PLOS GENETICS

RESEARCH ARTICLE

Parallel evolution leading to impaired biofilm formation in invasive *Salmonella* strains

Keith D. MacKenzie[1,2], Yejun Wang[3], Patrick Musicha[4,5,6], Elizabeth G. Hansen[1,2], Melissa B. Palmer[1,2], Dakoda J. Herman[1,2], Nicholas A. Feasey[5,7], Aaron P. White[1,2]*

1 Vaccine and Infectious Disease Organization-International Vaccine Centre, Saskatoon, SK., Canada, **2** Department of Microbiology and Immunology, University of Saskatchewan, Saskatoon, SK., Canada, **3** Department of Cell Biology and Genetics, Shenzhen University Health Science Center, Guangdong, China, **4** Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, United Kingdom, **5** Malawi Liverpool Wellcome Trust Clinical Research Programme, University of Malawi College of Medicine, Blantyre, Malawi, **6** Mahidol-Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand, **7** Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

* aaron.white@usask.ca

Citation: MacKenzie KD, Wang Y, Musicha P, Hansen EG, Palmer MB, Herman DJ, et al. (2019) Parallel evolution leading to impaired biofilm formation in invasive *Salmonella* strains. PLoS Genet 15(6): e1008233. https://doi.org/10.1371/journal.pgen.1008233

Editor: Xavier Didelot, University of Warwick, UNITED KINGDOM

Received: February 21, 2019

Accepted: June 7, 2019

Published: June 24, 2019

Data Availability Statement: All raw numerical data and statistical analysis has been deposited in figshare (https://figshare.com/) and is publicly available at: doi: https://doi.org/10.6084/m9.figshare.8220866.v1.

Funding: This work was supported by the Natural Sciences and Engineering Research Council of Canada, http://www.nserc-crsng.gc.ca/ (Grant #2017-05737 to APW; Alexander Graham Bell Canada Graduate Scholarship to KDM; Undergraduate Research Award to EGH), the

Abstract

Pathogenic *Salmonella* strains that cause gastroenteritis are able to colonize and replicate within the intestines of multiple host species. In general, these strains have retained an ability to form the rdar morphotype, a resistant biofilm physiology hypothesized to be important for *Salmonella* transmission. In contrast, *Salmonella* strains that are host-adapted or even host-restricted like *Salmonella enterica* serovar Typhi, tend to cause systemic infections and have lost the ability to form the rdar morphotype. Here, we investigated the rdar morphotype and CsgD-regulated biofilm formation in two non-typhoidal *Salmonella* (NTS) strains that caused invasive disease in Malawian children, *S*. Typhimurium D23580 and *S*. Enteritidis D7795, and compared them to a panel of NTS strains associated with gastroenteritis, as well as *S*. Typhi strains. Sequence comparisons combined with luciferase reporter technology identified key SNPs in the promoter region of *csgD* that either shut off biofilm formation completely (D7795) or reduced transcription of this key biofilm regulator (D23580). Phylogenetic analysis showed that these SNPs are conserved throughout the African clades of invasive isolates, dating as far back as 80 years ago. *S*. Typhi isolates were negative for the rdar morphotype due to truncation of eight amino acids from the C-terminus of CsgD. We present new evidence in support of parallel evolution between lineages of nontyphoidal *Salmonella* associated with invasive disease in Africa and the archetypal host-restricted invasive serovar; *S*. Typhi. We hypothesize that the African invasive isolates are becoming human-adapted and 'niche specialized' with less reliance on environmental survival, as compared to gastroenteritis-causing isolates.

Jarislowsky Chair in Biotechnology (APW), Saskatchewan Health Research Foundation (3866 to KDM); the University of Saskatchewan (Integrated Training Program in Infectious Diseases to MBP; Biomedical Research Award to DJH). 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

African clades of nontyphoidal *Salmonella* cause invasive disease on a daily basis and thousands of deaths each year. Although it is generally accepted that the transmission route for these organisms is fecal-oral, we know very little about their behaviour in the environment between hosts. In this paper, we have identified both a genotype and a phenotype that suggest environmental niche specialization that is distinct from lineages of *Salmonella* Typhimurium and *Salmonella* Enteritidis associated with industrialized food supply chains in resource-rich settings. We also compared with strains of *Salmonella* Typhi, which cause systemic typhoid fever infections exclusively in humans. In each invasive lineage, regulatory or structural gene mutations leading to loss or impairment of biofilm were identified, all associated with curli and cellulose production, the two main structures that comprise the biofilm matrix. This suggests that similar evolutionary pressures are acting on invasive *Salmonella* isolates. Public health strategies aimed at reducing the burden of invasive *Salmonella* disease must prevent transmission to vulnerable adults and children via water sanitation and hygiene practices–a process that starts with identification of environmental reservoirs. The results of our study will raise the profile of this neglected aspect of invasive salmonellosis and will challenge researchers and clinicians to search in new places for potential environmental reservoirs of these pathogens.

Introduction

The 2,600 serovars of the genus *Salmonella* have considerable genetic diversity, which permits them to occupy a wide variety of environmental and animal niches and to cause clinical presentation in humans ranging from asymptomatic carriage through enterocolitis and invasive disease. Most cases of human disease are caused by a few serovars of *Salmonella enterica*, which are loosely categorized as being invasive/typhoidal (serovars Typhi and Paratyphi A) or nontyphoidal. The nontyphoidal salmonellae (NTS) typically cause self-limiting enterocolitis and include common serovars such as *Salmonella* Typhimurium and *Salmonella* Enteritidis [1,2].

This simple clinical distinction breaks down in settings where there is high prevalence of immunosuppressive illness, such as sub-Saharan Africa (sSA). Here, NTS have emerged as a leading cause of bacterial bloodstream infection [3], or invasive nontyphoidal *Salmonella* (iNTS) disease. In common with typhoid fever, iNTS disease frequently presents without diarrheal symptoms, with non-focal febrile illness being the dominant clinical presentation [4]. This disease is responsible for an estimated 681,000 deaths per year, with nonspecific symptomology, multidrug resistance, and poor clinical outcomes despite correct diagnosis contributing to this significant mortality rate [5]. There is great urgency to better understand iNTS disease and reduce its impact in Africa and other areas of the world [6].

Most NTS infect a wide range of host species and are considered host-generalist pathogens [7]. In contrast, typhoidal serovars are exemplars of host-restricted pathogens, with marked genomic degradation in comparison to NTS serovars [8]. There is evidence that the same process is underway in strains of *S*. Typhimurium and *S*. Enteritidis associated with invasive disease in sub-Saharan Africa [9,10]. Many of these gene mutations affect metabolic processes involved in anaerobic respiration on unique carbon sources, a mechanism that is pivotal for the replication and outgrowth of *Salmonella* in the inflamed intestinal tract [11,12]. Whether this represents a random process related to the different geographical location of these strains,

convergent evolution with the typhoidal serovars towards an invasive rather than an enteric "lifestyle", or adaptation to distinct environmental niches remains an outstanding question.

Irrespective of invasive versus enteric lifestyle, all salmonellae are transmitted via the fecal-oral route, but key questions remain over how invasive strains interact with the environment between host colonization events [13,14]. Biofilm formation is proposed to aid in the survival and persistence of *Salmonella* cells during this environmental phase of the transmission cycle [15]. The most well-characterized format of *Salmonella* biofilm physiology has been termed the rdar (red, dry, and rough) morphotype, where a self-produced extracellular matrix interconnects cells and facilitates their adherence to abiotic and biotic surfaces [16–18]. Multiple cues, including ambient temperature, osmolarity, and nutrient availability, act via a complex regulatory network to activate CsgD, a member of the RpoS regulon and the primary transcriptional activator of the rdar morphotype [15,16]. CsgD in turn induces the expression of proteinaceous (curli fimbriae [19]), BapA [20]) and polysaccharide (cellulose [17,21], O-antigen capsule [18,22]) polymers that act as major contributors to the recalcitrant matrix structure associated with this phenotype. The genes for curli fimbriae and cellulose are highly conserved in *Salmonella* [23–25]; however, almost all Typhi and Paratyphi isolates are phenotypically negative for the rdar morphotype [23,26]. Therefore, loss of the rdar morphotype may represent an additional signature of host adaptation.

Landmark studies revealed genomic degradation within representative iNTS strains *S*. Typhimurium D23580 [27] and *S*. Enteritidis D7795 [10]–these observations were consistent with genetic signatures of host adaptation found in typhoidal and paratyphoidal *Salmonella* species [28,29]. As such, significant attention has focused on evaluating the virulence of these strains in laboratory models of *Salmonella* infection [9,30–32]. Comparatively less is known about the state of biofilm physiology of iNTS strains [33], and in particular the CsgD-regulated rdar morphotype, which we and other labs have postulated may be important for the persistence of nontyphoidal *Salmonella* species in non-host environments [15,23,25,26,34].

We wanted to determine if *S*. Typhimurium D23580 and *S*. Enteritidis D7795 can form biofilms, and if not, to determine if the corresponding mutations are conserved in other African strains associated with invasive disease. We selected typical Typhimurium and Enteritidis strains that cause enterocolitis to act as controls in our phenotypic screens and to rule out any serovar-specific genetic variations that do not affect biofilm formation. We included the archetypal Typhi strain (CT18) and modern versions of the H58 haplotype [35] to determine if patterns of gene loss or mutation are shared between invasive isolates.

Results

Evaluation of the biofilm phenotype for invasive nontyphoidal and typhoidal *Salmonella* strains

The rdar morphotype was first described by Römling and colleagues during their research into the expression of curli fimbriae in *S*. Typhimurium 14028 [16]. For the wildtype strain, the morphotype was reported as a temperature-restricted phenotype, with transcription of curli biosynthesis (*csg*) genes and characteristic patterns on the colony surface constrained to cells grown at 28˚C in nutrient-rich medium of low osmolarity. These conditions have since been used as a standard for screening *Salmonella* isolates for the rdar morphotype [26]. After three to five days of growth at 28˚C, nontyphoidal and typhoidal *Salmonella* strains were easily distinguishable based on the different appearances of their macrocolonies (Fig 1). We added Congo red dye to the agar to visualize the presence of curli fimbriae and cellulose and identify strains that possess the red, dry, and rough (rdar) morphotype [19]. Control strains of NTS were rdar-positive and red in colour, while *S*. Typhi strains were smooth and white (saw), or

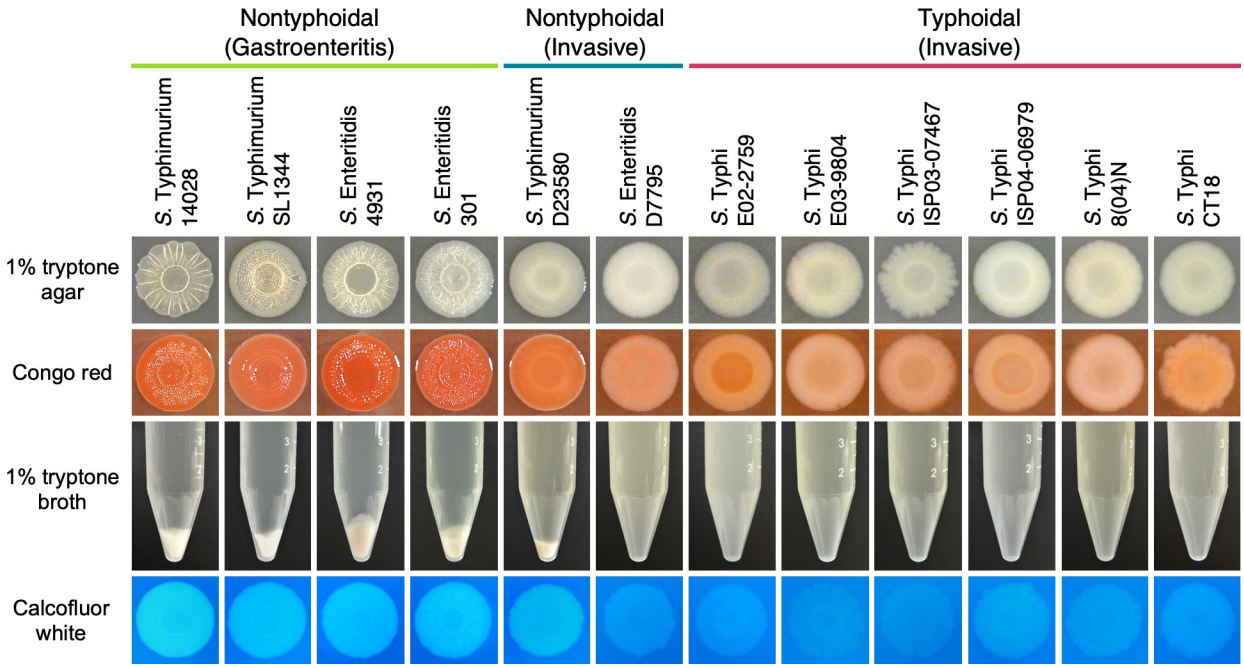

Fig 1. The biofilm phenotypes of representative nontyphoidal and typhoidal *Salmonella* strains. *Salmonella* strains that are known to cause gastroenteritis or enterocolitis (green), invasive disease (blue) or typhoid fever (red) were screened: for the ability to form the red, dry, and rough (rdar) morphotype, which presents as the formation of concentric rings and a wrinkled appearance on the surface of macrocolonies (panels 1 and 2, top); for the presence of multicellular, biofilm aggregates and planktonic cells in liquid cultures grown under biofilm-inducing conditions (middle panel; conical tubes); and for cellulose production, visualized as the white and fluorescent appearance of macrocolonies in the presence of calcofluor white dye (bottom panel).

https://doi.org/10.1371/journal.pgen.1008233.g001

rdar-negative. *S*. Typhimurium D23580 formed pale red colonies with an incomplete wrinkling pattern on the colony's surface, whereas *S*. Enteritidis D7795 was rdar-negative. Temperature-based restriction of curli expression can be overcome in some instances; examples of this include when strains carry certain mutations in the *csgD* promoter [16,36], or when nontyphoidal *Salmonella* are grown in iron-limiting conditions [16] or in the presence of human bile [37]. However, all NTS strains displayed a temperature-based restriction of curli expression [16], as they did not form rdar colonies at 37˚C and curli (*csgBAC*) transcription was at basal levels (S1 Fig).

Rdar morphotype-positive *Salmonella* strains grown in an *in vitro* flask model of biofilm development will produce two unique subpopulations—multicellular, biofilm aggregates and planktonic cells [38,39]—due to the bistable production of CsgD [40]. We reasoned that multicellular aggregation in liquid cultures would provide a clearer diagnostic for the presence of a functioning biofilm phenotype in both strong and moderate biofilm-producing strains. Cultures of *S*. Typhimurium D23580 contained visible multicellular aggregates and planktonic cell subpopulations (Fig 1). However, the aggregates made up a lower proportion of the population biomass compared to aggregates from other NTS and had observable differences in their physical structure (S2 Fig), suggesting a moderate but impaired biofilm phenotype. In contrast, cultures of *S*. Enteritidis D7795 and all typhoidal *Salmonella* strains consisted solely of planktonic cells. Cellulose production can be specifically assessed by supplementing T agar with calcofluor white dye and illuminating colonies with ultraviolet light [25]. Similar fluorescence intensity was observed for *S*. Typhimurium D23580 compared to other NTS, indicating the presence of cellulose production despite the partial rdar morphotype (Fig 1). In contrast, cellulose

production was comparably minimal or absent from macrocolonies of *S*. Enteritidis D7795 and all *S*. Typhi isolates.

Status of CsgD and curli production and RpoS synthesis and activity

The divergent *csg* operons are central to curli biosynthesis and include genes for curli fimbrial protein subunits (*csgBA*), transcriptional regulation (*csgD*), and curli assembly machinery (*csgC* and *csgEFG*) [41–43]. CsgD is involved in activating the transcription of both operons [16,19], making detection of curli fimbrial subunits a strong indicator of CsgD activity. We probed the lysates of cell subpopulations emerging from flask cultures for synthesis of CsgD and CsgA, the major subunit of curli fimbriae (Fig 2). We detected both proteins in lysates from multicellular aggregates, including *S*. Typhimurium D23580. In contrast, neither protein was detected for samples derived from *S*. Enteritidis D7795 or any of the *S*. Typhi strains.

We also probed the lysates for RpoS, the sigma factor that controls *csgD* expression [19] (S3A Fig). RpoS protein was detected in all nontyphoidal *Salmonella* strains, including both biofilm and planktonic cell subpopulations derived from biofilm-producing strains. *S*. Enteritidis D7795 appeared to have lower levels of RpoS compared to the other NTS strains. For the *S*. Typhi strains, there was significant variation in protein levels at the 24-hour time point used for sampling, with low levels of RpoS observed for *S*. Ty E03-9804 and *S*. Ty 8(04)N. Bacterial luciferase reporter technology allows for the systematic comparison of transcriptional activity across a wide variety of strains [24,44]. To evaluate RpoS activity, we tracked the expression of a synthetic, RpoS-dependent promoter-luciferase reporter construct in each strain during growth in microbroth cultures [15] (S3B–S3D Fig). All strains in our panel exceeded the standard RpoS activity threshold identified in our previous work with other *Salmonella* strains (i.e. 10,000 CPS) [24], except for *S*. Enteritidis D7795.

Screening for serovar- and strain-specific *cis* or *trans* variation in biofilm gene regulation

We hypothesized that variation in *csgD* expression levels could account for differences in the biofilm phenotypes we observed. The sequence of the 755-bp intergenic region between the *csgBAC* and *csgDEFG* operons was compared between each strain in our panel (Fig 3A). For *S*. Enteritidis D7795 and *S*. Typhimurium D23580, unique single nucleotide polymorphisms (SNPs) were identified within the intergenic region (Fig 3B). For *S*. Enteritidis D7795, a C-to-T SNP was observed in the regulatory region 47 bp upstream of the *csgD* transcriptional start site. For *S*. Typhimurium D23580, independent C-to-A and G-to-A transversion mutations were identified 80 bp and 189 bp upstream of the *csgD* transcriptional start site.

To assess the effects of sequence or '*cis*' differences in the *csg* intergenic region, we generated transcriptional reporters for the *csgD* and *csgB* promoters of all twelve strains. The reporters were transformed into *S*. Typhimurium 14028 to ensure that promoter activity was measured in a consistent cellular environment. Maximum expression levels were similar for all *csgD* or all *csgB* promoter-reporter constructs generated from NTS control strains and *S*. Typhi strains, suggesting that serovar-specific polymorphisms did not have a significant impact on promoter functionality (Fig 3C). We noted slightly lower activity from the *csgB* promoter from *S*. Typhimurium SL1344, but considered the promoter functional based on its overall expression profile (S4 Fig). In contrast, the activity of *csgD* promoters from *S*. Typhimurium D23580 and *S*. Enteritidis D7795 were significantly lower than native *S*. Typhimurium 14028 promoters (Fig 3C). *S*. Typhimurium D23580 had peak expression that was approximately three-fold lower than *S*. Typhimurium 14028, whereas *S*. Enteritidis D7795 had near background levels of expression (~1,000 counts per second) and appeared to be non-

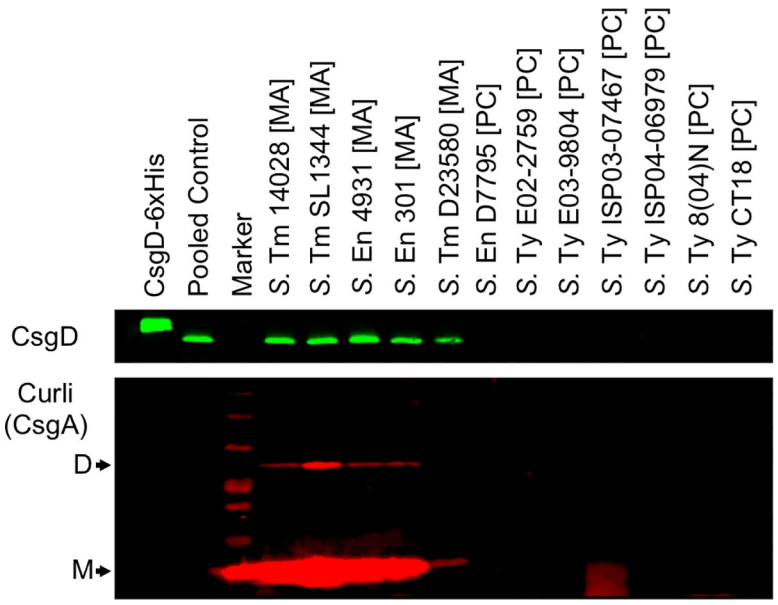

Fig 2. Detection of CsgD and CsgA protein synthesis in representative *Salmonella* strains by Western blot. Whole-cell lysates were derived from multicellular aggregates [MA] or planktonic cells [PC] harvested from flask cultures of nontyphoidal and typhoidal *Salmonella*. Lysates used for CsgD detection (top panel) were normalized based on total protein concentration. Purified CsgD-6xHis recombinant protein was used as a technical control for CsgD detection. Pooled control samples were derived from combining lysates obtained from the multicellular aggregates of *S*. Typhimurium 14028 and SL1344 and *S*. Enteritidis 4931 and 301 strains. Black arrows indicate the detection of CsgA subunit dimers (D) and monomers (M) (bottom panel). The data shown is representative of two biological replicates.

https://doi.org/10.1371/journal.pgen.1008233.g002

functional. In contrast, there was minimal difference between the activities of *csgB* promoters from the African NTS strains and the panel of control strains.

To determine if strains harboured mutations in *trans*-acting regulatory elements, each strain was transformed with a set of biofilm-associated transcriptional reporters derived from *S*. Typhimurium 14028. The *csgD* promoter was expressed in all *Salmonella* strains, though its activity was six-fold lower in *S*. Enteritidis D7795, *S*. Ty E02-2759, and *S*. Ty E03-9804 (Fig 3D). This result suggested that mutations in the regulatory network upstream of *csgD* could be partially responsible for loss of the rdar morphotype in these three strains. However, sequence analysis of the promoters and genes of six common regulators of *csgD* transcription did not reveal any unique sequence changes in these strains (i.e., positive regulators OmpR, MlrA, IHF, RstA; negative regulators CpxR, H-NS [45,46]). Expression of the *csgB* promoter showed a clearer distinction, as all control NTS strains and *S*. Typhimurium D23580 had significantly greater *csgB* promoter expression than rdar-negative *S*. Enteritidis D7795 and the *S*. Typhi strains (Fig 3D). A similar pattern was observed for the promoters of *adrA*, a gene which regulates cellulose production, and *cpxRA*, which is part of a two-component system that responds to envelope stress (S5 Fig). These *trans* patterns of gene expression correlated with a lack of CsgD in *S*. Enteritidis D7795 and *S*. Typhi strains.

A single, non-coding nucleotide shuts off the rdar morphotype for *Salmonella* serovar Enteritidis D7795. *S*. Enteritidis D7795 possessed a unique C-to-T transition in the *csgD* promoter region and also had a non-functional *csgD* promoter. We performed genome engineering to replace this SNP with corresponding sequence from rdar-positive *S*. Typhimurium 14028 (see Materials and Methods). We were able to track the presence of the SNP by digestion of the *csgD* promoter region with the restriction endonuclease, *Psi*I. Strains containing the SNP (i.e., TTATAA) had a unique 148 bp fragment compared to strains with the SNP corrected

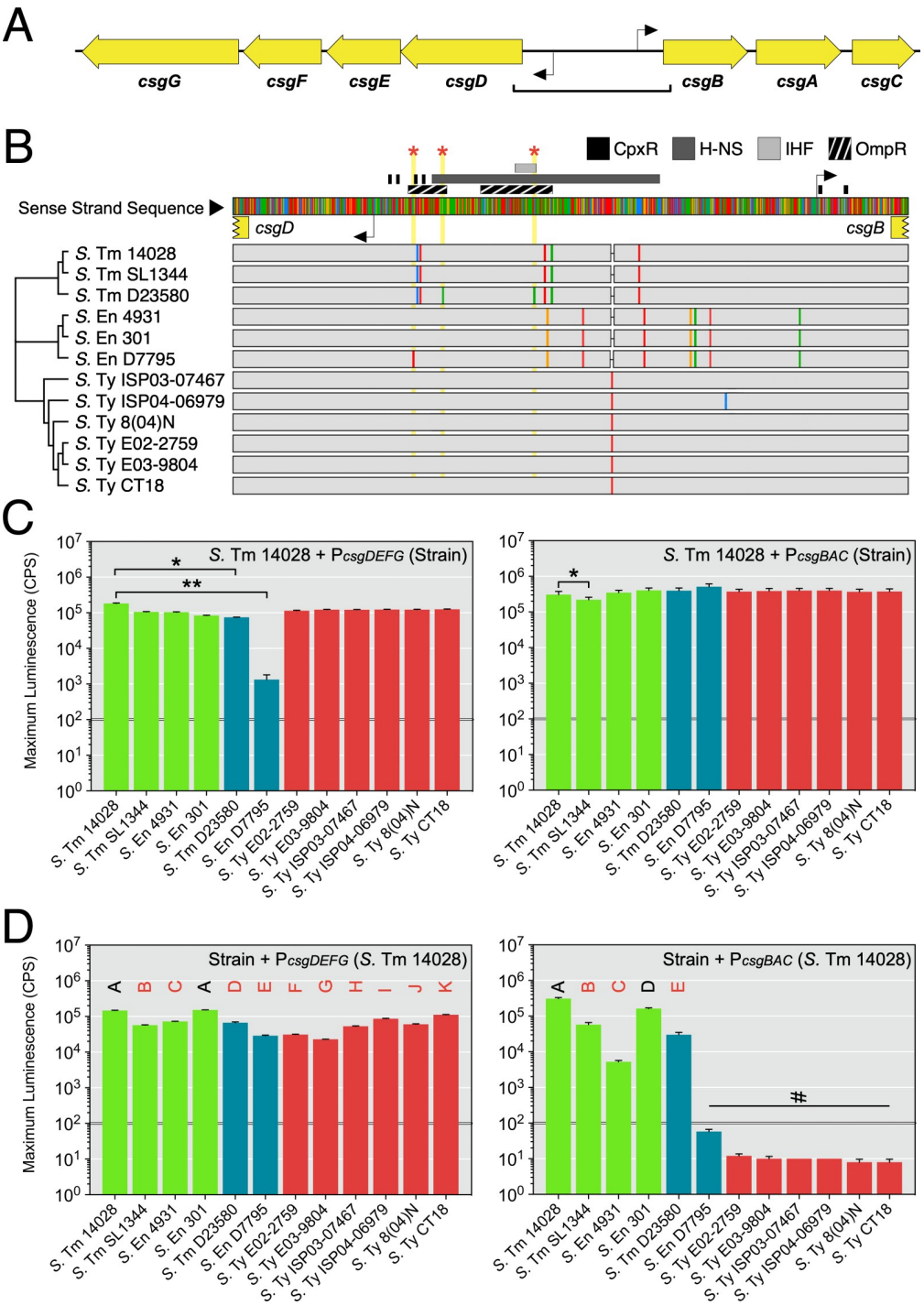

Fig 3. Evaluation of the *csgDEFG-csgBAC* intergenic region for changes in promoter sequence and activity. (A) Diagram representing the divergent *csg* operons. Transcriptional start sites are indicated as black elbow arrows. The square bracket indicates the region analyzed in (B). (B) Multiple sequence alignment of the intergenic and 5′ untranslated regions for the *csgDEFG* and *csgBAC* operons from *Salmonella* strains in this study. A neighbour-joining dendrogram was established based on bootstrapping parameters set to 1,000 replicates and a support threshold of 70%. Transcription factor binding sites that have been experimentally verified in *Salmonella* are indicated above the consensus sequence. Single nucleotide polymorphisms (SNPs) within each strain's DNA sequence are indicated as coloured rectangles within grey tracks (red, adenine; blue, cytosine; yellow, guanine; green, thymine). Red stars and yellow vertical blocks highlight SNP positions unique to *S.* Typhimurium D23580 and *S.* Enteritidis D7795. (C) Promoter-reporter fusion constructs were generated using *csgD* and *csgB* promoter sequences derived from the

Salmonella strains. The activity of each construct was evaluated in *S*. Typhimurium 14028 cells during 48 hours of growth. Graphed values represent the maximum reporter activity recorded in this period and is reported as counts per second (CPS). Statistical significance: *, $P < 0.05$; **, $P < 0.01$. (D) Constructs derived from *S*. Typhimurium 14028 *csgD* and *csgB* promoter sequences were introduced into each *Salmonella* strain. Letters above the bars indicate mean values that were statistically similar to (black font) or different from (red font) other mean values. #, values below the activity threshold as established in [24]. Each bar represents the mean value from three to five biological replicates. Error bars represent standard deviations.

https://doi.org/10.1371/journal.pgen.1008233.g003

(i.e., TTACAA) (Fig 4A). Strains of *S*. Enteritidis D7795 with the SNP corrected (-47T>C) had their biofilm phenotypes restored, as judged by the formation of rdar colonies on TCR agar and cellulose production on agar supplemented with calcofluor white (Fig 4B), in addition to the formation of biofilm aggregates when grown in liquid culture (S6 Fig). In contrast, *S*. Enteritidis D7795 strains that had retained the SNP after genome engineering were negative for the rdar morphotype and multicellular aggregate formation. As another test of the importance of this promoter SNP, we introduced the change into *S*. Typhimurium 14028 (-47C>T), which resulted in the loss of rdar colony morphology and cellulose production (Fig 4B).

To determine if the phenotypic changes observed were directly linked to changes in *csgD* promoter activity, we generated promoter-reporter constructs from each engineered strain. Presence of the 'T' SNP at position -47 correlated with functional inactivation of the *csgD* promoter, shown as a 48-hour time course for *S*. Enteritidis D7795 strains (Fig 4C) and a statistically significant drop in maximum promoter expression for the *S*. Typhimurium 14028 strains (Fig 4D). Together, these experiments provided evidence that a single polymorphism in a non-coding region of the *Salmonella* genome was capable of shutting off CsgD-regulated biofilm phenotypes.

Unique SNPs reduce *csgD* promoter activity in invasive NTS strain *Salmonella* serovar Typhimurium D23580

S. Typhimurium D23580 possessed two unique SNPs in the *csgD* promoter region and the promoter had reduced activity when compared to *S*. Typhimurium 14028. We used genome engineering to replace the SNP at the -80 bp position (D23580 -80A>C), which boosted *csgD* promoter activity approximately two-fold as compared to the D23580 parent strain (Fig 5A). Replacement of the SNP at -189 bp position (D23580 -80A>C, -189A>G) resulted in an additional boost in promoter activity to reach similar levels as the *csgD* promoter from wildtype *S*. Typhimurium 14028 (Fig 5A). Despite the increase in promoter activity, replacement of both *csgD* promoter SNPs (*S*. Typhimurium D23580-P*csgD* -80A>C, -189A>G) was unable to restore the rdar morphotype, and although the strain appeared to bind more Congo red than the parent *S*. Typhimurium D23580 strain, the difference was hard to quantitate (Fig 5B).

S. Typhimurium D23580 was first characterized as being rdar-negative or -intermediate due to the presence of a premature stop codon in *bcsG*, and the rdar morphotype was restored by over-expressing *bcsG* [47]. To evaluate the impact of the *bcsG* mutation, we introduced this SNP into the rdar-positive *S*. Typhimurium 14028 strain, which caused a loss of pattern formation on the colony surface and a drop in calcofluor binding intensity (Fig 6B). We concluded from this data that *S*. Typhimurium D23580 possessed multiple mutations that influence curli and cellulose production, leading to a reduced biofilm phenotype.

Conservation of biofilm-altering SNPs in invasive lineages of *S*. Typhimurium and *S*. Enteritidis

We performed *in silico* screening of additional serovar Enteritidis and Typhimurium strains isolated from Africa and other areas of the world [10,48] to determine how widespread the

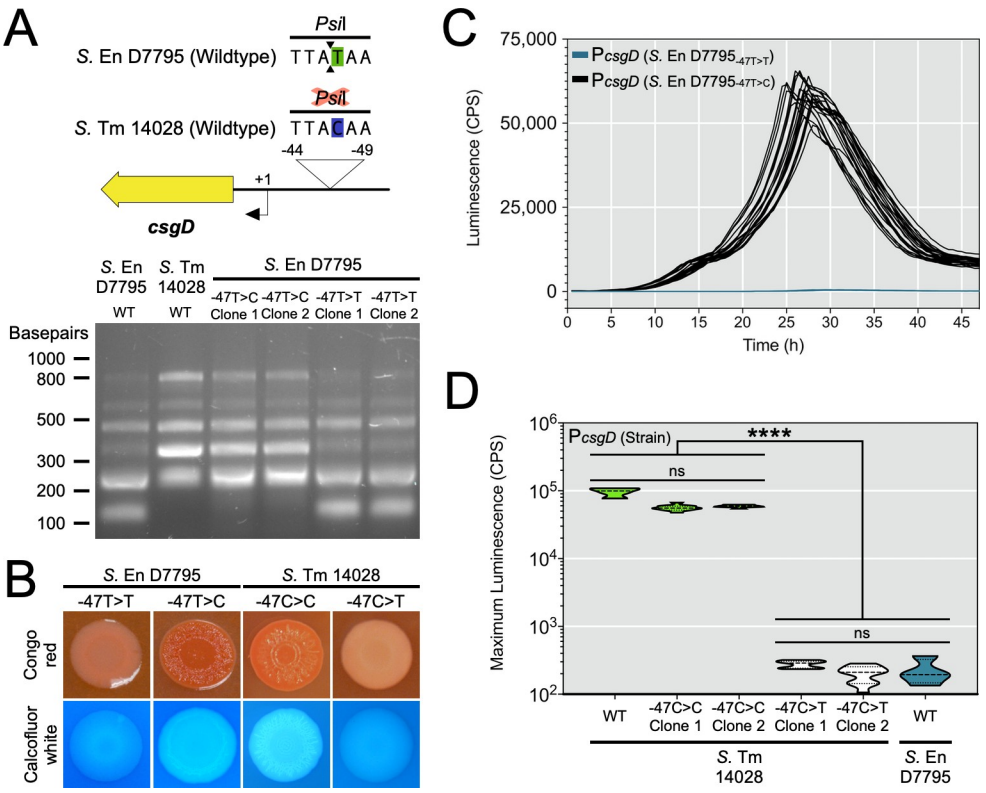

Fig 4. Biofilm phenotypes and *csgD* promoter activities following chromosomal replacement of strain-specific *csgD* promoter sequences in *S.* Enteritidis D7795 and *S.* Typhimurium 14028. (A) Genome engineering was used to replace part of the native *csgD* promoter sequence in *S.* Enteritidis D7795 with sequence from *S.* Typhimurium 14028. The same process was also used to replace part of the native sequence in *S.* Typhimurium 14028 cells with sequence from *S.* Enteritidis D7795. The 780-bp *csgD* promoter region was PCR amplified from the strains and clones listed, followed by digestion with *Psi*I, which has a recognition site overlapping the *S.* Enteritidis D7795 SNP-containing region. (B) Macrocolony phenotypes of *S.* Enteritidis D7795 and *S.* Typhimurium 14028 clones that either contain or do not contain the identified 'T' promoter SNP at position -47. (C) The *csgD* promoter from each *S.* Enteritidis D7795 clone was used to generate promoter luciferase reporters that were transformed into *S.* Typhimurium 14028 cells. Each line represents one biological replicate culture ($n = 22$) with measured promoter activity (CPS, counts per second) plotted versus time. (D) Maximum reporter activity recorded from *csgD* promoter luciferase constructs derived from *S.* Typhimurium 14028 clones and transformed into wildtype *S.* Typhimurium 14028 cells ($n = 11$ per reporter construct). The activity of wildtype (WT) *S.* Typhimurium 14028 and *S.* Enteritidis D7795 *csgD* promoter-reporter constructs were included in the assay ($n = 4$ per each construct). Violin plots show the frequency distribution of the data, with the dotted line representing the median value. ****, $P < 0.0001$.

https://doi.org/10.1371/journal.pgen.1008233.g004

biofilm-altering SNPs were. For *S.* Enteritidis, all 167 strains in the central/east African clade (i.e., HierBAPs-predicted clade from [10]), which includes D7795, possessed the inactivating *csgD* promoter SNP (Fig 6A). In contrast, 100% of the 250 isolates from the HierBAPs-predicted global clade associated with human enterocolitis and/or poultry farming did not have this SNP (Fig 6A; S3 Table). For *S.* Typhimurium, the *csgD* promoter mutation at position -189 was conserved in all 50 lineage I and 71 lineage II strains analyzed (Fig 6B; S4 Table), that were isolated from people in sSA within the past 30 years [48]. Lineage II, which includes strain D23580, is thought to have arisen from lineage I due to increased selection pressure from heavy antibiotic use [48]. Therefore, we predicted that lineage II isolates would carry some unique mutations. Consistent with this, all lineage II isolates possessed the *csgD* promoter SNP at position -80 as well as the premature stop codon in *bcsG*. Importantly, none of

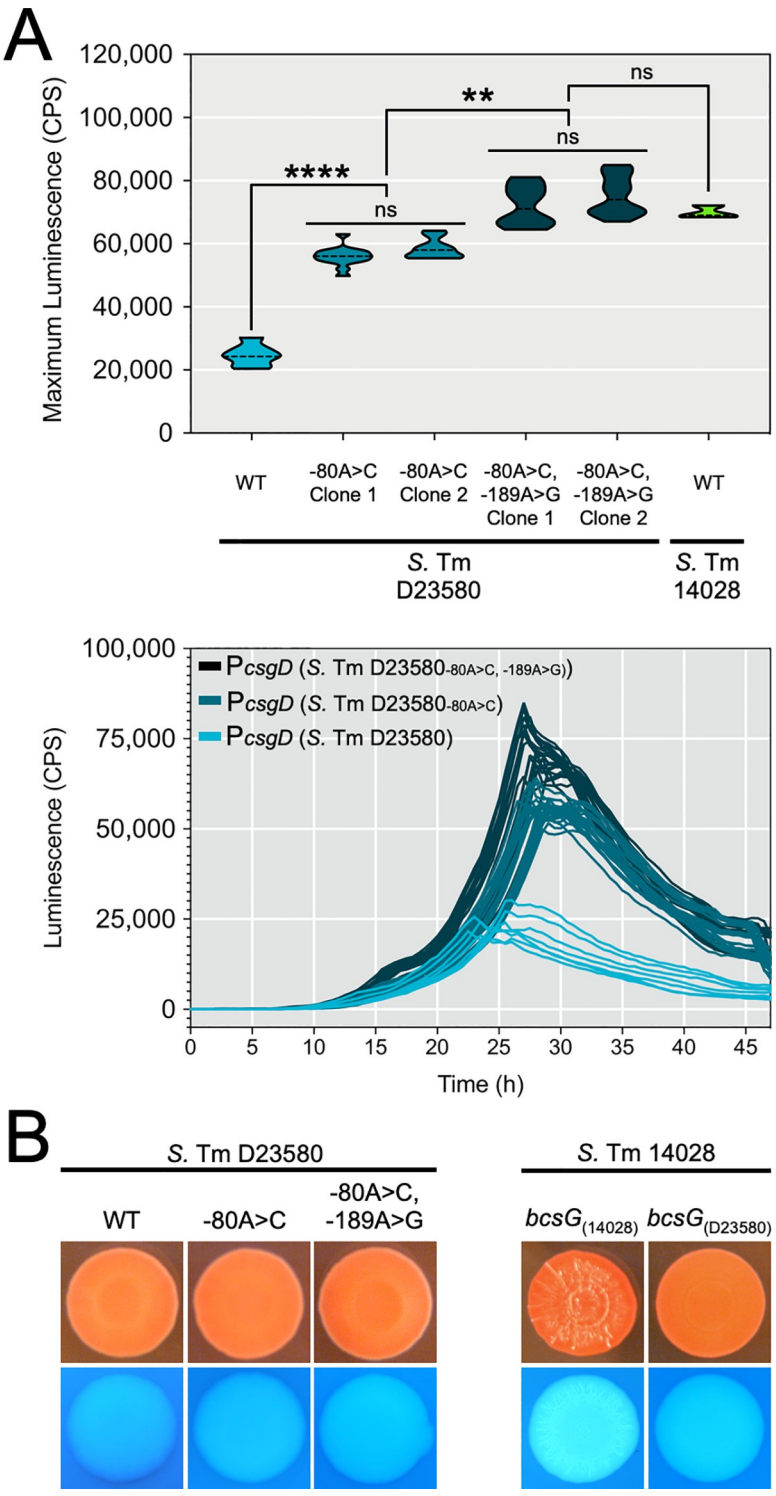

Fig 5. Biofilm phenotypes and *csgD* promoter activities following chromosomal replacement of strain-specific *csgD* promoter sequences or *bcsG* mutations in *S.* Typhimurium D23580 or *S.* Typhimurium 14028. (A) Top panel: Maximum activity (CPS; counts per second) was recorded for *csgD* promoter luciferase reporters derived from *S.* Typhimurium D23580 clones generated by genome engineering that either contained or did not contain the 'A' SNP at position -80 and 'A' SNP at position -189 relative to the *csgD* transcriptional start site. Reporters were transformed into *S.* Typhimurium 14028 and activity monitored during 48 hours of growth ($n = 28$ per reporter construct). Wildtype (WT) reporters from *S.* Typhimurium D23580 ($n = 24$) and *S.* Typhimurium 14028 ($n = 12$) were included as controls.

Violin plots show the frequency distribution of the data, with the dotted line representing the median value; **, $P < 0.01$; ****, $P < 0.0001$. Bottom panel: 48 h time-course expression profiles for *csgD* promoter luciferase reporters analyzed in the top panel, except for *S*. Typhimurium 14028; each line represents one biological replicate culture. (B) Macrocolony phenotypes of *S*. Typhimurium D23580 and *S*. Typhimurium 14028 clones containing native or replacement *csgD* promoter sequences (left panel) or *bcsG* alleles (right panel).

https://doi.org/10.1371/journal.pgen.1008233.g005

the 63 *S*. Typhimurium strains of other sequence types (i.e., ST-19, -34, -98, -128 or -568) possessed the biofilm-altering mutations (Fig 6B; S4 Table).

Loss of the rdar morphotype and multicellular aggregation in *Salmonella* serovar Typhi strains is due to truncated CsgD

Salmonella Typhi represents the best characterized group of strains that cause a more invasive disease than enterocolitis. All six *S*. Typhi strains that we analyzed had functional *csgD* promoters, and at least four strains appeared to have upstream regulatory components intact (Fig 3). In the *S*. Typhi CT18 genome sequence, Parkhill *et al*. identified a premature stop codon at the 3' end of *csgD*, which would result in truncation of eight amino acids from the C-terminal end of CsgD [49]. To analyze the functionality of this truncated *csgD* allele (i.e., *csgD*^CT18)

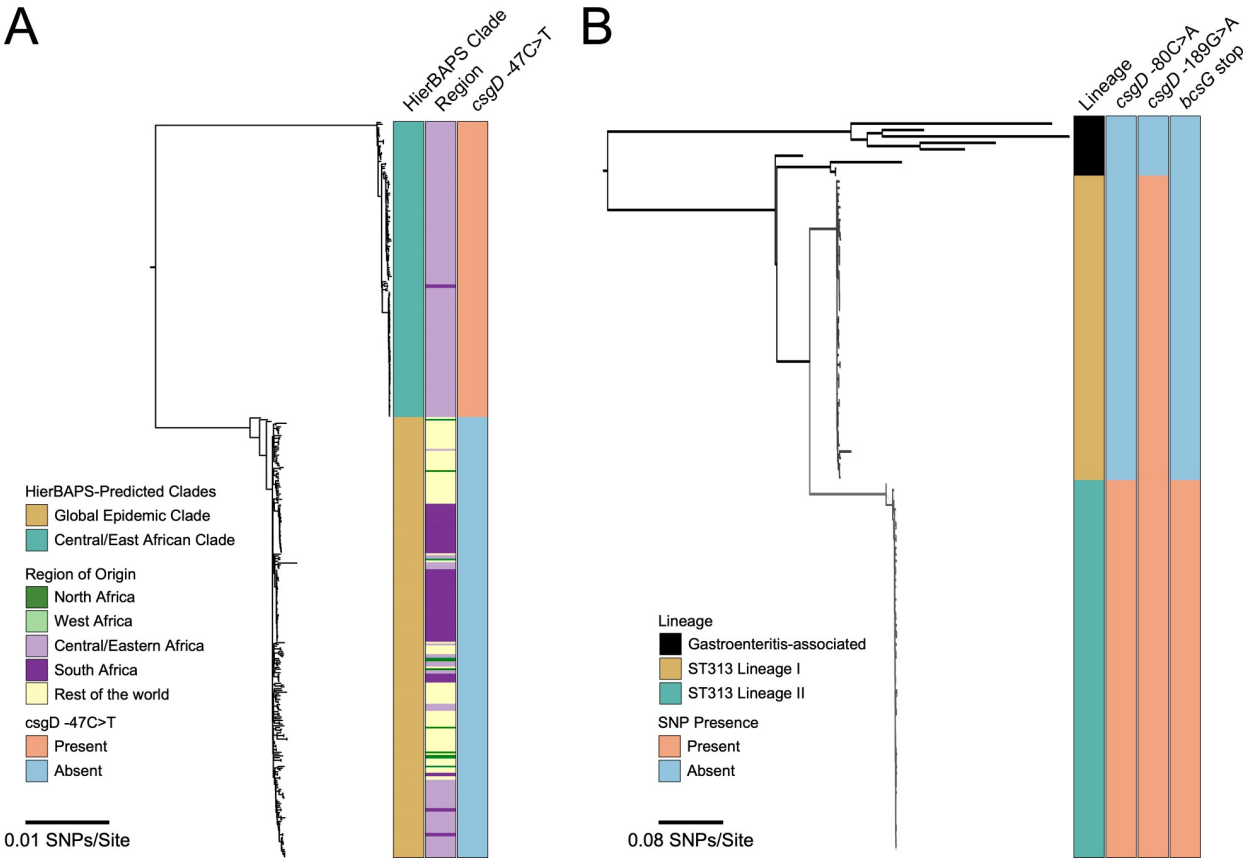

Fig 6. Conservation of *csgD* promoter and *bcsG* single nucleotide polymorphisms in invasive *S*. Enteritidis and *S*. Typhimurium lineages. Maximum likelihood phylogenic trees were constructed from bacterial genome sequences: (A) *S*. Enteritidis isolates from Feasey et al. [10], and (B) *S*. Typhimurium isolates from Okoro et al [48], keeping the same general tree shape for comparison purposes. (A) *S*. Enteritidis isolates were divided into the Central/East African clade (167 isolates) and global epidemic clade (250 isolates), with the distinct region of isolation shown along with presence or absence of the 'T' SNP. (B) *S*. Typhimurium isolates were divided into gastroenteritis-associated and invasive lineages (I and II), with the presence or absence of *csgD* promoter and *bcsG* polymorphisms shown.

https://doi.org/10.1371/journal.pgen.1008233.g006

compared to full-length *csgD* (i.e., *csgD*^14028), we performed complementation experiments in a *S*. Typhimurium 14028 Δ*csgD* strain background. The presence of *csgD*^14028 in a multi-copy plasmid resulted in a strain with characteristic rdar morphotype colonies, whereas the p3xFLAG control strain was smooth and white (Fig 7A). The presence of *csgD*^CT18 resulted in a strain that formed red colonies with rdar-intermediate morphology, indicating that the truncated CsgD protein was partially functional. When grown in liquid culture, the *csgD*^CT18 and p3xFLAG cultures were devoid of multicellular aggregates, whereas the *csgD*^14028 complemented culture produced both cell types (Fig 7A). We detected robust CsgD-3xFLAG synthesis in the *csgD*^14028 culture, whereas only a faint CsgD band could be detected in the *csgD*^CT18 culture (Fig 7B). When expression of *csgD*^CT18 was induced by addition of IPTG, it caused a boost in curli (*csgBAC*) gene promoter expression (Fig 7B) and aggregates were formed in liquid culture (Fig 7C). Expression of the *csgB* promoter in the induced *csgD*^CT18 cultures reached levels similar to, but still below the levels in the *csgD*^14028 uninduced cultures (Fig 6C). We were unable to measure *csgB* promoter activity in induced *csgD*^14028 cultures because growth ceased upon addition of IPTG. Together, these experiments demonstrated that truncated CsgD from *S*. Typhi CT18 had reduced functionality compared to full-length CsgD, but was able to restore both rdar and other CsgD-regulated biofilm phenotypes if expressed at higher levels. As a final test of functionality, we used genome engineering to introduce the premature stop codon in *csgD* into *S*. Typhimurium 14028, resulting in a strain that was rdar-negative with minimal Congo red binding and reduced cellulose production (Fig 7D). This experiment showed that the SNP at the 3' end of *csgD* was sufficient to disrupt CsgD-regulated biofilm phenotypes in *S*. Ty CT18.

We wanted to determine if there were other mutations that could potentially affect curli or cellulose production within our panel of twelve *Salmonella* strains. For curli, we performed sequence alignment of the entire 4,450 bp region containing the divergent *csgBAC* and *csgDEFG* operons (S7 Fig). Seventeen unique SNPs were identified in each of the six *S*. Typhi strains, however the only clear nonsynonymous mutation was the premature stop codon in *csgD* that was previously described [49]. Several serovar Typhimurium- and Enteritidis-specific SNPs were identified in the intergenic region and in the *csg* coding regions, however since the majority of these SNPs were found in biofilm-positive strains, we concluded that they likely do not pose a significant effect on *csg* function. For cellulose, we performed sequence alignment of the 14,273 bp region containing the divergent *bcsRQABZC* and *bcsEFG* operons (S8 Fig). Overall, this DNA region was less conserved than the *csg* region. All six *S*. Typhi strains shared numerous SNPs, including 21 non-synonymous changes in the *bcs* coding regions, and four premature stop codons in *bcsC*. *S*. Enteritidis D7795 had one unique SNP in *bcsC*, which resulted in a non-synonymous change that was shared with *S*. Typhi strains, and *S*. Typhimurium D23580 had one unique SNP leading to a premature stop codon in *bcsG*, as previously described [47].

Uniqueness of the identified SNPs influencing rdar morphotype biofilm formation in *Salmonella*

We wanted to determine if the SNPs identified in *S*. Enteritidis D7795, *S*. Typhimurium D23580 and S. Typhi were unique to these lineages or could be detected in isolates from other *S. enterica* subspecies *enterica* serovars. We analyzed the genomes of 248 isolates from 55 serovars, including representatives of the most common serovars associated with human disease [50] and host-adapted serovars such as Dublin, Choleraesuis and Gallinarum (S5 Table). The -47 C>T mutation of *S*. Enteritidis D7795 was found in one strain of serovar Weltevreden, one strain of serovar Anatum had a 110-bp deletion in the *csgD* promoter region comprising both the -47 C>T and -80 G>T mutations, and the -189 C>T mutation of *S*. Typhimurium

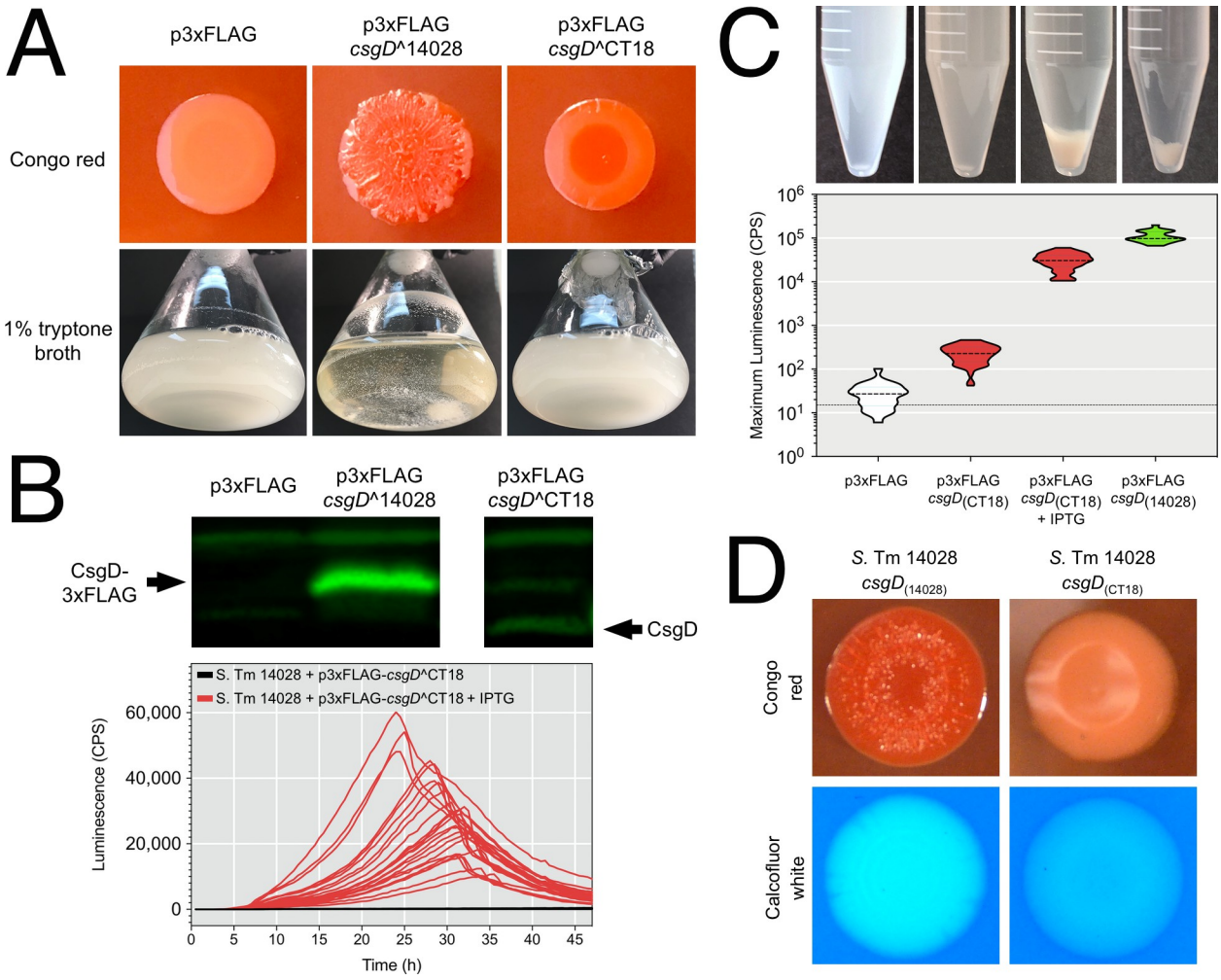

Fig 7. Functional analysis of the *S.* Typhi CT18 *csgD* allele. (A) The *csgD* alleles from *S.* Typhimurium 14028 and *S.* Typhi CT18 were cloned into p3xFLAG and transformed into *S.* Typhimurium 14028 Δ*csgD* cells. Colony morphology and flask cultures of uninduced cells were evaluated for biofilm phenotypes. (B) Top panel: Whole cell lysates were generated from cells acquired from flask cultures and probed for synthesis of CsgD via Western blot. Bottom panel: A *csgB* promoter-reporter construct was used to evaluate CsgD activity in *S.* Typhimurium 14028 Δ*csgD* cells harbouring p3xFLAG-*csgD*^CT18 with or without IPTG induction. (C) Top panel: Presence or absence of multicellular aggregates in flask cultures of *S.* Typhimurium 14028 Δ*csgD* cells containing p3xFLAG constructs (+/- IPTG) or p3xFLAG alone. Bottom panel: Maximum promoter activity of a *csgB* promoter-reporter construct measured in microbroth cultures corresponding to the tubes shown in the top panel. (D) Biofilm phenotypes of *S.* Typhimurium 14028 strains that contained the native *csgD*^14028 allele or the truncated csgD allele from *S.* Typhi CT18 following genome engineering.

https://doi.org/10.1371/journal.pgen.1008233.g007

D23580 was found in one strain of serovar Hillingdon and two strains of serovar Typhimurium (S5 Table). The SNP leading to a premature stop codon in *csgD* was unique to Typhi strains, and the G>A mutation causing a premature stop codon in *bcsG* was not detected in any strains (S5 Table). We also screened 82 West African serovar Enteritidis strains associated with invasive human disease [10] but none of the mutations were detected (S6 Table). With just a few exceptions, the SNPs identified in our biofilm screening appeared to be unique to the invasive lineages of *Salmonella* where they were originally detected.

Discussion

In this study, we identified critical changes that disrupt or reduce the rdar morphotype and other CsgD-regulated biofilm phenotypes in three invasive *Salmonella* lineages. For *S.*

Enteritidis D7795, a single promoter mutation was responsible for inactivating the *csgD* promoter and shutting off the rdar morphotype. The polymorphism was in one of the two most conserved bases in the OmpR recognition site (ACNTTTNGNTA'**C**'ANNTAT; [51]). This is predicted to knock out OmpR binding to a region which has been shown to be a major activating factor for *csgD* transcription [52]. Restoring biofilm formation was as simple as replacing this SNP, which re-activated the *csgD* promoter, overcoming low RpoS activity in *S*. Enteritidis D7795. We have observed mutations in this OmpR binding site before, in two strains of *Salmonella* serovar Arizonae that had lost biofilm formation, strains that we speculated were adapted for living within the snake intestine [24]. Conservation of the *csgD* promoter SNP in all 167 strains of the Central/East African clade of *S*. Enteritidis from sub-Saharan Africa, and lack of the SNP in 250 'global' *S*. Enteritidis isolates, is strong evidence that loss of this CsgD-regulated biofilm phenotype has being selectively maintained in the invasive population since the most recent common ancestor, circa 1945 [10].

S. Typhimurium D23580 had multiple mutations that influenced the rdar morphotype. Two SNPs were identified in the *csgD* promoter region, each causing a reduction in transcriptional activity. We hypothesized that reduced *csgD* promoter activity in *S*. Typhimurium D23580 could explain the reduced levels of curli production measured in the biofilm flask model. The third mutation was a premature stop codon in *bcsG* that was first identified by Singletary *et al*. in 2016 [47]. Recent work has shown that deletion of *bcsG* shuts off cellulose production [53]; however, in our hands, *S*. Typhimurium D23580 still produced measurable amounts of cellulose. BcsG has at least two functions, to add phosphatidylethanolamine (PE) to monomers of the growing cellulose chain [54], and to stabilize integration of BcsA into the inner membrane, a role that has been ascribed to the N-terminal half of the protein [53]. In *S*. Typhimurium D23580, the premature stop codon occurs at amino acid 247 in BcsG. If the N-terminal region of BcsG is synthesized, it would allow BcsA, the cellulose synthase enzyme, to reach native levels within the cytoplasmic membrane [53]. This could explain why *S*. Typhimurium D23580 can still produce moderate levels of cellulose. The effects of the *csgD* promoter mutations and *bcsG* truncation could have a great impact in the natural lifecycle of *Salmonella*, since they would reduce both the amount of biofilm produced and alter the physical structure of any multicellular aggregates. This could result in isolates that do not survive as well in the environment, as recently demonstrated for invasive *S*. Typhimurium isolates in Mali [33]. Conservation of all three polymorphisms in 100% of African *S*. Typhimurium lineage II isolates, collected from human patients in sSA within the past 30 years [48], is evidence of sustained selection against the rdar morphotype in this lineage.

For *Salmonella* serovar Typhi, we showed that a premature stop codon, resulting in loss of eight amino acids from the C-terminus of CsgD, was sufficient to shut off the rdar morphotype. The truncated CsgD protein had reduced activity, but was able to activate rdar-like morphologies and multicellular aggregation when expressed at higher levels. Therefore, any increases in *csgD* transcription, which can be caused by known promoter mutations [16,24] or potentially host-related environmental signals such as iron limitation [16] or the presence of human bile [37], may be enough to restore CsgD-regulated biofilm phenotypes in *S*. Typhi. To explain why CsgD was not detectable in *S*. Typhi, we hypothesize that the reduced activity of truncated CsgD was not able to activate the genetic feed-forward loop that normally amplifies CsgD production [39,55]. Key dimerization domains exist in the N-terminal half of CsgD [56], but it is not yet clear why the C-terminal truncation would reduce its activity. Despite the lack of a rdar morphotype, it should be noted that *S*. Typhi can form biofilms on human gallstones [57,58], a process which has been simulated with *S*. Typhimurium in a mouse model of chronic carriage; however, the nature of the extracellular matrix in such biofilms is still under investigation [37,59,60]. Our analysis highlighted the presence 4 premature stop codons in the *bcsC*

of the *S*. Typhi strains included in our panel [23]. BcsC is an essential enzyme in cellulose biosynthesis [61], with a C-terminal 'pore' domain that guides the growing cellulose chain out of the cell [62]. This suggests that *S*. Typhi strains could be negative for cellulose production irrespective of reduced CsgD function.

The impacts on the curli and cellulose systems in all three lineages of invasive *S. enterica* isolates provides strong evidence that parallel evolution has occurred [63]. With the association of phenotypic changes to mutations in promoter regions [64–66] or in transcriptional regulatory proteins that can act as bistable switches [40,67,68], we present further evidence that changes in gene expression can drive specialization or ecological divergence without significant changes in gene content [69,70]. We recently reported the increased expression of over 780 genes during the development of multicellular aggregates in flask cultures of *S*. Typhimurium 14028 [39]. The shift in the transcriptome of *Salmonella* cells contrasts sharply with both the small regulon of genes directly controlled by CsgD as well as the complex but limited number of regulatory factors that influence expression and synthesis of CsgD itself [34,45]. Variation in *cis* regulatory regions is hypothesized to have a reduced fitness cost compared to changes in coding sequence, since genetic plasticity is retained [71]. Within the three lineages of invasive *Salmonella* isolates, both regulatory and structural mutations have played a prominent role in loss or impairment of the rdar morphotype. It is possible that the accumulation of genetic changes has been aided by replication and circulation of African strains within immunocompromised hosts [72].

We do not fully understand what the selective pressures are that have led to loss or impairment of biofilm formation in invasive *Salmonella* isolates. The changes we identified were generally specific to the invasive lineages that were investigated and were not conserved across a wider variety of serovars and isolates. We know that enterocolitis-causing isolates replicate to high numbers in the intestine before passing out of the host and into the environment [7]. Replication is aided by *Salmonella*-induced inflammation, which destabilizes the normal microbiota and provides *Salmonella* with a selective advantage due to specific metabolic adaptations [30,73]. In contrast, *Salmonella* strains that cause systemic disease tend to have a stealth and persistence strategy and remain associated with the host for a longer duration of time [8]. The transition from the intestinal to systemic niche is thought to represent an evolutionary bottleneck for *Salmonella* [28], with significant losses in the functional gene repertoire consistently observed for invasive variants [29]. Bistable genetic networks in bacteria, such as the one described for CsgD, are often associated with the formation of two distinguishable phenotypes within a clonal population [74], which is thought to allow genotypes to persist in fluctuating environments [75]. It would make sense for enterocolitis-causing isolates to retain the CsgD network, because the presence of persistent (CsgD-ON) biofilm cells and virulent (CsgD-OFF) planktonic cells would likely improve the odds for future transmission events [15,23,26,39]. Systemic isolates might not require the CsgD network because they are increasingly reliant on human carriers for transmission, as noted for *S*. Typhi [28,57,58]. In general, the biofilm-altering SNPs identified in the invasive isolates were not found in lineages of *S*. Enteritidis or *S*. Typhimurium that typically cause enterocolitis in association with zoonotic transmission.

An alternative explanation for selection against the rdar morphotype is that the biofilm surface structures themselves are targets of the host immune system. Several independent studies have shown that *S*. Typhimurium strains that have lost curli and cellulose production are able to invade tissue culture cells more efficiently [76,77] and spread systemically *in vivo* [78,79]. Curli have been established as potent stimulators of the innate immune system that are recognized by Toll-like receptor 1 and 2 complexes [80,81], as well as intracellular NOD-like receptors [82]. Curli also have the ability to stimulate T-helper cell 17 (Th17) differentiation and increase expression of pro-inflammatory interleukin (IL) cytokines IL-17A, IL-22 and IL-1ß

[82,83]. Some questions still remain, however. Although there is evidence that cellulose can be produced *in vivo* [78,84], it has yet to be conclusively established if this constitutes a biofilm phenotype. It is also difficult to extrapolate how much of an immune response could be generated against *Salmonella* biofilms in immunocompromised, often HIV-positive patients in sSA.

Both *csgD* expression and the rdar morphotype are highly conserved across *Salmonella enterica* and *E. coli* [19,23,24,85], with notable exceptions including *Salmonella* serovars associated with host restriction and systemic disease (*S.* Typhi, *S.* Paratyphi A, and *S.* Gallinarum) and enteroinvasive *E. coli* and *Shigella* [23,26]. Loss of the rdar morphotype in *Salmonella* has been correlated with invasion of the intestinal epithelial lining [23]. There are numerous examples of *Salmonella* biofilm formation providing a survival or persistence advantage under conditions of stress, such as desiccation, nutrient deprivation and disinfection [24,25,86–88]. The impairment or inactivation of the rdar morphotype in the African invasive lineages suggests that their lifestyle could be distinct from lineages of *S.* Typhimurium and *S.* Enteritidis associated with industrialized food supply chains in resource rich settings. Although it is generally accepted that the transmission route for these invasive organisms is fecal-oral, we know very little about their behaviour in the environment between hosts. Based on the data presented, we hypothesize that the African invasive NTS isolates are becoming human-adapted, as has been speculated by other researchers [89,90]. Increased knowledge about the ecological niches that harbor these specialized strains as well as their transmission patterns will be critical for developing public health measures to reduce the morbidity and mortality associated with invasive salmonellosis.

Materials and methods

Bacterial strains, media, and growth conditions

The bacterial strains used in this study are listed in S1 Table. For standard growth, strains were inoculated from frozen stocks onto LB agar (lysogeny broth, 1% NaCl, 1.5% agar) and grown overnight at 37˚C. One isolated colony was used to inoculate 5 mL LB broth and the culture was incubated for 18 hours at 37˚C with agitation at 200 RPM. For colony morphology assays, overnight cultures of each strain were normalized to an optical density of 1.0 at 600 nm and 2 μL were spotted onto 1% tryptone medium containing 1.5% Difco agar (T agar) [36]. To visualize the rdar morphotype, T agar was supplemented with 40–60 μg mL^{-1} Congo red. To visualize cellulose production, T agar was supplemented with calcofluor white (fluorescent brightener 28; Sigma-Aldrich Canada) at a final concentration of 200 μg mL^{-1} [25]. To analyze liquid culture growth under biofilm-inducing conditions, 1 x 10^9 CFU were inoculated into 100 mL of 1% tryptone, pH 7.4, and incubated at 28˚C for 24 or 48 hours with agitation at 200 rpm.

Generation of bacterial luciferase reporters

The pCS26 and pU220 reporter plasmids containing *csgDEFG*, *csgBAC*, and *adrA* promoter sequences from *S.* Typhimurium 14028 fused to the *luxCDABE* operon from *Photorhabdus luminescens* have been described previously [15,24]. The RpoS-dependent reporter plasmid sig38H4 contains the *luxCDABE* operon preceded by a synthetic promoter designed based on the alignment of multiple RpoS-controlled promoters [15]. To generate the pCS26-*cpxR* promoter-*luxCDABE* construct, the *cpx* intergenic region was PCR amplified from *S.* Typhimurium 14028 genomic DNA using primers cpxR1 and cpxR2 (S2 Table) and Phusion high-fidelity DNA polymerase (New England BioLabs), with reaction conditions outlined by the manufacturer. The desired PCR product was purified (Geneaid PCR cleanup kit), sequentially digested with *Xho*I and *Bam*HI (New England BioLabs), and ligated using T4 DNA ligase

(New England BioLabs) into the pCS26 vector cut with *Xho*I and *Bam*HI. To generate *csgDEFG* and *csgBAC* luciferase reporters from each *Salmonella* strain, the *csg* intergenic region was PCR amplified from genomic DNA using primers agfD1 and agfD2 (S2 Table). The PCR products were then ligated into either pCS26 (XhoI-BamHI) or pU220 (BamHI-XhoI) to generate the *csgB* and *csgD* promoter-reporter constructs, respectively.

Luciferase reporter assays

LB overnight cultures of *Salmonella* strains were diluted 1 in 600 in a final volume of 150 μL of 1% tryptone broth supplemented with 50 μg mL^{-1} Kn in 96-well clear-bottom black plates (Costar #9520; Corning Life Sciences). To minimize evaporation of the medium during the assay, cultures were overlaid with 50 μL of mineral oil. Cultures were assayed for absorbance (590 nm, 0.1 s) and luminescence (1s; in counts per second [CPS]) every 30 min during growth at 28˚C with agitation in a Victor X3 multilabel plate reader (Perkin-Elmer).

SDS-PAGE and western blotting

For planktonic cell samples, approximately 5 x 10^{10} cells were sedimented by centrifugation (11,000 x g; 10 min; 4˚C). For biofilm aggregate samples, approximately 30 mg samples were resuspended in 1 mL of water and homogenized with a glass tissue grinder (Product #7727–2, Corning Life Sciences) for 25 dounces, prior to centrifugation (10,000 x g; 1 min) to sediment the cell material. Sedimented samples were resuspended in 1 mL of SDS-PAGE sample buffer without 2-mercaptoethanol and bromophenol blue and boiled for 5 min. Using the DC protein assay (Bio-Rad Laboratories), cell lysates were normalized to a final protein concentration of 3 mg/mL. Bromophenol blue (0.0002% final concentration) and 2-mercaptoethanol (0.2% final concentration) were added to each lysate before loading 15 μg of total protein per lane. SDS-PAGE was performed with a 5% stacking gel and a 12 or 15% resolving gel. Proteins were transferred to nitrocellulose for 40 min at 25 V using a Trans-Blot SD semi-dry transfer cell (Bio-Rad Laboratories) in tris-glycine buffer supplemented with methanol. To detect curli fimbriae, cell debris was sedimented following boiling in SDS-PAGE sample buffer, washed twice with 500 μL of distilled water, suspended in 250 μL of 90% formic acid, frozen and lyophilized [36]. The dried samples were resuspended in SDS-PAGE sample buffer and loaded directly without boiling into each SDS-PAGE gel lane. CsgD was detected using a CsgD-specific monoclonal antibody at a 1-in-6 dilution of tissue culture supernatant (ImmunoPrecise Antibodies Ltd., Victoria, BC). To detect RpoS protein, a commercially available mouse polyclonal immune serum recognizing epitope 33 to 256 of *E. coli* RpoS was used at a 1-in-2000 dilution (BioLegend; 1RS1). CsgA, the major subunit of curli fimbriae, was detected by using a rabbit polyclonal serum raised against whole purified curli [36]. GroEL was used as a protein-loading control and was detected with a 1-in-60,000 dilution of rabbit polyclonal immune serum (Sigma-Aldrich; G6532). Secondary antibodies IRDye 800CW Goat anti-Mouse immunoglobulin G (IgG) or 680RD Goat anti-rabbit IgG (Mandel Scientific) were used at a 1-in-10,000 dilution and detected using the the Odyssey CLx imaging system and Image Studio 4.0 software package (Li-Cor Biosciences).

Reference genome sequences

Whole genome sequences were obtained from the National Centre for Biotechnology Information (NCBI) via the following accession numbers: NC_016810 (*S*. Typhimurium SL1344); NC_016854 (*S*. Typhimurium D23580); NC_003198 (*S*. Typhi CT18). Mapped assemblies for *S*. Typhi H58 haplotype strains E02-2759, E03-9804, ISP03-07467, ISP04-06979 were available from http://www.sanger.ac.uk/Projects/S_typhi [28]. The *S*. Typhi 8(04)N genome was

available from the European Nucleotide Archive under the name SGB112 and associated with the assembly number GCA_001362315.1. The sequence for *S*. Enteritidis D7795 was available from the Public Health England Pathogens BioProject on NCBI (accession number PRJNA248792).

Chromosomal DNA isolation, genome sequencing and sequence alignments

Overnight liquid cultures of *Salmonella* serovar Enteritidis strains 301 and ATCC 4931 were sub-cultured 1 in 100 in 200 mL LB broth and grown at 37˚C for 2.5 hours. Approximately 7 x 10^8 cells were centrifuged (6000 x g, 10 minutes, 4˚C), resuspended in 25 mM Tris, 1 mM EDTA solution (pH 8.0) to a total volume of 3.5 mL, and then treated with 10 mg lysozyme (Sigma-Aldrich; L6876) and 200 units of mutanolysin (Sigma-Aldrich; M9901) for 1 hour at 37˚C. For cell lysis, each sample was treated with 50 μL of 25% SDS, 1 mg proteinase K (Applied Biosystems; AM23548), and 125 μL of a 5M sodium chloride solution and incubated at 65˚C for 30 minutes. Nucleic acid was isolated from cell lysates using a series of phenol:chloroform:isoamyl and phenol:chloroform extractions, precipitated by the addition of ammonium acetate (at a final concentration of 2M), washed with ethanol, and resuspended in Tris-EDTA solution (10 mM Tris, 1 mM EDTA, pH 8.0). To remove RNA, RNase A was added to each sample at a final concentration of 0.2 mg/mL and incubated for 1 hour at 37˚C. Samples were purified once more by phenol:chloroform:isoamyl extraction, precipitated with ammonium acetate, washed with ethanol, and resuspended in a final volume of 200 μL of Tris-EDTA solution.

Purified chromosomal DNA samples were fragmented by cup horn sonication with a high-intensity ultrasonic processor (Vibra-Cell, Danbury, CT) for 10 cycles of a 30-second pulses and 2 minute rest. DNA libraries were prepared from 1 μg of fragmented DNA using the NEBNext Ultra DNA Library Prep Kit for Illumina (New England BioLabs; E7645S) and NEBNext Multiplex Oligos for Illumina (Index Primer Set 2) (New England BioLabs; E7500S) according to the manufacturer's protocols. Adaptor-ligated DNA was size-selected between 400 and 500 bp total library size (length of insert sequence + adaptor sequence) following kit instructions. DNA samples were assessed for quality, purity, and integrity using a NanoDrop ND-1000 spectrophotometer (Fisher Scientific) and an Agilent 2100 Bioanalyzer with a High Sensitivity DNA chip (Agilent Technologies; 5067–4626). To ensure efficient adaptor ligation, samples were analyzed by quantitative PCR using the KAPA Library Quantification Kit for the Illumina platform (KAPA Biosystems; KK4824). Samples were sequenced using the MiSeq Reagent Kit version 3, 600 cycles (2 x 300 bp read length) (Illumina; MS-102-3003).

Genome sequencing assembly and DNA sequence alignments were performed using the Geneious Pro v8.1.5 software package (Biomatters). Paired-end sequence reads from Entertidis strains 4931 and 301 were assembled into contigs based on mapping to the *S*. Enteritidis reference genome P125109 (NC_011294). Sequence alignments were performed using ClustalW and an IUB cost matrix (gap open cost of 15, gap extend cost of 6.66). A phylogenetic tree for strains included in this study was constructed based on the *csg* operon region by using the Geneious Tree Builder program [91], with the Tamura-Nei model of genetic substitution and the neighbour-joining algorithm with bootstrapping (1000 replicates and support threshold of 70%).

Genome engineering in *S*. Enteritidis D7795, *S*. Typhimurium D23580 and *S*. Typhimurium 14028

The I-*Sce*I suicide plasmid system developed by Victor de Lorenzo and colleagues [92] was used for genome engineering. The following fragments were PCR amplified: 1) *csg* intergenic

region from *S*. Typhimurium 14028 or *S*. Enteritidis D7795 using primers agfD3-FWD and agfD4-REV (S2 Table); 2) *csg* intergenic region from *S*. Typhimurium 14028 using primers agfD5-FWD and agfD6-REV (S2 Table); 3) *bcsG*-containing region from *S*. Typhimurium 14028 or *S*. Typhimurium D23580 using primers bcsG-checkF and bcsG-checkR (S2 Table); and 4) *csgD*-containing region from *S*. Typhi CT18 using primers csgDORFstartPstI and csgDreplaceREV (S2 Table). Phusion high-fidelity DNA polymerase was used for amplification, following reaction conditions outlined by the manufacturer (New England BioLabs). PCR products were purified, digested with *Bam*HI and *Pst*I (New England Biolabs) and ligated into *Bam*HI/*Pst*I-digested pSEVA212 [94]. Clones corresponding to each product were selected in *E. coli* S17-1 (λpir) and mating was performed to move the plasmid constructs into *S*. Enteritidis D7795, *S*. Typhimurium D23580 or *S*. Typhimurium 14028. Merodiploid strains with pSEVA212 plasmid constructs inserted into the genome were selected by growth on M9 minimal agar supplemented with 1mM MgSO4, 0.2% glucose and 100 μg mL^{-1} kanamycin (M9-Glc-Kn100) and confirmed by re-streaking onto M9-Glc-Kn100 agar. Purified pSEVA628S (200–300 ng) [93] was transformed into merodiploid strains by electroporation with selection on Luria agar supplemented with 1 mM m-toluate and 20 μg mL^{-1} gentamicin. Resulting colonies were re-streaked onto Luria agar supplemented with 20 μg mL^{-1} gentamicin, streaked on Luria agar supplemented with 50 μg mL^{-1} kanamycin to confirm loss of the pSEVA212 plasmid, and streaked on TCR plates (1% tryptone, 1.5% agar, 40 μg mL^{-1} Congo red) to check the biofilm phenotype. Colonies were selected from TCR plates and grown at 37˚C for two overnight growth steps without gentamicin to generate cells that lack pSEVA628S. Final colonies were streaked onto (1) Luria agar, (2) Luria agar + 50 μg mL^{-1} Kn, (3) Luria agar + 20 μg mL^{-1} gentamicin, and (4) TCR plates to select the desired phenotypes. To confirm the genotypes, *csgD* promoter-, *bcsG*- or *csgD*-containing regions were PCR-amplified from resulting strains and Sanger sequencing was performed (Eurofins MWG Operon; Louisville, Kentucky, USA).

Generating p3xFLAG vector constructs

The pFLAG-CTC expression vector (Sigma Aldrich #E8408) was maintained in *E. coli* DH10B. Purified pFLAG-CTC was digested with *Sal*I and *Xho*I restriction enzymes. A synthetic polylinker generated from phosphorylated oligonucleotides 3xFLAG-linkerA and 3xFLAG-linkerB (S2 Table) was ligated into the digested vector to generate p3xFLAG. The *csgD* open reading frame was PCR amplified from *S*. Typhimurium 14028 or S. Typhi CT18 genomic DNA using primers csgD-ORF-start and csgD-ORF-end (S2 Table), followed by digestion with *Xho*I and *Bam*HI. The *csgD* fragments were ligated into *Xho*I- and *Bgl*II-digested plasmid to generate p3xFLAG/*csgD*^14028 and p3xFLAG/*csgD*^CT18. Insertion of the 3xFLAG polylinker and *csgD* ORF was confirmed by DNA sequencing. The *csgD*^CT18 allele does not have the 3xFLAG linker attached because of the premature stop codon in the 3' end of csgD. The resulting p3xFLAG vectors were transformed into *S*. Typhimurium Δ*csgD* prior to analysis of biofilm formation.

In silico detection of identified polymorphisms in Salmonella serovar Enteritidis and serovar Typhimurium strains

S. Enteritidis D77 is one of 167 isolates recently sequenced and described as being part of a distinct clade of *S*. Enteritidis featuring genomic degradation and geographical restriction to central and eastern Africa [10]. Similarly, *S*. Typhimurium D23580 is part of a unique lineage of *S*. Typhimurium, consisting of isolates from sSA [48]. Genome assemblies of the *S*. Enteritidis and *S*. Typhimurium isolates were investigated for the presence or absence of *csgD* promoter

or bcsG SNPs through *in silico* PCR. Primers of the *csgD* sequence with a SNP in the promoter region were searched for in the genome assemblies using in_silico_pcr.py script (https://github.com/simonrharris/in_silico_pcr), allowing zero changes for a match. Names and accession numbers for each strain included in the SNP screening are listed in S3 and S4 Tables. A similar approach was followed to screen the genome sequences of 248 diverse *S. enterica* isolates [94], as well as 82 strains of serovar Enteritidis that are part of the West Africa lineage of strains originally described by Feasey et al. [10].

Statistical analysis

Statistical analysis was performed using GraphPad Prism versions 7.0c and 8.0.2. Data collected from promoter-reporter luciferase assays was reported as the maximum luciferase expression measured during the course of the experiment, and was expressed as the mean ± the standard deviation. This data was logarithmically transformed and evaluated for normal distribution using the Shapiro-Wilk normality test. If the data was normally distributed, comparisons of the mean maximum luciferase expression levels obtained from multiple experiments were performed using an ordinary one-way ANOVA with post-hoc analysis via Holm-Sidak's multiple comparisons test with statistical significance set at $p < 0.05$. If the data was not normally distributed, comparisons were performed using the Kruskal-Wallis test with post-hoc analysis via Dunn's multiple comparison test with statistical significance set at a p value of 0.05.

Data availablility

All numerical data and statistical analysis has been deposited in figshare (https://figshare.com/) and is publicly available at doi: https://doi.org/10.6084/m9.figshare.8220866.v1. The Illumina paired end sequence reads comprising the genome sequences of *S. enterica* suspecies *enterica* serovar Enteritidis strains have been deposited in the Sequence Read Archives (strain 301—SAMN11956692; strain ATCC 4931—SAMN11956691).

Supporting information

S1 Fig. Non-typhoidal *Salmonella* strains display temperature restriction of the rdar morphotype. (A) Morphological comparison of colonies grown for five days at 28˚C or 37˚C on T agar supplemented with 40 μg mL^{-1} Congo red. (B) Maximum expression from a curli-specific reporter (*csgBAC*::*luxCDABE*) in each strain grown for 48 h at 28˚C or 37˚C. Luciferase expression was measured every 30 min during continuous culture.
(TIF)

S2 Fig. Qualitative comparison of multicellular aggregates produced by *S*. Typhimurium 14028 and D23580 strains. Multicellular aggregates and planktonic cells formed from liquid cultures of *S*. Typhimurium 14028 and *S*. Typhimurium D23580. Images match those presented in Fig 1, but have been enhanced to emphasize aggregates within the samples. Aggregates formed by *S*. Typhimurium D23580 cells appear structurally distinct from aggregates formed by *S*. Typhimurium 14028 cells.
(TIF)

S3 Fig. Detection of RpoS sigma factor protein synthesis and activity. (A) Whole cell lysates were generated from multicellular aggregates and planktonic cells isolated from flask cultures of *Salmonella* strains after 24 hours of growth and probed for synthesis of RpoS. Lysates were normalized by total protein concentration. GroEL was used as a loading control to ensure that equal amounts of protein were loaded into each sample lane. (B) RpoS activity was evaluated

by measuring luminescence from a synthetic, RpoS-dependent promoter-reporter construct expressed in each *Salmonella* strain during 48 hours of growth. Graphed values represent the maximum reporter activity recorded and is reported as counts per second (CPS). (C) Absorbance measurements represent strain growth in microaerophilic conditions in 96-well microtiter plates. Each curve represents one biological replicate; $n = 4$ per strain. (D) Time course of RpoS-dependent promoter activity in each indicated strain.
(TIF)

S4 Fig. Expression profile of a *csgB* promoter-reporter construct derived from *S*. Typhimurium SL1344 sequence in *S*. Typhimurium 14028 cells. Each curve represents one biological replicate; $n = 3$.
(TIF)

S5 Fig. Activity of *S*. Typhimurium 14028 *adrA* and *cpxRA* promoters in representative nontyphoidal and typhoidal *Salmonella* strains. Promoter-reporter constructs derived from *S*. Typhimurium 14028 *adrA* and *cpxRA* promoter sequences were introduced into each of the strains. Letters above the bars indicate mean values that were statistically similar to (black font) or different from (red font) other mean values. #, values below the activity threshold as established in [24]. Each bar represents the mean value from 3–5 independent biological replicates and error bars represent the standard deviations.
(TIF)

S6 Fig. Detecting multicellular aggregates in liquid cultures of *S. Enteritidis* D7795 strains generated through genome engineering. Clones that contain (-47T>T) or do not contain (-47T>C) the identified 'T' promoter SNP at position -47 were grown in flasks of 1% tryptone for 24 hours before being evaluated for the ability to form aggregates in liquid cultures.
(TIF)

S7 Fig. Sequence comparison of the *csg* operons in nontyphoidal and typhoidal *Salmonella* strains. Simplified multiple sequence alignment of the *csgDEFG* and *csgBAC* operons highlighting serovar- and strain-specific single nucleotide polymorphisms representing both nonsynonymous and synonymous mutations: *S*. Typhimurium and D23580 (yellow), *S*. Enteritidis and D7795 (light blue), and *S*. Typhi (pink). Other highlighted changes (black) include the SNP in D7795 that inactivates *csgD* transcription, and the SNP in *S*. Typhi strains that introduces a premature stop codon in *csgD*, yielding a CsgD protein that is truncated by 8 amino acids. The long black bar represents the DNA region with nucleotide position numbers listed above and *csg* genes shown below as yellow-boxed arrows. Special sequence features involved in operon regulation are highlighted above and below the black bar: -35 and -10 promoter regions (black elbow arrows), H-NS binding region (grey box), CpxR-binding sites (black bars), and OmpR binding regions (hatched boxes).
(TIF)

S8 Fig. Sequence comparison of the *bcs* operons in nontyphoidal and typhoidal *Salmonella* strains. Simplified multiple sequence alignment of the *bcsRQABZC* and *bcsEFG* operons highlighting non-synonymous serovar- and strain-specific single nucleotide polymorphisms. Non-synonymous SNPs are shown in yellow; black bars indicate SNPs that yield premature stop codons (i.e. *bcsG* in D23580; four SNPs in *bcsC* in *S*. Typhi strains). The long black bar represents the DNA region with nucleotide position numbers listed above and *bcs* genes shown below as yellow-boxed arrows. *bcsR*, *bcsF* and a small hypothetical protein are shown without their names listed.
(TIF)

S1 Table. Strains and plasmids used in this study.
(PDF)

S2 Table. Oligonucleotides used in this study.
(PDF)

S3 Table. List of *S.* Enteritidis isolates corresponding to Fig 6A that were analyzed for presence of *csgD* promoter SNP.
(XLSX)

S4 Table. List of *S.* Typhimurium isolates corresponding to Fig 6B that were analyzed for presence of *csgD* promoter and *bcsG* SNPs.
(XLSX)

S5 Table. 248 isolates representing 55 serovars of *S. enterica* subspecies *enterica* that were screened for SNPs in the *csgD* promoter, *csgD* ORF, and *bcsG* ORF.
(XLSX)

S6 Table. *S.* Enteritidis isolates associated with invasive human disease in West Africa that were screened for SNPs in the *csgD* promoter, *csgD* ORF, and *bcsG* ORF.
(XLSX)

Acknowledgments

The authors wish to thank Gordon Dougan, Robert Kingsley and members of their laboratories for sending *Salmonella* strains; Victor de Lorenzo for providing I-*Sce*I plasmids; Nicolas Wenner and Jay Hinton for providing the I-*Sce*I genome engineering protocols; Bill Kay and Mike Surette for critical reading of the manuscript. Published as VIDO manuscript #844.

Author Contributions

Conceptualization: Keith D. MacKenzie, Yejun Wang, Aaron P. White.

Data curation: Keith D. MacKenzie, Yejun Wang, Patrick Musicha, Aaron P. White.

Formal analysis: Keith D. MacKenzie, Yejun Wang, Patrick Musicha, Nicholas A. Feasey, Aaron P. White.

Funding acquisition: Aaron P. White.

Investigation: Keith D. MacKenzie, Yejun Wang, Patrick Musicha, Elizabeth G. Hansen, Melissa B. Palmer, Dakoda J. Herman, Aaron P. White.

Methodology: Keith D. MacKenzie, Yejun Wang, Aaron P. White.

Project administration: Aaron P. White.

Resources: Yejun Wang, Patrick Musicha, Nicholas A. Feasey, Aaron P. White.

Software: Yejun Wang, Patrick Musicha.

Supervision: Aaron P. White.

Validation: Keith D. MacKenzie, Yejun Wang, Aaron P. White.

Visualization: Keith D. MacKenzie, Melissa B. Palmer, Aaron P. White.

Writing – original draft: Keith D. MacKenzie, Nicholas A. Feasey, Aaron P. White.

Writing – review & editing: Keith D. MacKenzie, Yejun Wang, Patrick Musicha, Nicholas A. Feasey, Aaron P. White.

References

1. Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleesschauwer B, et al. World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. PLOS Med. 2015 Dec 3; 12(12):e1001921. https://doi.org/10.1371/journal.pmed.1001921 PMID: 26633831
2. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal *Salmonella* gastroenteritis. Clin Infect Dis. 2010 Mar 15; 50(6):882–9. https://doi.org/10.1086/650733 PMID: 20158401
3. Reddy EA, Shaw AV, Crump JA. Community-acquired bloodstream infections in Africa: a systematic review and meta-analysis. Lancet Infect Dis. 2010; 10(6):417–32. https://doi.org/10.1016/S1473-3099(10)70072-4 PMID: 20510282
4. Feasey NA, Dougan G, Kingsley RA, Heydermann RS, Gordon MA. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. The Lancet. 2012 Jun 30; 379(9835):2489–99.
5. Ao TT, Feasey NA, Gordon MA, Keddy KH, Angulo FJ, Crump JA. Global Burden of Invasive Nontyphoidal *Salmonella* Disease, 2010. Emerg Infect Dis. 2015; 21(6):941.
6. Haselbeck AH, Panzner U, Im J, Baker S, Meyer CG, Marks F. Current perspectives on invasive nontyphoidal Salmonella disease. Curr Opin Infect Dis. 2017; 30(5).
7. Bäumler A, Fang FC. Host specificity of bacterial pathogens. Cold Spring Harb Perspect Med. 2013 Dec 1; 3(12):a010041–1. https://doi.org/10.1101/cshperspect.a010041 PMID: 24296346
8. Dougan G, Baker S. *Salmonella enterica* Serovar Typhi and the Pathogenesis of Typhoid Fever. Annu Rev Microbiol. 2014 Sep 8; 68(1):317–36.
9. Okoro CK, Barquist L, Connor TR, Harris SR, Clare S, Stevens MP, et al. Signatures of Adaptation in Human Invasive Salmonella Typhimurium ST313 Populations from Sub-Saharan Africa. PLoS Negl Trop Dis. 2015 Mar 24; 9(3):e0003611. https://doi.org/10.1371/journal.pntd.0003611 PMID: 25803844
10. Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, et al. Distinct *Salmonella* Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. Nat Genet. 2016 Aug 22; 48(10):1211–7. https://doi.org/10.1038/ng.3644 PMID: 27548315
11. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, et al. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature. 2010 Sep 23; 467(7314):426–9. https://doi.org/10.1038/nature09415 PMID: 20864996
12. Nuccio S-P, Bäumler AJ. Comparative Analysis of Salmonella Genomes Identifies a Metabolic Network for Escalating Growth in the Inflamed Gut. mBio. 2014 May 1; 5(2):e00929–14. https://doi.org/10.1128/mBio.00929-14 PMID: 24643865
13. Lawley TD, Bouley DM, Hoy YE, Gerke C, Relman DA, Monack DM. Host Transmission of *Salmonella enterica* Serovar Typhimurium Is Controlled by Virulence Factors and Indigenous Intestinal Microbiota. Infect Immun. 2008 Jan 1; 76(1):403. https://doi.org/10.1128/IAI.01189-07 PMID: 17967858
14. Winfield MD, Groisman EA. Role of Nonhost Environments in the Lifestyles of *Salmonella* and *Escherichia coli*. Appl Environ Microbiol. 2003 Jul 1; 69(7):3687. https://doi.org/10.1128/AEM.69.7.3687-3694.2003 PMID: 12839733
15. White AP, Gibson DL, Kim W, Kay WW, Surette MG. Thin Aggregative Fimbriae and Cellulose Enhance Long-Term Survival and Persistence of *Salmonella*. J Bacteriol. 2006 Apr 17; 188(9):3219–27. https://doi.org/10.1128/JB.188.9.3219-3227.2006 PMID: 16621814
16. Römling U, Sierralta WD, Eriksson K, Normark S. Multicellular and aggregative behaviour of *Salmonella typhimurium* strains is controlled by mutations in the *agfD* promoter. Mol Microbiol. 1998 Apr; 28(2):249–64. PMID: 9622351
17. Zogaj X, Nimtz M, Rohde M, Bokranz W, Römling U. The multicellular morphotypes of *Salmonella typhimurium* and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. Mol Microbiol. 2001 Mar; 39(6):1452–63. PMID: 11260463
18. Gibson DL, White AP, Snyder SD, Martin S, Heiss C, Azadi P, et al. *Salmonella* Produces an O-Antigen Capsule Regulated by AgfD and Important for Environmental Persistence. J Bacteriol. 2006 Nov 1; 188(22):7722–30. https://doi.org/10.1128/JB.00809-06 PMID: 17079680
19. Römling U, Bian Z, Hammar M, Sierralta WD, Normark S. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. J Bacteriol. 1998 Feb; 180(3):722–31. PMID: 9457880

20. Latasa C, Roux A, Toledo-Arana A, Ghigo J-M, Gamazo C, Penadés JR, et al. BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar Enteritidis. Mol Microbiol. 2005 Oct 14; 58(5):1322–39. https://doi.org/10.1111/j.1365-2958.2005.04907.x PMID: 16313619

21. Zakikhany K, Harrington CR, Nimtz M, Hinton JCD, Römling U. Unphosphorylated CsgD controls biofilm formation in *Salmonella enterica* serovar Typhimurium. Mol Microbiol. 2010 Jul 6; 77(3):771–86. https://doi.org/10.1111/j.1365-2958.2010.07247.x PMID: 20545866

22. White AP, Gibson DL, Collinson SK, Banser PA, Kay WW. Extracellular Polysaccharides Associated with Thin Aggregative Fimbriae of *Salmonella enterica* Serovar Enteritidis. J Bacteriol. 2003 Aug 29; 185(18):5398–407. https://doi.org/10.1128/JB.185.18.5398-5407.2003 PMID: 12949092

23. Römling U, Bokranz W, Rabsch W, Zogaj X, Nimtz M, Tschäpe H. Occurrence and regulation of the multicellular morphotype in Salmonella serovars important in human disease. Int J Med Microbiol. 2003 Jan 1; 293(4):273–85. https://doi.org/10.1078/1438-4221-00268 PMID: 14503792

24. White AP, Surette MG. Comparative Genetics of the rdar Morphotype in *Salmonella*. J Bacteriol. 2006 Dec 4; 188(24):8395–406. https://doi.org/10.1128/JB.00798-06 PMID: 17012381

25. Solano C, García B, Valle J, Berasain C, Ghigo J-M, Gamazo C, et al. Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. Mol Microbiol. 2002 Feb 28; 43(3):793–808. PMID: 11929533

26. MacKenzie KD, Palmer MB, Köster WL, White AP. Examining the Link between Biofilm Formation and the Ability of Pathogenic Salmonella Strains to Colonize Multiple Host Species. Front Vet Sci. 2017 Aug 25; 4:307–19.

27. Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, et al. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. Genome Res. 2009 Dec 1; 19(12):2279–87. https://doi.org/10.1101/gr.091017.109 PMID: 19901036

28. Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill F-X, Goodhead I, et al. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. Nat Genet. 2008 Aug; 40 (8):987–93. https://doi.org/10.1038/ng.195 PMID: 18660809

29. Holt KE, Thomson NR, Wain J, Langridge GC, Hasan R, Bhutta ZA, et al. Pseudogene accumulation in the evolutionary histories of *Salmonella enterica* serovars Paratyphi A and Typhi. BMC Genomics. 2009; 10(1):36.

30. Yang J, Barrila J, Roland KL, Kilbourne J, Ott CM, Forsyth RJ, et al. Characterization of the Invasive, Multidrug Resistant Non-typhoidal Salmonella Strain D23580 in a Murine Model of Infection. PLoS Negl Trop Dis. 2015 Jun 19; 9(6):e0003839. https://doi.org/10.1371/journal.pntd.0003839 PMID: 26091096

31. Carden S, Okoro C, Dougan G, Monack D. Non-typhoidal Salmonella Typhimurium ST313 isolates that cause bacteremia in humans stimulate less inflammasome activation than ST19 isolates associated with gastroenteritis. FEMS Pathogens and Disease. 2014 Dec 24; 73(4):2247.

32. Parsons BN, Humphrey S, Salisbury AM, Mikoleit J, Hinton JCD, Gordon MA, et al. Invasive Non-Typhoidal Salmonella Typhimurium ST313 Are Not Host-Restricted and Have an Invasive Phenotype in Experimentally Infected Chickens. PLoS Negl Trop Dis. 2013 Oct 10; 7(10):e2487. https://doi.org/10.1371/journal.pntd.0002487 PMID: 24130915

33. Ramachandran G, Aheto K, Shirtliff ME, Tennant SM. Poor biofilm-forming ability and long-term survival of invasive Salmonella Typhimurium ST313. FEMS Pathog Dis. 2016 Jun 23; 74(5):ftw049.

34. Steenackers H, Hermans K, Vanderleyden J, De Keersmaecker SCJ. Salmonella biofilms: An overview on occurrence, structure, regulation and eradication. Food Res Int. 2012 Feb 29; 45(2):502–31.

35. Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ, Feasey NA, et al. Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella* Typhi identifies inter- and intracontinental transmission events. Nat Genet. 2015 May 11; 47(6):632–9. https://doi.org/10.1038/ng.3281 PMID: 25961941

36. Collinson SK, Emödy L, Müller KH, Trust TJ, Kay WW. Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*. J Bacteriol. 1991 Aug 1; 173(15):4773–81. https://doi.org/10.1128/jb.173.15.4773-4781.1991 PMID: 1677357

37. González JF, Tucker L, Fitch J, Wetzel A, White P, Gunn JS. Human bile-mediated regulation of *Salmonella* curli fimbriae. J Bacteriol. 2019 Apr 1:1–32.

38. White AP, Gibson DL, Grassl GA, Kay WW, Finlay BB, Vallance BA, et al. Aggregation via the Red, Dry, and Rough Morphotype Is Not a Virulence Adaptation in *Salmonella enterica* Serovar Typhimurium. Infect Immun. 2008 Feb 21; 76(3):1048–58. https://doi.org/10.1128/IAI.01383-07 PMID: 18195033

39. MacKenzie KD, Wang Y, Shivak DJ, Wong CS, Hoffman LJL, Lam S, et al. Bistable Expression of CsgD in *Salmonella enterica* Serovar Typhimurium Connects Virulence to Persistence. Infect Immun. 2015 May 12; 83(6):2312–26. https://doi.org/10.1128/IAI.00137-15 PMID: 25824832

40. Grantcharova N, Peters V, Monteiro C, Zakikhany K, Römling U. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar Typhimurium. J Bacteriol. 2010 Jan; 192(2):456–66. https://doi.org/10.1128/JB.01826-08 PMID: 19897646

41. Hammar MR, Arnqvist A, Bian Z, Olsén A, Normark S. Expression of two *csg* operons is required for production of fibronectin- and Congo red-binding curli polymers in *Escherichia coli* K-12. Mol Microbiol. 1995 Nov; 18(4):661–70.

42. Barnhart MM, Chapman MR. Curli Biogenesis and Function. Annu Rev Microbiol. 2006 Oct; 60(1):131–47.

43. Taylor JD, Zhou Y, Salgado PS, Patwardhan A, McGuffie M, Pape T, et al. Atomic Resolution Insights into Curli Fiber Biogenesis. Structure. 2011 Sep 7; 19(9):1307–16. https://doi.org/10.1016/j.str.2011.05.015 PMID: 21893289

44. Shivak DJ, MacKenzie KD, Watson NL, Pasternak JA, Jones BD, Wang Y, et al. A Modular, Tn 7-Based System for Making Bioluminescent or Fluorescent *Salmonella* and *Escherichia coli* Strains. Appl Environ Microbiol. 2016 Jul 29; 82(16):4931–43. https://doi.org/10.1128/AEM.01346-16 PMID: 27260360

45. Ogasawara H, Yamada K, Kori A, Yamamoto K, Ishihama A. Regulation of the *Escherichia coli csgD* promoter: interplay between five transcription factors. Microbiology. 2010 Aug 2; 156(8):2470–83.

46. Ogasawara H, Yamamoto K, Ishihama A. Regulatory role of MlrA in transcription activation of *csgD*, the master regulator of biofilm formation in *Escherichia coli*. FEMS Microbiol Lett. 2010 Sep 27; 312 (2):160–8. https://doi.org/10.1111/j.1574-6968.2010.02112.x PMID: 20874755

47. Singletary LA, Karlinsey JE, Libby SJ, Mooney JP, Lokken KL, Tsolis RM, et al. Loss of Multicellular Behavior in Epidemic African Nontyphoidal Salmonella enterica Serovar Typhimurium ST313 Strain D23580. mBio. 2016 May 4; 7(2):e02265–15. https://doi.org/10.1128/mBio.02265-15 PMID: 26933058

48. Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN, et al. Intracontinental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. Nat Genet. 2012 Nov; 44(11):1215–21. https://doi.org/10.1038/ng.2423 PMID: 23023330

49. Parkhill J, Dougan G, James KD, Thomson NR, Pickard D, Wain J, et al. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature. 2001 Oct 25; 413 (6858):848–52. https://doi.org/10.1038/35101607 PMID: 11677608

50. Dieckmann R, Malorny B. Rapid Screening of Epidemiologically Important *Salmonella enterica* subsp. enterica Serovars by Whole-Cell Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry. Appl Environ Microbiol. 2011 Jun 7; 77(12):4136–46. https://doi.org/10.1128/AEM.02418-10 PMID: 21515723

51. Huang K-J, Igo MM. Identification of the Bases in the *ompF* Regulatory Region, which Interact with the Transcription Factor OmpR. J Mol Biol. 1996; 262(5):615–28. https://doi.org/10.1006/jmbi.1996.0540 PMID: 8876642

52. Gerstel U, Park C, Römling U. Complex regulation of *csgD* promoter activity by global regulatory proteins. Mol Microbiol. 2004 Jan 28; 49(3):639–54.

53. Sun L, Vella P, Schnell R, Polyakova A, Bourenkov G, Li F, et al. Structural and Functional Characterization of the BcsG Subunit of the Cellulose Synthase in *Salmonella typhimurium*. J Mol Biol. 2018; 430 (18, Part B):3170–89.

54. Thongsomboon W, Serra DO, Possling A, Hadjineophytou C, Hengge R, Cegelski L. Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose. Science. 2018 Jan 19; 359 (6373):334. https://doi.org/10.1126/science.aao4096 PMID: 29348238

55. Gualdi L, Tagliabue L, Landini P. Biofilm formation-gene expression relay system in *Escherichia coli*: modulation of sigmaS-dependent gene expression by the CsgD regulatory protein via sigmaS protein stabilization. J Bacteriol. 2007 Nov; 189(22):8034–43. https://doi.org/10.1128/JB.00900-07 PMID: 17873038

56. Wen Y, Ouyang Z, Devreese B, He W, Shao Y, Lu W, et al. Crystal structure of master biofilm regulator CsgD regulatory domain reveals an atypical receiver domain. Prot Sci. Aug 22; 26(10):2073–82.

57. Gunn JS, Marshall JM, Baker S, Dongol S, Charles RC, Ryan ET. *Salmonella* chronic carriage: epidemiology, diagnosis, and gallbladder persistence. Trends Microbiol. 2014 Nov 1; 22(11):648–55. https://doi.org/10.1016/j.tim.2014.06.007 PMID: 25065707

58. Crawford RW, Rosales-Reyes R, Ramirez-Aguilar MDLL, Chapa-Azuela O, Alpuche-Aranda C, Gunn JS. Gallstones play a significant role in *Salmonella* spp. gallbladder colonization and carriage. Proc Natl Acad Sci USA. 2010 Mar 2; 107(9):4353–8. https://doi.org/10.1073/pnas.1000862107 PMID: 20176950

59. Neiger MR, González JF, Gonzalez-Escobedo G, Kuck H, White P, Gunn JS. Pathoadaptive alteration of *Salmonella* biofilm formation in response to the gallbladder environment. J Bacteriol. 2019 Apr 8;:1–37.

60. Crawford RW, Reeve KE, Gunn JS. Flagellated but Not Hyperfimbriated *Salmonella enterica* Serovar Typhimurium Attaches to and Forms Biofilms on Cholesterol-Coated Surfaces. J Bacteriol. 2010 May 26; 192(12):2981–90. https://doi.org/10.1128/JB.01620-09 PMID: 20118264

61. Wong HC, Fear AL, Calhoon RD, Eichinger GH, Mayer R, Amikam D, et al. Genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci USA. 1990 Oct 1; 87(20):8130. https://doi.org/10.1073/pnas.87.20.8130 PMID: 2146681

62. Römling U, Galperin MY. Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends Microbiol. 2015 Sep 1; 23(9):545–57. https://doi.org/10.1016/j.tim.2015.05.005 PMID: 26077867

63. Stern DL. The genetic causes of convergent evolution. Nat Rev Genet. 2013 Nov; 14(11):751–64. https://doi.org/10.1038/nrg3483 PMID: 24105273

64. Hammarlöf DL, Kröger C, Owen SV, Canals R, Lacharme-Lora L, Wenner N, et al. Role of a single noncoding nucleotide in the evolution of an epidemic African clade of *Salmonella*. Proc Natl Acad Sci USA. 2018 Mar 13; 115(11):E2614. https://doi.org/10.1073/pnas.1714718115 PMID: 29487214

65. Osborne SE, Walthers D, Tomljenovic AM, Mulder DT, Silphaduang U, Duong N, et al. Pathogenic adaptation of intracellular bacteria by rewiring a *cis*-regulatory input function. Proc Natl Acad Sci USA. 2009 Mar 10; 106(10):3982. https://doi.org/10.1073/pnas.0811669106 PMID: 19234126

66. Ilyas B, Mulder DT, Little DJ, Elhenawy W, Banda MM, Pérez-Morales D, et al. Regulatory Evolution Drives Evasion of Host Inflammasomes by Salmonella Typhimurium. Cell Reports. 2018; 25(4):825–5. https://doi.org/10.1016/j.celrep.2018.09.078 PMID: 30355489

67. Chin CY, Tipton KA, Farokhyfar M, Burd EM, Weiss DS, Rather PN. A high-frequency phenotypic switch links bacterial virulence and environmental survival in Acinetobacter baumannii. Nat Microbiol. 2018; 3(5):563–9. https://doi.org/10.1038/s41564-018-0151-5 PMID: 29693659

68. Turner KH, Vallet-Gely I, Dove SL. Epigenetic Control of Virulence Gene Expression in Pseudomonas aeruginosa by a LysR-Type Transcription Regulator. PLoS Genet. 2009 Dec 18; 5(12):e1000779. https://doi.org/10.1371/journal.pgen.1000779 PMID: 20041030

69. Konstantinidis KT, Ramette A, Tiedje JM. The bacterial species definition in the genomic era. Philos Trans R Soc Lond B Biol Sci. 2006 Nov 29; 361(1475):1929–40. https://doi.org/10.1098/rstb.2006.1920 PMID: 17062412

70. Cohan FM. Transmission in the Origins of Bacterial Diversity, From Ecotypes to Phyla. Microbiol Spectr. 2017 Oct 3; 5(5):1–26.

71. Mayo AE, Setty Y, Shavit S, Zaslaver A, Alon U. Plasticity of the *cis*-Regulatory Input Function of a Gene. PLoS Biol. 2006 Mar 28; 4(4):e45. https://doi.org/10.1371/journal.pbio.0040045 PMID: 16602820

72. Klemm EJ, Gkrania-Klotsas E, Hadfield J, Forbester JL, Harris SR, Hale C, et al. Emergence of host-adapted *Salmonella* Enteritidis through rapid evolution in an immunocompromised host. Nat Microbiol. 2016 Mar; 1(3):15023. https://doi.org/10.1038/nmicrobiol.2015.23 PMID: 27127642

73. Thiennimitr P, Winter SE, Bäumler AJ. *Salmonella*, the host and its microbiota. Curr Opin Microbiol. 2012 Feb 1; 15(1):108–14. https://doi.org/10.1016/j.mib.2011.10.002 PMID: 22030447

74. Veening J-W, Smits WK, Kuipers OP. Bistability, Epigenetics, and Bet-Hedging in Bacteria. Annu Rev Microbiol. 2008 Oct; 62(1):193–210.

75. Ackermann M. A functional perspective on phenotypic heterogeneity in microorganisms. Nat Rev Microbiol. 2015 Aug; 13(8):497–508. https://doi.org/10.1038/nrmicro3491 PMID: 26145732

76. Ahmad I, Lamprokostopoulou A, Le Guyon S, Streck E, Barthel M, Peters V, et al. Complex c-di-GMP Signaling Networks Mediate Transition between Virulence Properties and Biofilm Formation in Salmonella enterica Serovar Typhimurium. PLoS ONE. 2011 Dec 2; 6(12):e28351–15. https://doi.org/10.1371/journal.pone.0028351 PMID: 22164276

77. Lamprokostopoulou A, Monteiro C, Rhen M, Römling U. Cyclic di-GMP signalling controls virulence properties of *Salmonella enterica* serovar Typhimurium at the mucosal lining. Environ Microbiol. 2010 Jan; 12(1):40–53. https://doi.org/10.1111/j.1462-2920.2009.02032.x PMID: 19691499

78. Pontes MH, Lee E-J, Choi J, Groisman EA. *Salmonella* promotes virulence by repressing cellulose production. Proc Natl Acad Sci USA. 2015 Apr 21; 112(16):5183–8. https://doi.org/10.1073/pnas.1500989112 PMID: 25848006

79. Oppong GO, Rapsinski GJ, Newman TN, Nishimori JH, Biesecker SG, Tükel Ç. Epithelial Cells Augment Barrier Function via Activation of the Toll-Like Receptor 2/Phosphatidylinositol 3-Kinase Pathway

upon Recognition of *Salmonella enterica* Serovar Typhimurium Curli Fibrils in the Gut. Infect Immun. 2013 Feb 1; 81(2):478. https://doi.org/10.1128/IAI.00453-12 PMID: 23208603

80. Tükel Ç, Raffatellu M, Humphries AD, Wilson RP, Andrews-Polymenis HL, Gull T, et al. CsgA is a pathogen-associated molecular pattern of *Salmonella enterica* serotype Typhimurium that is recognized by Toll-like receptor 2. Mol Microbiol. 2005 Aug 26; 58(1):289–304. https://doi.org/10.1111/j.1365-2958.2005.04825.x PMID: 16164566

81. Tükel Ç, Nishimori JH, Wilson RP, Winter MG, Keestra AM, Van Putten JPM, et al. Toll-like receptors 1 and 2 cooperatively mediate immune responses to curli, a common amyloid from enterobacterial biofilms. Cell Microbiol. 2010 May 18; 12(10):1495–505. https://doi.org/10.1111/j.1462-5822.2010.01485.x PMID: 20497180

82. Rapsinski GJ, Wynosky-Dolfi MA, Oppong GO, Tursi SA, Wilson RP, Brodsky IE, et al. Toll-Like Receptor 2 and NLRP3 Cooperate To Recognize a Functional Bacterial Amyloid, Curli. Infect Immun. 2015 Jan 14; 83(2):693–701. https://doi.org/10.1128/IAI.02370-14 PMID: 25422268

83. Nishimori JH, Newman TN, Oppong GO, Rapsinski GJ, Yen J-H, Biesecker SG, et al. Microbial Amyloids Induce Interleukin 17A (IL-17A) and IL-22 Responses via Toll-Like Receptor 2 Activation in the Intestinal Mucosa. Infect Immun. 2012 Dec 1; 80(12):4398. https://doi.org/10.1128/IAI.00911-12 PMID: 23027540

84. Amarasinghe JJ, D 039 Hondt RE, Waters CM, Mantis NJ. Exposure of Salmonella enterica Serovar Typhimurium to a Protective Monoclonal IgA Triggers Exopolysaccharide Production via a Diguanylate Cyclase-Dependent Pathway. Infect Immun. 2013 Mar 1; 81(3):653. https://doi.org/10.1128/IAI.00813-12 PMID: 23230292

85. Smith RD, Price EJ, Burby EP, Blanco PL, Chamberlain J, Chapman RM. The Production of Curli Amyloid Fibers Is Deeply Integrated into the Biology of *Escherichia coli*. Biomolecules. 2017; 7(4).

86. Anriany YA, Weiner RM, Johnson JA, De Rezende CE, Joseph SW. *Salmonella enterica* Serovar Typhimurium DT104 Displays a Rugose Phenotype. Appl Environ Microbiol. 2001 Sep 1; 67(9):4048. https://doi.org/10.1128/AEM.67.9.4048-4056.2001 PMID: 11526004

87. Scher K, Römling U, Yaron S. Effect of Heat, Acidification, and Chlorination on *Salmonella enterica* Serovar Typhimurium Cells in a Biofilm Formed at the Air-Liquid Interface. Appl Environ Microbiol. 2005 Mar 1; 71(3):1163. https://doi.org/10.1128/AEM.71.3.1163-1168.2005 PMID: 15746314

88. Finn S, Condell O, McClure P, Amézquita A, Fanning S. Mechanisms of survival, responses and sources of Salmonella in low-moisture environments. Front Microbiol. 2013; 4:331. https://doi.org/10.3389/fmicb.2013.00331 PMID: 24294212

89. Kariuki S, Onsare RS. Epidemiology and Genomics of Invasive Nontyphoidal *Salmonella* Infections in Kenya. Clin Infect Dis. 2015 Nov 1; 61 Suppl 4(suppl 4):S317–24.

90. Kariuki S, Revathi G, Kariuki N, Kiiru J, Mwituria J, Muyodi J, et al. Invasive multidrug-resistant non-typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic transmission? J Med Microbiol. 2006; 55(5):585–91.

91. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012 Jun 11; 28(12):1647–9. https://doi.org/10.1093/bioinformatics/bts199 PMID: 22543367

92. Martínez-García E, de Lorenzo V. Engineering multiple genomic deletions in Gram-negative bacteria: analysis of the multi-resistant antibiotic profile of *Pseudomonas putida* KT2440. Environ Microbiol. 2011 Aug 24; 13(10):2702–16. https://doi.org/10.1111/j.1462-2920.2011.02538.x PMID: 21883790

93. Silva-Rocha R, Martínez-García E, Calles B, Chavarría M, Arce-Rodríguez A, las Heras de A, et al. The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. Nucleic Acids Res. 2012 Nov 22; 41(D1):D666–75.

94. Wang Y, Chen X, Hu Y, Zhu G, White AP, Köster W. Evolution and Sequence Diversity of FhuA in *Salmonella* and *Escherichia*. Infect Immun. 2018 Nov; 86(11):181–12.