

A



B

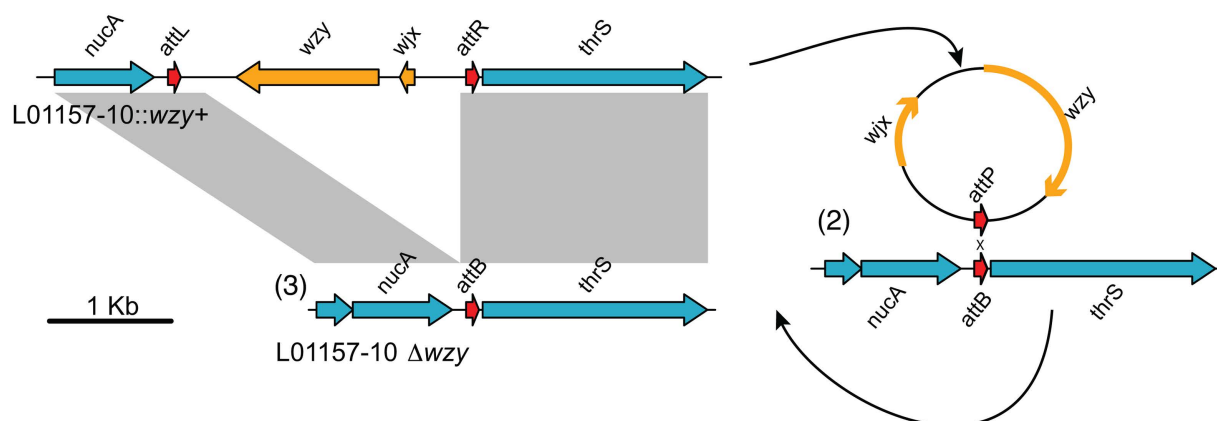


Fig 4. Dependency of *wzy* locus deletion and formation of a circularised *wzy* locus molecule on *attL^{wzy}* repeat sequence. (A) PCR amplification of the *wzy* locus from *S. Typhimurium* DT8 strain L01157-10, a spontaneous phage resistant variant of L01157-10 (L01157-10 Δwzy) and a variant of L01157-10 in which the *attL^{wzy}* repeat sequence was deleted. Inward facing oligonucleotide primers (green arrows) flanking the *wzy* locus (top panel) and outward facing oligonucleotide primers (purple arrows) from within the deleted region (lower panel) were used in separate amplification reactions. Note: these are different primers than those used in Fig 3. (B) Model for the genotypes resulting from excision and circularisation of the *wzy* locus.

<https://doi.org/10.1371/journal.pgen.1011688.g004>

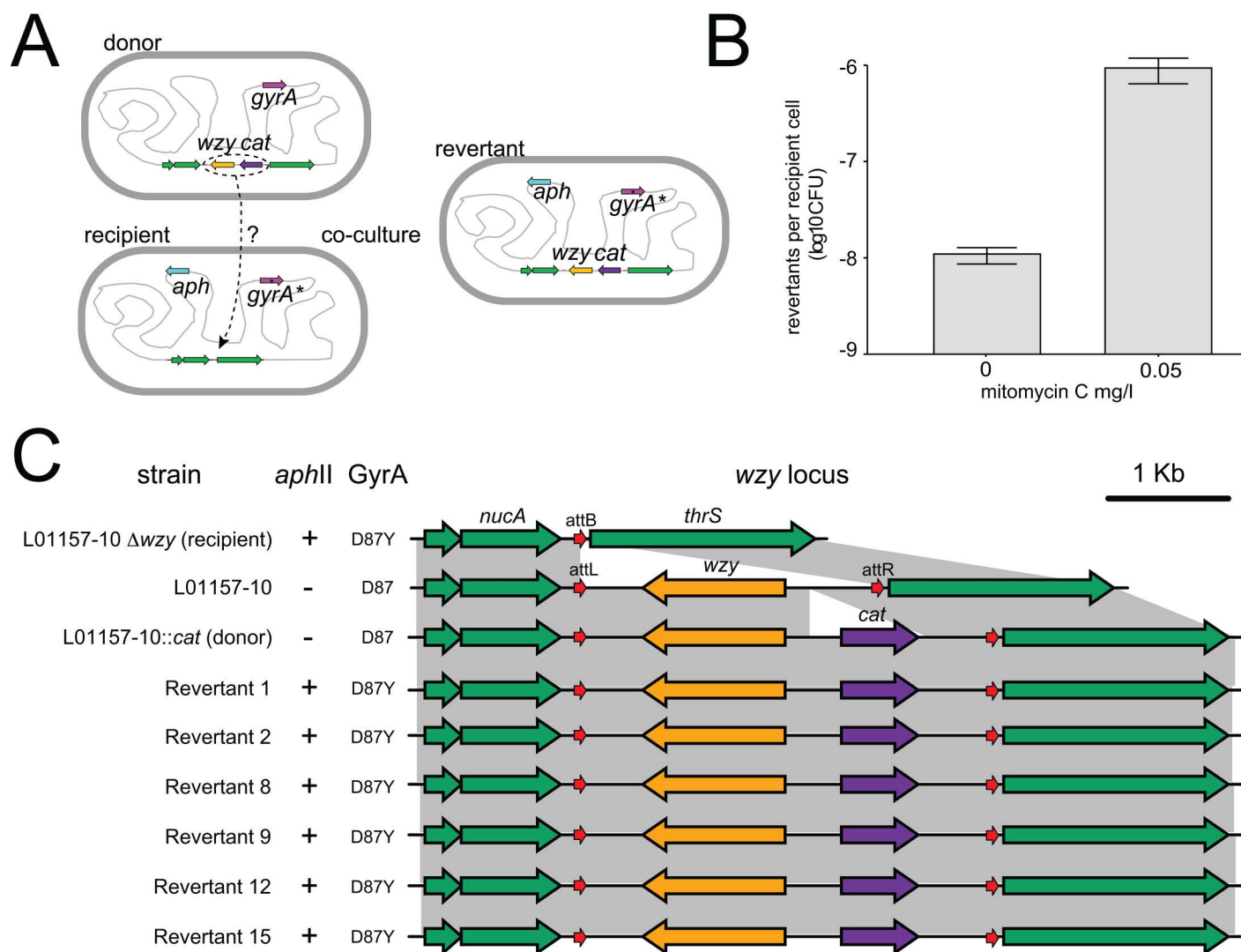


Fig 5. Transmission of *wzy* locus from *S. Typhimurium* wildtype L01157-10 to *S. Typhimurium* L01157-10 Δ*wzy* during co-culture. (A) Genotypes and genomic context of selectable markers and the *wzy* locus in L01157-10 *wzy*::*cat* donor, L01157-10 Δ*wzy* *aph*, *Nal*^r recipient and L01157-10 *wzy*::*cat* *aph*, *Nal*^r revertant strains. (B) Reversion frequency in the presence or absence of mitomycin C supplementation.

<https://doi.org/10.1371/journal.pgen.1011688.g005>

Ancestral state reconstruction of DT8/ DT30 clade is consistent with bidirectional switching between states.

Our laboratory data was consistent with a bidirectional excision and reversion of the *wzy* locus. To investigate the dynamics of phage sensitivity changes in natural populations of *S. Typhimurium*, we investigated the dynamics of DT8 and DT30 phenotype by estimating ancestral state at each bifurcating ancestral node and along each branches of the ML phylogenetic tree of DT8 and DT30 strain constructed previously in this study. We used two independent and complementary methods, a maximum likelihood approach and a Bayesian Markov-Chain with Monte-Carlo sampling (MCMC). The maximum likelihood estimation at each node inferred switches from DT8 to DT30 and from DT30 to DT8 (Fig 6A). Most ancestral nodes were predicted to be DT8, but there were multiple predictions of transitions from DT8 to DT30 and back to DT8 (Fig 6B). A total of 45 changes in phage resistance phenotype were predicted, consisting of 26 switches from DT8 to DT30 and 19 switches from DT30 to DT8.

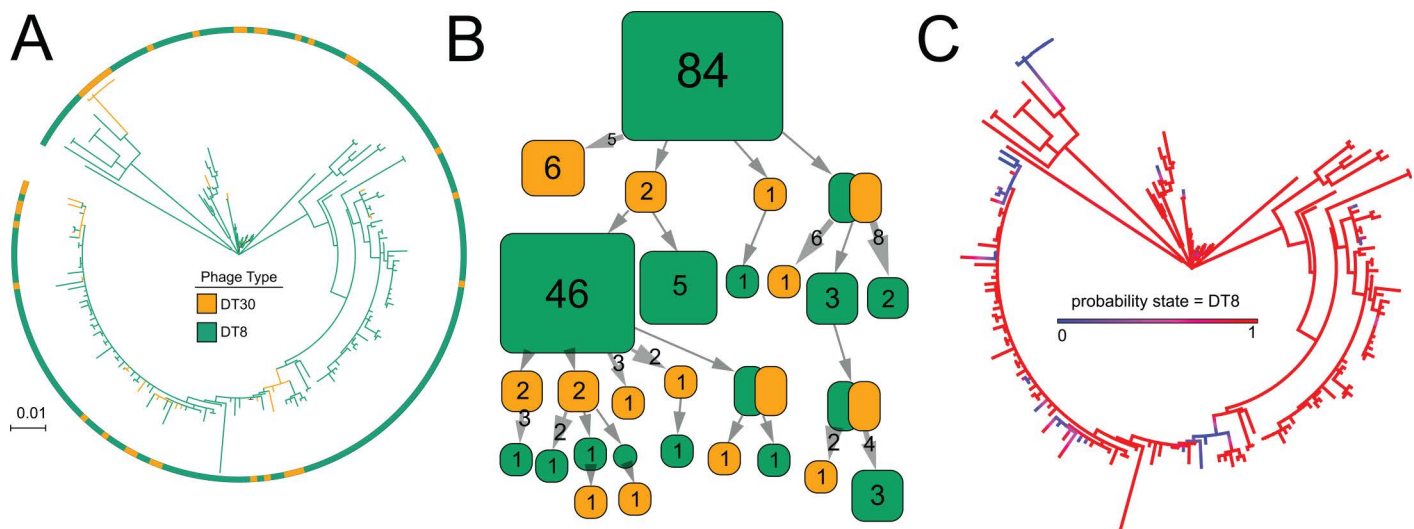


Fig 6. Ancestral state reconstruction of phage type for *S. Typhimurium* strains within the DT8/DT30 clade. A) Maximum likelihood phylogenetic reconstruction of marginal ancestral states with empirical posterior probabilities used to infer an estimated history along the phylogeny. The outer ring indicated phage type DT8 (green) or DT30 (orange). The colours indicated at points in the phylogenetic tree are estimates of whether the state is bacteriophage resistant (DT30, orange) or bacteriophage susceptible (DT8, green). B) Collapsed-branch node graph to visualise the possible changes in state along lineages in (A). Filled squares indicate the state within a particular collapsed branch section of a tree, and how many tips track back to the section for DT8 (green) and DT30 (orange). Arrows indicate the direction of evolution, with the root node at the top and descendant nodes at intervals moving downwards. C) Probability density map indicating the probability that a state across the reconstructed phylogeny is DT8 from 1000 stochastic maps with periodic sampling from a distribution of possible trees with all DT8 probabilities shown as colours. Red indicates a probability of 1, and blue a probability of 0 that the ancestor was DT8.

<https://doi.org/10.1371/journal.pgen.1011688.g006>

Using a complementary Bayesian approach of ancestral estimation, the common ancestor was also predicted to be DT8 (Fig 6C). A mean of 39.659 changes in phenotype state between DT8 and DT30 was estimated based on 1000 simulated trees. Again changes were predicted to be bidirectional with an average of 25.754 switches from DT8 to DT30, and 13.905 for DT30 to DT8. To investigate whether the highest likelihood model where phage type transition rates were unequal was significant due to increased sampling of DT8 strains, rather than based on phylogenetic structure, we performed a permutation test. A total of 18 permutations out of 20 were significantly different from the data (S2A Fig), indicating 90% confidence that the observed data did not occur by chance. The Bayesian MCMC and ML approaches produced models that were highly correlated as indicated by an R^2 of 0.675 for the probability of outcomes for each common ancestor being DT30 from scaled maximum likelihood ancestral states and stochastic mapping (S2B Fig).

Together these analyses were consistent with multiple state changes from DT8 to DT30, back to DT8 consistent with a reversible mechanism consistent with the phage-resistant DT30 phenotype occurring sporadically throughout the population.

Discussion

Co-evolution of bacteria and phage is well-documented *in vitro* [22] and there is mounting evidence that phage predation impacts population structure [28,33]. Phage typing has been used historically to differentiate between closely related strains of *S. Typhimurium* and *S. Enteritidis* based on sensitivity to a panel of phage preparations [66]. Investigation of phage type in the context of phylogenetic relationship of strains is one approach to understand the impact of phage predation on bacterial populations and the mechanisms that ensure fitness and survival. In a subclade of *S. Typhimurium* containing strains adapted to circulation in ducks, APHA surveillance data and our phylogenetic analysis of a subset of

genomes indicated that around 80% were of DT8 and 20% DT30 that were sporadically distributed within the phylogenetic structure. The DT30 strains were characterised by a decrease in sensitivity to phage that use O-antigen as their primary receptor. In *S. Typhimurium*, less than 50–75% of isolates within epidemic clades have the dominant phage type [28] indicating frequent switching in phage sensitivity profile. In the case of the currently dominant pandemic multidrug resistant *S. Typhimurium* ST34 clone [35,67], of 723 strains isolated from human infection in the UK in 2014, around 70% were of the dominant phage type DT193 [28]. The remaining were of 14 different phage types. In this case the ancestor was predicted to be DT120, the second most common phage type in 2014. The switch from DT120 to DT193 was due to acquisition of a prophage, mTmII and resulted in a decrease in sensitivity phage and clonal expansion [28]. Decreased sensitivity to phage predation may be a common characteristic of broad host range livestock associated *S. Typhimurium*, perhaps due to a fitness advantage from diminished phage predation [28]. In contrast, the DT8/DT30 clade contained strains of DT30 distributed within the population of DT8, with little evidence of clonal expansion of the type with reduced phage sensitivity.

The reason for the difference in apparent fitness of DT30 investigated in this study and DT193 studied previously, as assessed by their relative level of clonal expansion [28] is likely to be the result the distinct mechanism of decreased sensitivity to phage. In this study, the *wzy* locus was deleted in 41% of DT30 strains, accounting for the change in sensitivity to phage using O-antigen as receptor and the change in phage type. Mutations in the coding sequence of *wzy* was also present in other DT30 strains. The *wzy* gene encodes an O-antigen polymerase required for the elaboration of long chain O-antigen on the surface that forms a barrier [68] that contributes to resistance to lysozyme [69], oxidative stress [70], colicins [71] and bile acids [72], and also affects the interaction with the complement system [73]. It is therefore not surprising that mutation of *wzy* results in attenuation of virulence in mice, albeit not as attenuated as mutations in *waaG*, *waal*, *waaJ*, *wbaP* and *waaL*, other genes involved in O-antigen biosynthesis that would also be expected to result in resistance to phage that use it as a receptor [74]. In contrast, mTmII lysogeny responsible for the switch from DT120 to DT193 and decreased sensitivity to typing phage by an as yet unknown mechanism appears to have little or no fitness cost.

Genetic and phenotypic heterogeneity is a recurring theme in the interaction of bacteria with the host, environment and bacteriophage. The most common mechanism that generates heterogeneity in an otherwise clonal population is phase variation that results from slip strand mis-pairing, sequence inversion or differential methylation (reviewed in [75]). In bacterial pathogens heterogeneity generated from phase variation is important for functions including evasion of cross-immunity [76], persistence in the host [77], biofilm formation and detachment [78], bistable expression of energetically expensive type III secretion systems [79], and protection against bacteriophage and exogenous DNA [12,80]. The role of heterogeneity in defence against bacteriophage is particularly evident for genes involved in modification of O-antigen, a common receptor for bacteriophage. For example, it has been proposed that phase variation of the *lic2A* gene of *Haemophilus influenzae* encodes a glycosyltransferase that modifies O-antigen by the addition of galactose moiety is involved in a herd-resistance-like phenomenon resulting from reduced phage titres affecting predator-prey dynamics [81]. *S. Typhimurium* also encodes multiple glycosyl transferase genes that modify O-antigen and an O-antigen chain length modulating gene whose expression are also under phase variation that contribute to resistance to bacteriophage [80,82,83]. We describe a potentially novel form of phase variation in which heterogeneity is generated by reversible deletion of the *wzy* gene required to elaborate long chain O-antigen of LPS.

The mechanism for excision of the *wzy* locus is not known but we establish that it requires at least attL, one of a pair of 113bp direct repeats flanking the *wzy* locus. The presence of attL and attR suggest that the mechanism may be analogous to other excision and integration systems that employ integrase and in some case with an excisionase [65]. However, direct repeat sequences flanking the *wzy* locus are unlike the att repeat sequences of phage that are inverted repeats and often form attB and attC sequences with limited sequence similarity except for short (3–10 bp)

sequences where homologous recombination occurs. The *wzy* locus repeats bear some resemblance to the FRT and *loxP* sites of the site-specific recombinase systems from *Saccharomyces cerevisiae* and P1 phage, respectively, that can operate to excise or integrate sequence based on direct repeat sequences, catalysed by a site-specific recombinase. However, FRT and *LoxP* sequences are also distinct from the *wzy* locus attL and attR, each being only 34 bp and composed of two symmetrical 13 bp recognition sequences with a variable central sequence. In comparison, the *wzy* locus repeats are 113 bp and lack symmetry within the sequence. We were unable to identify site-specific recombinase target sequences with a similar length and characteristics to that of the attL and attR. The relatively long and identical repeats may alternatively support the possibility that excision and integration was mediated by RecA dependent or independent recombination. RecA dependent recombination occurs between repeat sequences of greater than the minimal efficient processing segment (MEP) of 23–27 bp (RecA/RecBC) and 44–90 bp (RecA/RecF) with little increase in frequency above around 100 bp [84]. The greater the distance between repeats or decrease in sequence identity of the repeat sequence decreases the efficiency of recombination, and our observations are therefore consistent with such a mechanism. A RecA independent mechanism involving crossover between replicating sister strands, may also play a role [85], although it is not clear if this is capable of forming a circular excised molecule as observed for the *wzy* locus.

The potential for reversion of *S. Typhimurium* strains with the *wzy* gene deletion back to wild type by transfer of *wzy* from bystander *S. Typhimurium* is a critical observation that identifies this as a mechanism that is potentially analogous to classical phase variation [75]. The mechanism of *wzy* locus transfer is not known but the frequency increases in the presence of mitomycin C that is known to induce the SOS response and activate prophage [86]. Prophage mediated genetic transfer has been observed in *S. Typhimurium* that was induced by the antibiotic carbadox [87] and prophage-like gene transfer agent (GTAs) in several bacterial species, that are thought to be cryptic prophage that retain the ability to package host DNA and carry out generalised transduction [88]. Circularisation of the *wzy* locus in the donor strain may increase the frequency of transfer by protecting the molecule from endonucleases. Induction of the SOS response also increases expression of RecA and RecN that may be involved in homologous recombination at the attB site of the recipient and the attC site of the circularised *wzy* locus [89]. However, it also remains possible that an integrase is responsible for this genetic event and this too has the potential to be up-regulated as part of the SOS response.

The observation that *wzy* deletion results in resistance to phage that use O-antigen as a receptor and the potential for subsequent reversion raises the possibility that this is a mechanism for recovery following phage predation. We determined that in populations of *S. Typhimurium* DT8 cultured in the absence of phage, approximately 1 in 500 had deleted the *wzy* locus, providing a pool of resistant variants that in the case of predation by phage using O-antigen as a receptor will survive the initial attack. However, *wzy* mutants have a fitness cost in the host and potentially in the environment and therefore reversion is key to this being an evolutionarily stable mechanism. While a circular form of the *wzy* locus was detectable in cultures of DT8 L01157-10 this appeared to be transient since this form was not detected in a variant selected by virtue of resistance to phage and that had excised the *wzy* locus. Reversion by reintegration of the *wzy* locus from the circular form in the same cell cannot occur in this case, although we cannot discount the possibility that reintegration of the circular form could occur.

Modelling of ancestral states in the population structure of *S. Typhimurium* DT8 and DT30 strains was consistent with the reversion of Δwzy DT30 strains to DT8. In the case of *S. Typhimurium* Δwzy variants to enable reversion, it is essential that *S. Typhimurium* with an intact *wzy* locus is also present in the population following phage predation. Importantly, we determined that following culture with a phage that used O-antigen as a receptor, although the frequency of Δwzy variants in the population increased to 1 in 30, still around 97% of surviving *S. Typhimurium* had an intact *wzy* locus. The *wzy*⁺ variants in these cultures were likely resistant to the phage due to mutations in other

O-antigen biosynthesis genes such a single base insertion in the *pgm* gene that resulted in a frame shift predicted to truncate this gene described previously [12]. The potential for *wzy*⁺ and *wzy*⁻ variants to survive in the host or environment for long enough to allow for the predating phage to disperse is not known and this is a key unanswered question. During infections of the host, strains of *S. Typhimurium* with mutations in various O-antigen biosynthesis genes were able to colonise the Peyer's patches, liver and spleen of mice following oral inoculation, but at lower levels compared to the wild-type strain. It is also noteworthy that mutation of *wzy* had the least impact of six genes that were tested on colonisation [74] suggesting that although Δwzy variants may be present in only a relatively small proportion of the surviving population following phage predation, the *wzy*⁺ proportion may increase over time. The ability of strains harbouring mutations in various O-antigen biosynthesis genes to persist in the intestine is not known and this is an important factor that will determine the opportunity for reversion to occur. Another important factor is the dynamics of clearance of phage from the intestine once all phage sensitive strains have been killed and this is also currently not known.

Changes in O-antigen length due to phase variation of genes encoding chain length determination protein *opvAB* has also been described in *S. Typhimurium* and may represent a complementary system to that the *wzy* deletion [83]. Phase variation results in a sub-population of *Salmonella* which have *opvAB*^{on} and *opvAB*^{off} allowing defence against certain phage that utilise O-antigen as a receptor, that was also recognised to come with an associated cost to virulence. In our study we did not investigate *opvAB* expression and therefore cannot compare the relative impact on evading phage predation with that of *wzy* deletion. It remains an open question as to how these two systems complement each other in evading phage predation.

All *S. Typhimurium* strains that we investigated contained the att sequence. Although we did not determine the presence of the att sequence in other serotypes, it is likely to be conserved in *Salmonella enterica* serovars of serogroups B1, D1 and A and at least some subspecies II serovars that encode an alternative *wzy* (*rfc*) encoded outside of the *rfb* locus [64,90]. Serogroup B1, D1 and A serovars are thought to have acquired the *wzy* in the *rfc* locus and the ancestral *wzy* in the *rfb* locus has been deleted leaving a sequence remnant [64,74]. Although only a subset of *S. enterica* serovars have the *wzy* (*rfc*) gene, these represent serotypes that account for the vast majority of human and livestock infections. For example, five of the top ten serovars accounting for non-typhoidal infections in people in the UK in 2022 are of serogroups B1 or D1 including Enteritidis, Typhimurium, Agona, Paratyphi B (Java) and Saintpaul, that together account for over half of all infections. Typhoidal *Salmonella* is a major cause of morbidity and mortality in South Asia, Africa and Latin America and *S. Typhi* and *S. Paratyphi A* that cause these infections both encode *wzy* in the *rfc* locus. Previous report of deletion of the *wzy* locus in strains of *S. Enteritidis* indicate that the deletion mechanism described in our study are likely to occur in other serovars with the *wzy* gene in the *rfc* locus [91].

In this study we investigated the deletion of *wzy*, the selection of a *wzy*⁻ subpopulation during predation by phage that use O-antigen as a receptor and presented evidence of subsequent reversion to wild type. In this context we propose a working model in which phage predation results in the killing of sensitive *S. Typhimurium* and the survival of a subpopulation that have either deleted the *wzy* gene or have mutations in other O-antigen biosynthesis genes but retain the *wzy* locus (Fig 7). We propose that during infection or in the environment following phage predation, the absence of phage sensitive *S. Typhimurium* will result in clearance of the phage. The surviving subpopulation can therefore transfer the *wzy* locus from strains with alternate O-antigen mutations into Δwzy *S. Typhimurium* reconstituting wild type *S. Typhimurium*. In this scenario, we would expect failure of phage treatment that could be tackled with repeated or continuous application of the phage. Alternatively, treatment with combinations of phage that exhibit collateral sensitivity due to the targeting of alternative receptors may also be effective in countering the proposed *wzy* deletion mechanism. These considerations are important to the effective deployment of phage as novel antimicrobials against *Salmonella*.

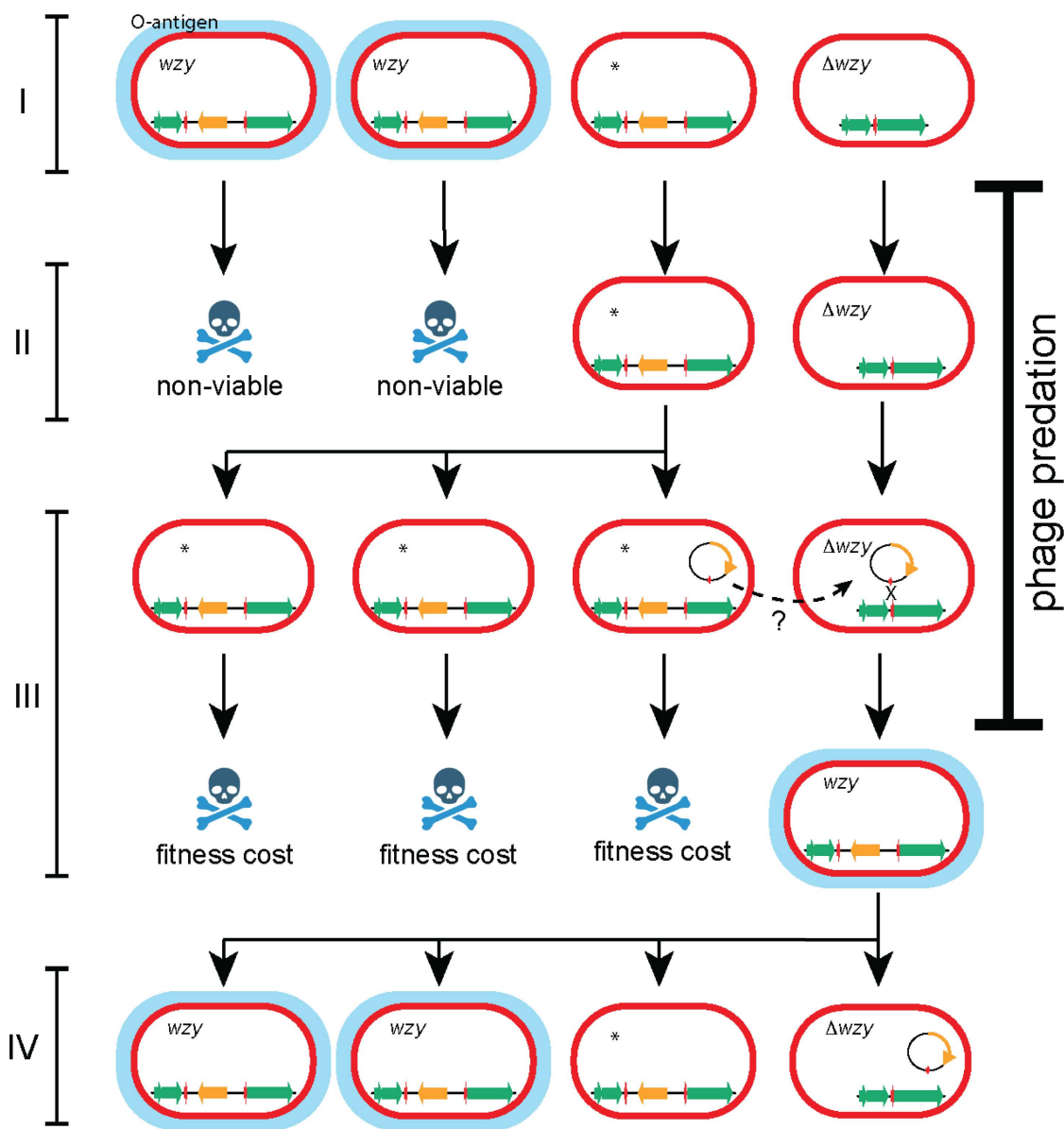


Fig 7. A model for how reversible deletion of the *wzy* may contribute to resistance to phage killing and recovery following transient predation by a lytic phage that recognises O-antigen as primary receptor. Ovals represent bacteria of wild type genotype (*wzy*) expressing O-antigen (blue oval) or genotype variants with Δwzy or variants with mutations in other O-antigen biosynthesis genes marked *, lacking O-antigen. Vertical arrows represent the outcome of replication, phage predation or reduced fitness. Transfer of the *wzy* locus by an unknown mechanism (broken line arrow). I) Population of *S. Typhimurium* DT8 composed of wild type (*wzy*), Δwzy or variants with mutations in other O-antigen biosynthesis genes marked *. II) On phage predation wild type *S. Typhimurium* are killed and Δwzy and variants with mutations in mutations in other O-antigen biosynthesis genes survive. III) The lack of O-antigen confers a fitness cost and are non-viable in the medium to long term. On cessation of phage predation wild type *S. Typhimurium* can be reconstituted by transfer of *wzy* from variants with mutations in mutations in other O-antigen biosynthesis genes. IV) Reconstitution of the *wzy* locus results in viable *S. Typhimurium* DT8 identical to ancestors prior to phage predation that replicates giving rise to a mixed population of Δwzy or variants with mutations in other O-antigen biosynthesis genes marked. Created in BioRender. Kolenda, R. (2025) <https://BioRender.com/o00p885>.

<https://doi.org/10.1371/journal.pgen.1011688.g007>

Supporting information

S1 Fig. Genetic Map showing the position and sequence of attL and attR. Filled arrows indicate the *nucA* and *thrS* genes (blue) that flank the *wzy* gene (orange). The location of the attL and attR sequence are indicated (red).
(EPS)

S2 Fig. Permutation test to assess if the probability distributions of each node being DT30 from each permutation were significantly different to the observed data. (A) permutation test with 20 DT8/DT30 phenotype permutations and tested against the observed data using Mann-Whitney-Wilcoxon tests were performed and $p > 0.05$ are indicated (*) or not significant (N.S). (B) scatter plot where individual nodes (blue circles) represent the DT30 state probability at each node determined by SIMMAP or PastML methodology. Best fit for the linear model is indicated (red broken line).
(EPS)

S1 Table. Strains and genome sequence accession used in this study.
(XLSX)

S2 Table. Oligonucleotide primers used for genetic engineering of *S. Typhimurium* or PCR amplification in this study.
(XLSX)

S3 Table. Genome sequence of 134 representative strains of *S. Typhimurium* and presence of *wzy* and att sequence determined using ARIBA.(A).
(XLSX)

S1 File. Scripts used in this study. Commands used for running Snippy, RAxML, Pyseer, and R scripts used for ancestral state reconstruction with phytools and simmap, and pastML.
(DOCX)

S2 File. Expected sequence and amplicon sequence of circularised *wzy* locus. Sequence 'circular_wzy_locus' is the expected sequence of the *wzy* locus following hypothetical circularisation by recombination of the attL and attR sequences in the L01157-10 genome. Sequence 'consensus_PCR' is the consensus of the forward and reverse sequences of the amplicon from the hypothetical circularised *wzy* locus in [Fig 4A](#).
(TXT)

S3 File. Alignment of circular_wzy_locus and consensus_PCR. Alignment of sequences using Muscle software.
(TXT)

S4 File. Numerical data for graphs in [Figs 1](#), [3](#) and [5](#).
(XLSX)

Author contributions

Conceptualization: Oliver JD Charity, Gaetan Thilliez, Robert A Kingsley.

Data curation: Oliver JD Charity, Haider Al-Khanaq, Rafał Kolenda, Matt Bawn.

Formal analysis: Oliver JD Charity, Gaetan Thilliez, Rafał Kolenda.

Funding acquisition: Liljana Petrovska, Robert A Kingsley.

Investigation: Oliver JD Charity, Gaetan Thilliez, Robert A Kingsley.

Methodology: Oliver JD Charity, Gaetan Thilliez, Haider Al-Khanaq, Luke Acton, Matt Bawn, Robert A Kingsley.

Project administration: Robert A Kingsley.

Resources: Liljana Petrovska, Robert A Kingsley.

Software: Oliver JD Charity, Gaetan Thilliez, Rafał Kolenda, Matt Bawn.

Supervision: Robert A Kingsley.

Visualization: Oliver JD Charity, Gaetan Thilliez, Robert A Kingsley.

Writing – original draft: Oliver JD Charity, Gaetan Thilliez, Robert A Kingsley.

Writing – review & editing: Gaetan Thilliez, Haider Al-Khanaq, Liljana Petrovska, Robert A Kingsley.

References

- Branchu P, Bawn M, Kingsley RA. Genome variation and molecular epidemiology of salmonella enterica serovar typhimurium pathovariants. *Infect Immun*. 2018;86(8):e00079-18. <https://doi.org/10.1128/IAI.00079-18> PMID: [29784861](#)
- Salmonella in livestock production in gb—2015 2016. APHA; 2016. <https://www.gov.uk/government/publications/salmonella-in-livestock-production-in-great-britain-2015>
- Woolhouse M, Ward M, van Bunnik B, Farrar J. Antimicrobial resistance in humans, livestock and the wider environment. *Philos Trans R Soc Lond B Biol Sci*. 2015;370(1670):20140083. <https://doi.org/10.1098/rstb.2014.0083> PMID: [25918441](#)
- Pereira C, Moreirinha C, Lewicka M, Almeida P, Clemente C, Cunha Â, et al. Bacteriophages with potential to inactivate Salmonella Typhimurium: Use of single phage suspensions and phage cocktails. *Virus Res*. 2016;220:179–92. <https://doi.org/10.1016/j.virusres.2016.04.020> PMID: [27126773](#)
- Wall SK, Zhang J, Rostagno MH, Ebner PD. Phage therapy to reduce preprocessing salmonella infections in market-weight swine. *Appl Environ Microbiol*. 2010;76(1):48–53.
- Hooton SPT, Atterbury RJ, Connerton IF. Application of a bacteriophage cocktail to reduce Salmonella Typhimurium U288 contamination on pig skin. *Int J Food Microbiol*. 2011;151(2):157–63. <https://doi.org/10.1016/j.ijfoodmicro.2011.08.015> PMID: [21899907](#)
- Bai J, Jeon B, Ryu S. Effective inhibition of Salmonella Typhimurium in fresh produce by a phage cocktail targeting multiple host receptors. *Food Microbiol*. 2019;77:52–60. <https://doi.org/10.1016/j.fm.2018.08.011> PMID: [30297056](#)
- Saha D, Mukherjee R. Ameliorating the antimicrobial resistance crisis: phage therapy. *IUBMB Life*. 2019.
- Kortright KE, Chan BK, Koff JL, Turner PE. Phage therapy: a renewed approach to combat antibiotic-resistant bacteria. *Cell Host Microbe*. 2019;25(2):219–32. <https://doi.org/10.1016/j.chom.2019.01.014> PMID: [30763536](#)
- Tang F, Zhang P, Zhang Q, Xue F, Ren J, Sun J, et al. Isolation and characterization of a broad-spectrum phage of multiple drug resistant Salmonella and its therapeutic utility in mice. *Microb Pathog*. 2019;126:193–8. <https://doi.org/10.1016/j.micpath.2018.10.042> PMID: [30408490](#)
- Florez M, Sanchez A, Moreno A. Composition comprising bacteriophages for reducing, eliminating and/or preventing salmonella enteritidis, salmonella typhimurium and salmonella paratyphi b. Google Patents; 2019.
- Acton L, Pye HV, Thilliez G, Kolenda R, Matthews M, Turner AK, et al. Collateral sensitivity increases the efficacy of a rationally designed bacteriophage combination to control Salmonella enterica. *J Virol*. 2024;98(3):e0147623. <https://doi.org/10.1128/jvi.01476-23> PMID: [38376991](#)
- Georjon H, Bernheim A. The highly diverse antiphage defence systems of bacteria. *Nat Rev Microbiol*. 2023;21(10):686–700. <https://doi.org/10.1038/s41579-023-00934-x> PMID: [37460672](#)
- Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat Rev Microbiol*. 2020;18(2):113–9. <https://doi.org/10.1038/s41579-019-0278-2> PMID: [31695182](#)
- Samson JE, Magadán AH, Sabri M, Moineau S. Revenge of the phages: defeating bacterial defences. *Nat Rev Microbiol*. 2013;11(10):675–87. <https://doi.org/10.1038/nrmicro3096> PMID: [23979432](#)
- Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol Rev*. 2014;38(5):916–31. <https://doi.org/10.1111/1574-6976.12072> PMID: [24617569](#)
- Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, et al. Antagonistic coevolution accelerates molecular evolution. *Nature*. 2010;464(7286):275–8. <https://doi.org/10.1038/nature08798> PMID: [20182425](#)
- Pal C, Maciá MD, Oliver A, Schachar I, Buckling A. Coevolution with viruses drives the evolution of bacterial mutation rates. *Nature*. 2007;450(7172):1079–81. <https://doi.org/10.1038/nature06350> PMID: [18059461](#)
- Gómez P, Buckling A. Bacteria-phage antagonistic coevolution in soil. *Science*. 2011;332(6025):106–9. <https://doi.org/10.1126/science.1198767> PMID: [21454789](#)
- Gencay YE, Gambino M, Prüssing TF, Brøndsted L. The genera of bacteriophages and their receptors are the major determinants of host range. *Environ Microbiol*. 2019;21(6):2095–111. <https://doi.org/10.1111/1462-2920.14597> PMID: [30888719](#)
- Lenski RE, Levin BR. Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities *Am Nat*. 1985;125(4):585–602. <https://doi.org/10.1086/284364>

22. Brockhurst MA, Morgan AD, Fenton A, Buckling A. Experimental coevolution with bacteria and phage. The *Pseudomonas fluorescens*--Phi2 model system. *Infect Genet Evol*. 2007;7(4):547–52. <https://doi.org/10.1016/j.meegid.2007.01.005> PMID: [17320489](#)
23. Hall AR, Scanlan PD, Morgan AD, Buckling A. Host-parasite coevolutionary arms races give way to fluctuating selection. *Ecol Lett*. 2011;14(7):635–42. <https://doi.org/10.1111/j.1461-0248.2011.01624.x> PMID: [21521436](#)
24. Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol Rev*. 2014;38(5):916–31. <https://doi.org/10.1111/1574-6976.12072> PMID: [24617569](#)
25. Raetz CRH, Whitfield C. Lipopolysaccharide endotoxins. *Annu Rev Biochem*. 2002;71:635–700. <https://doi.org/10.1146/annurev.biochem.71.110601.135414> PMID: [12045108](#)
26. Kim M, Ryu S. Spontaneous and transient defence against bacteriophage by phase-variable glucosylation of O-antigen in *Salmonella enterica* serovar Typhimurium. *Mol Microbiol*. 2012;86(2):411–25. <https://doi.org/10.1111/j.1365-2958.2012.08202.x> PMID: [22928771](#)
27. Brockhurst MA, Rainey PB, Buckling A. The effect of spatial heterogeneity and parasites on the evolution of host diversity. *Proc Biol Sci*. 2004;271(1534):107–11.
28. Charity OJ, Acton L, Bawn M, Tassinari E, Thilliez G, Chattaway MA, et al. Increased phage resistance through lysogenic conversion accompanying emergence of monophasic *Salmonella* Typhimurium ST34 pandemic strain. *Microb Genom*. 2022;8(11):mgen000897. <https://doi.org/10.1099/mgen.0.000897> PMID: [36382789](#)
29. Anderson ES, Ward LR, Saxe MJ, de Sa JD. Bacteriophage-typing designations of *Salmonella* typhimurium. *J Hyg (Lond)*. 1977;78(2):297–300. <https://doi.org/10.1017/s0022172400056187> PMID: [321679](#)
30. Anderson E, Ward L, de Saxe M, De Sa J. Bacteriophage-typing designations of *Salmonella* typhimurium. *Epidemiol Infect*. 1977;78(2):297–300.
31. Reuter S, Ellington MJ, Cartwright EJP, Köser CU, Török ME, Gouliouris T, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. *JAMA Intern Med*. 2013;173(15):1397–404. <https://doi.org/10.1001/jamainternmed.2013.7734> PMID: [23857503](#)
32. Mohammed M, Casjens SR, Millard AD, Harrison C, Gannon L, Chattaway MA. Genomic analysis of Anderson typing phages of *Salmonella* Typhimurium: towards understanding the basis of bacteria-phage interaction. *Sci Rep*. 2023;13(1):10484. <https://doi.org/10.1038/s41598-023-37307-6> PMID: [37380724](#)
33. Cowley LA, Dallman TJ, Jenkins C, Sheppard SK. Phage predation shapes the population structure of shiga-toxigenic *Escherichia coli* O157:H7 in the UK: an evolutionary perspective. *Front Genet*. 2019;10:763. <https://doi.org/10.3389/fgene.2019.00763> PMID: [31543896](#)
34. Branchu P, Charity OJ, Bawn M, Thilliez G, Dallman TJ, Petrovska L, et al. SGI-4 in monophasic *salmonella* typhimurium ST34 Is a Novel ICE that enhances resistance to copper. *Front Microbiol*. 2019;10:1118. <https://doi.org/10.3389/fmicb.2019.01118> PMID: [31178839](#)
35. Petrovska L, Mather A, AbuOun M, Branchu P, Harris S, Connor T. Microevolution of monophasic *Salmonella* Typhimurium during epidemic, United Kingdom, 2005–2010. *Emerg Infect Dis*. 2016;22(4):617.
36. Ashton PM, Peters T, Ameh L, McAleer R, Petrie S, Nair S. Whole genome sequencing for the retrospective investigation of an outbreak of *salmonella* typhimurium dt 8. *PLoS Curr*. 2015;7.
37. Garvey P, McKeown P, Kelly P, Cormican M, Anderson W, Flack A. Investigation and management of an outbreak of *salmonella* typhimurium dt8 associated with duck eggs, Ireland 2009 to 2011. *Eurosurveillance*. 2013;18(12):20416. <https://doi.org/10.2807/ese.18.12.20416-en>
38. APHA. *Salmonella* in livestock production in Great Britain, 2023/2024. Available from: <https://www.gov.uk/government/publications/salmonella-in-livestock-production-in-great-britain-2016>.
39. APHA. *Salmonella* in Livestock Production in GB 2013. UK: Animal and Plant Health Agency; 2014. Available from: www.gov.uk/government/publications
40. Bawn M, Alikhan N-F, Thilliez G, Kirkwood M, Wheeler NE, Petrovska L, et al. Evolution of *Salmonella enterica* serotype Typhimurium driven by anthropogenic selection and niche adaptation. *PLoS Genet*. 2020;16(6):e1008850. <https://doi.org/10.1371/journal.pgen.1008850> PMID: [32511244](#)
41. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754–60. <https://doi.org/10.1093/bioinformatics/btp324> PMID: [19451168](#)
42. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:12073907*. 2012.
43. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. *Bioinformatics*. 2011;27(15):2156–8. <https://doi.org/10.1093/bioinformatics/btr330> PMID: [21653522](#)
44. Seemann T. Snippy: rapid haploid variant calling and core SNP phylogeny. 2015.
45. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*. 2006;22(21):2688–90. <https://doi.org/10.1093/bioinformatics/btl446> PMID: [16928733](#)
46. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A*. 2000;97(12):6640–5. <https://doi.org/10.1073/pnas.120163297> PMID: [10829079](#)
47. Kirkwood M, Vohra P, Bawn M, Thilliez G, Pye H, Tanner J, et al. Ecological niche adaptation of *Salmonella* Typhimurium U288 is associated with altered pathogenicity and reduced zoonotic potential. *Commun Biol*. 2021;4(1):498. <https://doi.org/10.1038/s42003-021-02013-4> PMID: [33893390](#)
48. Sharan SK, Thomason LC, Kuznetsov SG, Court DL. Recombineering: a homologous recombination-based method of genetic engineering. *Nat Protoc*. 2009;4(2):206–23. <https://doi.org/10.1038/nprot.2008.227> PMID: [19180090](#)

49. Holden ER, Wickham GJ, Webber MA, Thomson NM, Trampari E. Donor plasmids for phenotypically neutral chromosomal gene insertions in Enterobacteriaceae. *Microbiology (Reading)*. 2020;166(12):1115–20. <https://doi.org/10.1099/mic.0.000994> PMID: [33226934](#)
50. Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One*. 2008;3(11):e3647. <https://doi.org/10.1371/journal.pone.0003647> PMID: [18985154](#)
51. Pagel M, Meade A, Barker D. Bayesian estimation of ancestral character states on phylogenies. *Syst Biol*. 2004;53(5):673–84. <https://doi.org/10.1080/10635150490522232> PMID: [15545248](#)
52. Bollback JP. SIMMAP: stochastic character mapping of discrete traits on phylogenies. *BMC Bioinformatics*. 2006;7:88. <https://doi.org/10.1186/1471-2105-7-88> PMID: [16504105](#)
53. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol*. 2011;3(2):217–23. <https://doi.org/10.1111/j.2041-210x.2011.00169.x>
54. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*. 2007;23(1):127–8. <https://doi.org/10.1093/bioinformatics/btl529> PMID: [17050570](#)
55. Ishikawa SA, Zhukova A, Iwasaki W, Gascuel O. A fast likelihood method to reconstruct and visualize ancestral scenarios. *Mol Biol Evol*. 2019;36(9):2069–85. <https://doi.org/10.1093/molbev/msz131> PMID: [31127303](#)
56. Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics*. 2018;34(24):4310–2. <https://doi.org/10.1093/bioinformatics/bty539> PMID: [30535304](#)
57. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol*. 2016;17(1):132. <https://doi.org/10.1186/s13059-016-0997-x> PMID: [27323842](#)
58. Lees JA, Vehkala M, Välimäki N, Harris SR, Chewapreecha C, Croucher NJ, et al. Sequence element enrichment analysis to determine the genetic basis of bacterial phenotypes. *Nat Commun*. 2016;7:12797. <https://doi.org/10.1038/ncomms12797> PMID: [27633831](#)
59. Langmead B, Salzberg S. Bowtie2. *Nat Methods*. 2013;9:357–9.
60. Quinlan AR. BEDTools: the swiss-army tool for genome feature analysis. *Curr Protoc Bioinformatics*. 2014;47:11.12.1–34. <https://doi.org/10.1002/0471250953.bi1112s47> PMID: [25199790](#)
61. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol*. 2016;8(1):28–36. <https://doi.org/10.1111/2041-210x.12628>
62. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*. 2004;5:113. <https://doi.org/10.1186/1471-2105-5-113> PMID: [15318951](#)
63. Mohammed M, Cormican M. Whole genome sequencing provides possible explanations for the difference in phage susceptibility among two *Salmonella* Typhimurium phage types (DT8 and DT30) associated with a single foodborne outbreak. *BMC Res Notes*. 2015;8:728. <https://doi.org/10.1186/s13104-015-1687-6> PMID: [26613761](#)
64. Reeves PR, Cunneen MM, Liu B, Wang L. Genetics and evolution of the *Salmonella* galactose-initiated set of o antigens. *PLoS One*. 2013;8(7):e69306. <https://doi.org/10.1371/journal.pone.0069306> PMID: [23874940](#)
65. Groth AC, Calos MP. Phage integrases: biology and applications. *J Mol Biol*. 2004;335(3):667–78. <https://doi.org/10.1016/j.jmb.2003.09.082> PMID: [14687564](#)
66. Rabsch W, Truepschuch S, Windhorst D, Gerlach RG. Typing phages and prophages of *Salmonella*. Norfolk, UK: Caister Academic Press; 2011.
67. Ingle DJ, Ambrose RL, Baines SL, Duchene S, Gonçalves da Silva A, Lee DYJ, et al. Evolutionary dynamics of multidrug resistant *Salmonella* enterica serovar 4,[5],12:i:- in Australia. *Nat Commun*. 2021;12(1):4786. <https://doi.org/10.1038/s41467-021-25073-w> PMID: [34373455](#)
68. Blasco P, Patel DS, Engström O, Im W, Widmalm G. Conformational dynamics of the lipopolysaccharide from *Escherichia coli* O91 revealed by nuclear magnetic resonance spectroscopy and molecular simulations. *Biochemistry*. 2017;56(29):3826–39. <https://doi.org/10.1021/acs.biochem.7b00106> PMID: [28609625](#)
69. Bao Y, Zhang H, Huang X, Ma J, Logue CM, Nolan LK, et al. O-specific polysaccharide confers lysozyme resistance to extraintestinal pathogenic *Escherichia coli*. *Virulence*. 2018;9(1):666–80. <https://doi.org/10.1080/21505594.2018.1433979> PMID: [29405825](#)
70. Zheng Y, Wang H, Huang L, Zhang T, Zong B, Ren X, et al. Effect of O antigen ligase gene mutation on oxidative stress resistance and pathogenicity of NMEC strain RS218. *Microb Pathog*. 2019;136:103656. <https://doi.org/10.1016/j.micpath.2019.103656> PMID: [31400443](#)
71. Tran ENH, Papadopoulos M, Morona R. Relationship between O-antigen chain length and resistance to colicin E2 in *Shigella flexneri*. *Microbiology (Reading)*. 2014;160(Pt 3):589–601. <https://doi.org/10.1099/mic.0.074955-0> PMID: [24425769](#)
72. Crawford RW, Keestra AM, Winter SE, Xavier MN, Tsois RM, Tolstikov V, et al. Very long O-antigen chains enhance fitness during *Salmonella*-induced colitis by increasing bile resistance. *PLoS Pathog*. 2012;8(9):e1002918. <https://doi.org/10.1371/journal.ppat.1002918> PMID: [23028318](#)
73. Grossman N, Schmetz M, Foulds J, Klima E, Jimenez-Lucho V, Leive L. Lipopolysaccharide size and distribution determine serum resistance in *Salmonella montevideo*. *J Bacteriol*. 1987;169(2):856–63.
74. Kong Q, Yang J, Liu Q, Alamuri P, Roland KL, Curtiss R 3rd. Effect of deletion of genes involved in lipopolysaccharide core and O-antigen synthesis on virulence and immunogenicity of *Salmonella enterica* serovar typhimurium. *Infect Immun*. 2011;79(10):4227–39. <https://doi.org/10.1128/IAI.05398-11> PMID: [21768282](#)

75. van der Woude MW. Phase variation: how to create and coordinate population diversity. *Curr Opin Microbiol*. 2011;14(2):205–11. <https://doi.org/10.1016/j.mib.2011.01.002> PMID: [21292543](#)
76. vanderWoude M, Baumlér A. Phase and antigenic variation in bacteria. *Clin Microbiol Rev*. 2004;17(3):581–611.
77. Norris TL, Bäumlér AJ. Phase variation of the *lpf* operon is a mechanism to evade cross-immunity between *Salmonella* serotypes. *Proc Natl Acad Sci U S A*. 1999;96(23):13393–8. <https://doi.org/10.1073/pnas.96.23.13393> PMID: [10557331](#)
78. Tormo MÁ, Úbeda C, Martí M, Maiques E, Cucarella C, Valle J, et al. Phase-variable expression of the biofilm-associated protein (Bap) in *Staphylococcus aureus*. *Microbiology (Reading)*. 2007;153(Pt 6):1702–10. <https://doi.org/10.1099/mic.0.2006/005744-0> PMID: [17526828](#)
79. Diard M, Garcia V, Maier L, Remus-Emsermann MNP, Regoes RR, Ackermann M, et al. Stabilization of cooperative virulence by the expression of an avirulent phenotype. *Nature*. 2013;494(7437):353–6. <https://doi.org/10.1038/nature11913> PMID: [23426324](#)
80. Broadbent SE, Davies MR, van der Woude MW. Phase variation controls expression of *Salmonella* lipopolysaccharide modification genes by a DNA methylation-dependent mechanism. *Mol Microbiol*. 2010;77(2):337–53. <https://doi.org/10.1111/j.1365-2958.2010.07203.x> PMID: [20487280](#)
81. Turkington CJR, Morozov A, Clokie MRJ, Bayliss CD. Phage-resistant phase-variant sub-populations mediate herd immunity against bacteriophage invasion of bacterial meta-populations. *Front Microbiol*. 2019;10:1473. <https://doi.org/10.3389/fmicb.2019.01473> PMID: [31333609](#)
82. Davies MR, Broadbent SE, Harris SR, Thomson NR, van der Woude MW. Horizontally acquired glycosyltransferase operons drive salmonellae lipopolysaccharide diversity. *PLoS Genet*. 2013;9(6):e1003568. <https://doi.org/10.1371/journal.pgen.1003568> PMID: [23818865](#)
83. Cota I, Sánchez-Romero MA, Hernández SB, Pucciarelli MG, García-Del Portillo F, Casadesús J. Epigenetic control of salmonella enterica O-antigen chain length: a tradeoff between virulence and bacteriophage resistance. *PLoS Genet*. 2015;11(11):e1005667. <https://doi.org/10.1371/journal.pgen.1005667> PMID: [26583926](#)
84. Watt VM, Ingles CJ, Urdea MS, Rutter WJ. Homology requirements for recombination in *Escherichia coli*. *Proc Natl Acad Sci U S A*. 1985;82(14):4768–72. <https://doi.org/10.1073/pnas.82.14.4768> PMID: [3161076](#)
85. Lovett ST, Drapkin PT, Suter VA Jr, Gluckman-Peskind TJ. A sister-strand exchange mechanism for recA-independent deletion of repeated DNA sequences in *Escherichia coli*. *Genetics*. 1993;135(3):631–42. <https://doi.org/10.1093/genetics/135.3.631> PMID: [8293969](#)
86. DeMarini DM, Lawrence BK. Prophage induction by DNA topoisomerase II poisons and reactive-oxygen species: role of DNA breaks. *Mutat Res*. 1992;267(1):1–17. [https://doi.org/10.1016/0027-5107\(92\)90106-c](https://doi.org/10.1016/0027-5107(92)90106-c) PMID: [1373845](#)
87. Bearson BL, Allen HK, Brunelle BW, Lee IS, Casjens SR, Stanton TB. The agricultural antibiotic carbadox induces phage-mediated gene transfer in salmonella. *Front Microbiol*. 2014;5:52.
88. Stanton TB. Prophage-like gene transfer agents—novel mechanisms of gene exchange for *Methanococcus*, *Desulfovibrio*, *Brachyspira*, and *Rhodobacter* species. *Anaerobe*. 2007;13(2):43–9. <https://doi.org/10.1016/j.anaerobe.2007.03.004> PMID: [17513139](#)
89. Maslowska KH, Makiela-Dzibenska K, Fijalkowska IJ. The SOS system: A complex and tightly regulated response to DNA damage. *Environ Mol Mutagen*. 2019;60(4):368–84. <https://doi.org/10.1002/em.22267> PMID: [30447030](#)
90. Collins LV, Hackett J. Molecular cloning, characterization, and nucleotide sequence of the *rfc* gene, which encodes an O-antigen polymerase of *Salmonella typhimurium*. *J Bacteriol*. 1991;173(8):2521–9. <https://doi.org/10.1128/jb.173.8.2521-2529.1991> PMID: [1707412](#)
91. Guard-Petter J, Parker CT, Asokan K, Carlson RW. Clinical and veterinary isolates of *Salmonella enterica* serovar enteritidis defective in lipopolysaccharide O-chain polymerization. *Appl Environ Microbiol*. 1999;65(5):2195–201. <https://doi.org/10.1128/AEM.65.5.2195-2201.1999> PMID: [10224019](#)