Genetic plasticity of the \textit{Shigella} virulence plasmid is mediated by intra- and inter-molecular events between insertion sequences

Giulia Pilla, Gareth McVicker, Christoph M. Tang*  
Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom  
* christoph.tang@path.ox.ac.uk

Abstract

Acquisition of a single copy, large virulence plasmid, pINV, led to the emergence of \textit{Shigella} spp. from \textit{Escherichia coli}. The plasmid encodes a Type III secretion system (T3SS) on a 30 kb pathogenicity island (PAI), and is maintained in a bacterial population through a series of toxin:antitoxin (TA) systems which mediate post-segregational killing (PSK). The T3SS imposes a significant cost on the bacterium, and strains which have lost the plasmid and/or genes encoding the T3SS grow faster than wild-type strains in the laboratory, and fail to bind the indicator dye Congo Red (CR). Our aim was to define the molecular events in \textit{Shigella flexneri} that cause loss of Type III secretion (T3S), and to examine whether TA systems exert positional effects on pINV. During growth at 37˚C, we found that deletions of regions of the plasmid including the PAI lead to the emergence of CR-negative colonies; deletions occur through intra-molecular recombination events between insertion sequences (ISs) flanking the PAI. Furthermore, by repositioning MvpAT (which belongs to the VapBC family of TA systems) near the PAI, we demonstrate that the location of this TA system alters the rearrangements that lead to loss of T3S, indicating that MvpAT acts both globally (by reducing loss of pINV through PSK) as well as locally (by preventing loss of adjacent sequences). During growth at environmental temperatures, we show for the first time that pINV spontaneously integrates into different sites in the chromosome, and this is mediated by inter-molecular events involving \textit{IS1294}. Integration leads to reduced PAI gene expression and impaired secretion through the T3SS, while excision of pINV from the chromosome restores T3SS function. Therefore, pINV integration provides a reversible mechanism for \textit{Shigella} to circumvent the metabolic burden imposed by pINV. Intra- and inter-molecular events between ISs, which are abundant in \textit{Shigella} spp., mediate plasticity of \textit{S. flexneri} pINV.

Author summary

\textit{Shigella flexneri} is the leading cause of bacillary dysentery worldwide. Key to its virulence is a large 210 kb single copy plasmid, pINV, which encodes a Type III Secretion System (T3SS) on a 30 kb pathogenicity island (PAI). When \textit{S. flexneri} is grown on solid media...
containing Congo red (CR), virulent, T3SS-expressing colonies appear red (CR+). Colonies of bacteria are white and large (CR-) if they lose T3SS expression; thus, the T3SS imposes a significant metabolic burden on S. flexneri. Within the laboratory, spontaneous emergence of CR- colonies is observed, but the molecular events responsible have not been defined. We characterised CR- bacteria that arise during growth at 37°C and 21°C, and demonstrate that recombination between insertion sequences (ISs) on pINV results in loss of the PAI. Furthermore, we demonstrate that MvpAT, a member of the VapBC family of toxin:antitoxin systems encoded on pINV, is responsible for both plasmid maintenance through post-segregational killing, and retention of adjacent sequences. We show for the first time that ISs on the plasmid and chromosome mediate inter-molecular recombination events, resulting in spontaneous and reversible integration of pINV into the chromosome; following integration, T3SS expression is down-regulated. Therefore, integration/excision results in phenotypic heterogeneity that provides a bet-hedging strategy for Shigella to circumvent the metabolic burden associated with retaining virulence genes.

Introduction

The genus Shigella is a major cause of diarrhoeal disease worldwide, and is responsible for around 188 million cases and 600,000 deaths each year [1, 2]. Most infections occur in low income countries where contaminated water and inadequate sanitation promote the transmission of the bacterium [2, 3]. Shigella is a human-specific pathogen that is divided into four species: Shigella dysenteriae, Shigella flexneri, Shigella sonnei and Shigella boydii [3]. Although the prevalence of each species depends on the geographic region, S. flexneri remains the leading cause of endemic shigellosis worldwide [2].

The four species of Shigella have emerged from Escherichia coli following the acquisition of a large plasmid, pINV, a 213 kb element that is essential for virulence [4]. pINV is a single copy, non-conjugative element that consists of a patchwork of pathogenesis-associated and plasmid maintenance genes, separated by regions of repeated sequences such as insertion sequence (IS) elements [5]. Indeed, ISs are highly abundant in S. flexneri, and account for 53% of pINV-encoded genes and 6.7% of all chromosomal sequence [5, 6].

Genes present on pINV enable the bacterium to invade intestinal epithelial cells, escape into the host cell cytosol, undergo cell-to-cell spread, and induce pyroptosis in macrophages [3, 7, 8]. Most of the virulence genes on pINV are located in a 30 kb pathogenicity island (PAI), which encodes components of a Type III Secretion System (T3SS), a molecular syringe that delivers bacterial effector proteins into the host cytoplasm [5, 9], with most secreted effectors also encoded by genes in the PAI. Expression of the T3SS is highly regulated and responds to specific environmental cues such as temperature [10], pH [11], osmolarity [10], oxygen [12], and iron concentrations [13]. Temperature is a key signal for Shigella [14], as it distinguishes between free-living and host-associated environments. The Shigella T3SS is activated at temperatures found in the gastrointestinal tract [14]; a rise in temperature to 37°C relieves H-NS repression of the pINV-encoded regulator VirF [15]. In turn, VirF activates the expression of another regulator, VirB, which is encoded on the PAI and controls expression of genes for the T3SS and its effectors [16].

For any single copy plasmid, its replication must be matched with the division of the chromosome, and active partitioning systems are needed to ensure that each daughter cell receives a copy of the plasmid on division. Furthermore plasmids can be maintained in bacterial populations through post-segregational killing (PSK) mechanisms, typically consisting of toxin:antitoxin (TA) systems which eliminate cells lacking a plasmid after division.
*S. flexneri* pINV possesses specific systems to prevent plasmid loss. To date, two partitioning systems, ParAB and StbAB, have been identified by sequence analysis [5], and three functional TA systems, MvpAT, CcdAB and GmvAT have been characterised in more detail [17–19]. Type II TA systems, such as those found on pINV, are composed of genes encoding a toxic protein and a protein antidote. In general, the toxin is more stable than the antitoxin, with the antitoxin specifically degraded by proteases belonging to Lon or Clp families [20]. Typically, the antitoxin is produced at higher levels so that once degraded, it is rapidly replenished [21]. In the presence of the plasmid, the antitoxin counteracts the activity of the toxin, preventing cell death. However, once the plasmid is lost, the relatively unstable antitoxin is degraded and no longer replaced, leaving the toxin free to arrest cell growth, resulting in PSK [22]. MvpAT is the most characterized TA system on pINV and belongs to the VapBC family of TA systems [17–19]. The toxin MvpT is a site-specific endonuclease that stalls translation by cleaving tRNA\(^{f_{MET}}\) [23]. MvpAT is essential for pINV maintenance at 37°C, while GmvAT confers pINV stability at environmental temperatures [19]. However, aside from the influence of temperature, the need for multiple TA systems on pINV and other plasmids remains unclear.

Colonies of virulent *S. flexneri* expressing a T3SS bind Congo red (CR) when grown on solid media containing this dye, giving rise to a CR\(^+\) phenotype [24]. As the T3SS is specifically expressed at 37°C [10, 25], binding to CR is only evident for colonies cultured at this temperature. In the laboratory, *S. flexneri* can spontaneously lose expression of its T3SS, resulting in white, avirulent colonies (CR\(^-\) phenotype) [26]. Of note, expression of the T3SS represents a high metabolic burden for *Shigella*, evident from the higher growth rate of CR\(^-\) strains compared with CR\(^+\) bacteria at 37°C [26, 27]. Although CR binding is widely used to distinguish between virulent and non-virulent *Shigella* [14], the molecular events that lead to *S. flexneri* becoming CR\(^-\) have not been characterised, although examples of segregational instability (i.e. loss of the entire plasmid), or structural instability (i.e. undefined deletions and/or rearrangements of pINV) have been described [14, 26–28].

Our aim was to define the genetic events that underlie the plasticity of pINV leading to the appearance of CR\(^-\) colonies during growth of *S. flexneri* at host and environmental temperatures. The majority of events occurring at 37°C result from structural instability of pINV, resulting in loss of the T3SS following intra-molecular events between ISs flanking the PAI. We also show that the TA system *mvpAT* not only contributes to segregational stability of pINV as described previously [19, 29], but also exerts local effects and prevents the loss of adjacent sequences. During growth at 21°C, we show for the first time that pINV spontaneously integrates into the chromosome. Chromosomal integration of *S. flexneri* pINV has been described previously [30], but only following exposure of bacteria with curing agents, or introduction of an incompatible plasmid. Furthermore, we observed that spontaneous integration occurs by inter-molecular events between copies of IS1294 present on pINV and the chromosome, and leads to reduced expression of the T3SS regulatory cascade (virF and virB); CR\(^+\) revertants in which the plasmid had excised were recovered from strains with pINV integration. Therefore, integration provides a reversible mechanism for expression of the T3SS. Our findings highlight the importance of ISs in remodelling of *S. flexneri* pINV, and provide a framework for understanding changes in plasmid behaviour that influence the evolution and maintenance of virulence in this important human pathogen.

**Results**

*MvpAT contributes to retention of pINV at 37°C*

To examine the impact of MvpAT on the nature of pINV instability in *S. flexneri*, we introduced a point mutation into *mvpT* (*mvpTD7A*) to abolish its activity while leaving *mvpAT* in its...
original position on pINV [23] with cat marker downstream of mvpAT for selection; the strain was designated native mvpAT D7A, with “native” referring to its location on pINV. We also constructed an isogenic strain, native mvp WT, with the resistance marker in the same location, but downstream of a wild-type copy of mvpAT. The emergence of CR colonies from S. flexneri M90T, native mvp D7A, and native mvp WT was assessed by measuring the proportion of CR colonies after approximately 50 generations of growth at 37˚C, the temperature of the human intestinal tract.

There was no significant difference in the number of CR bacteria arising from native mvp WT and wild-type S. flexneri M90T, demonstrating that introduction of the cat cassette has no impact on the rate of emergence of CR bacteria (Fig 1A). Furthermore, consistent with previous work [19, 26], inactivation of mvpAT resulted in a significant increase in the proportion of CR colonies after 50 generations, which reached almost 90% in native mvp D7A; the average loss of CR binding did not exceed 60% of colonies of M90T or native mvp WT following the same number of generations (native mvp WT vs. native mvp D7A, p < 0.001, Fig 1A).

Fig 1. Emergence of CR colonies in S. flexneri results from loss of the PAI at 37˚C. (A) Proportion of CR colonies in S. flexneri M90T, native mvp WT and native mvp D7A after approximately 50 generations at 37˚C. Solid line: mean (n = six biological replicates). (B) Percentage of CR colonies lacking specified virulence-related genes. Loss of the entire plasmid is inferred by loss of the origin of replication; “others” refers to CR colonies that contain virB, virF and the origin of replication. Eight independent CR colonies obtained from native mvp WT grown in six biological repeats (total 48 colonies) were analysed by multiplex PCR. *** p < 0.001; **** p < 0.0001; n.s., not significant; values analysed with one-way ANOVA, Tukey multiple comparisons test. (C) BLAST Ring Image Generator (BRIG 0.95 and BLASTN v2.2.29) alignment of plasmid sequences from twenty CR colonies that had emerged from native mvp WT at 37˚C; each ring represents the plasmid from an independent CR colony. S. flexneri M90T pWR100 (inner black ring) shown as the reference.

https://doi.org/10.1371/journal.pgen.1007014.g001
Next we assessed whether MvpAT has any effect on the nature of pINV instability. Eight CR colonies arising from native \( mvp^{WT} \) or native \( mvp^{D7A} \) were isolated on six separate occasions (i.e. a total of 48 colonies), and examined by PCR for the presence of \( virB \), \( virF \), and the \( ori \) on pINV, and \( hns \) as a chromosomal control (Fig 1B). Loss of either \( virB \) and \( virF \) is sufficient to render \( Shigella \) CR [26], and \( virB \) is located within the T3SS PAI [5]. Amplification of the replication origin (\( ori \)) was used to monitor the presence of pINV. We found that \( virB \) was the only gene not amplified by multiplex PCR from 92% of CR colonies emerging from \( S. flexneri \) M90T (Fig 1B). A similar result was obtained for native \( mvp^{WT} \), demonstrating that the presence of \( cat \) downstream of \( mvpAT \) does not affect the nature of emerging CR colonies (loss of \( virB \) in \( S. flexneri \) M90T vs. native \( mvp^{WT} \), \( p > 0.9999 \)). However, the profile of CR bacteria arising from native \( mvp^{D7A} \) was distinct from \( S. flexneri \) M90T and native \( mvp^{WT} \); loss of the entire plasmid was the major cause of loss of CR binding in native \( mvp^{D7A} \), accounting 92% of CR colonies (Fig 1B; native \( mvp^{D7A} \) vs. native \( mvp^{WT} \), \( p < 0.0001 \)); loss of \( virB \) alone (as detected by multiplex PCR) was associated with only 8% of CR colonies in native \( mvp^{D7A} \).

Taken together, these data confirm that \( mvpAT \) is fundamental for plasmid stability at 37°C [19, 29] and contributes to the maintenance of the entire plasmid, consistent with its role in PSK. Inactivation of MvpAT by introducing a non-toxic allele of \( mvpT \) leads to a dramatic increase in the CR phenotype, mainly due to loss of pINV (i.e. segregational instability). However, in contrast to a previous study of a different \( S. flexneri \) strain [26], we found that wild-type M90T \( S. flexneri \) CR bacteria mostly arise following deletions involving the PAI rather than loss of the entire plasmid.

**Spontaneous loss of the T3SS PAI occurs by at least four distinct routes**

To further characterise the molecular events responsible for the emergence of CR \( S. flexneri \), we performed whole genome sequencing of 20 CR colonies independently derived from native \( mvp^{WT} \) at 37°C (Fig 1C); plasmid sequences were aligned with BLAST Ring Image Generator with pINV from \( S. flexneri \) M90T as the reference [9]. Results confirmed that CR binding is mainly lost at 37°C following loss of the T3SS PAI and not loss of pINV (Fig 1C). Out of 20 CR colonies, 17 lacked sequences unique to the PAI, with the remaining three displaying deletions involving the PAI rather than loss of the entire plasmid.

To define the sequences mediating loss of the PAI, we performed PCR to amplify regions flanking the deletions (Fig 2 and S3 Table for primer sequences), and analysed the products by restriction enzyme digestion and sequencing. We identified four events leading to loss of the PAI, each involving a pair of homologous ISs (Fig 2A). The most frequent deletion (Variant 1) occurred in 11 out of 17 CR colonies and involved two copies of IS\( Sf l 4 \), spanning positions 87,663–90,277 bp and 164,461–167,075 bp of pINV [9]. The second most frequent deletion event (Variant 2) occurred in four of colonies and included two copies of IS\( 1 2 9 4 \), located at positions 58,970–60,658 bp and 134,268–135,956 bp. Variant 3 and 4 each accounted for a single CR colony and employed copies of IS\( 6 0 0 \) (83,866–85,129 bp and 132,311–133,574 bp) and IS\( 1 2 9 4 \) (48,065–49,417 bp and 134,268–135,956 bp), respectively. Taken together these results demonstrate that ISs on pINV are “hot spots” for recombination, resulting in deletion of different regions of the plasmid that include the PAI (S1 Table).

**MvpAT governs loss of local sequences**

The reason why \( S. flexneri \) pINV has three functional TA systems is unclear [19]. As we found that distinct regions on pINV can be lost by IS-mediated recombination, we hypothesised that...
the presence of multiple TA systems could be due to a positional effect of these elements on plasmid dynamics. To assess whether the position of mvpAT influences plasmid stability, we deleted mvpAT from pINV and introduced either a wild-type or a non-functional version (mvpD7A) of mvpAT adjacent to the T3SS PAI, at nt. 100,792 of pINV [9], generating ectopic mvpWT and ectopic mvpD7A, respectively (Fig 3A). Similar to native mvpD7A, when the non-functional allele was introduced into the ectopic site, the inactivity of MvpAT led to a large population of CR- bacteria emerging after 50 generations (~ 80% of the total population, Fig 3B, ectopic mvpWT vs. ectopic mvpD7A, p < 0.0001), most of which resulted from loss of the entire plasmid. This demonstrates that MvpAT contributes to retention of pINV, irrespective of its position on the plasmid.

However, there was a dramatic reduction in the number of CR- bacteria emerging from ectopic mvpWT compared with native mvpWT at 37°C (Fig 3B, p < 0.0001). Indeed, over approximately 50 generations, the proportion of CR- colonies in ectopic mvpWT did not exceed 5%, compared with ~ 50% for native mvpWT (Fig 3B). Furthermore, there was a striking change in the events leading to CR- colonies when mvpAT is positioned next to the PAI. By
multiplex PCR, loss of *virB* represented a significantly lower proportion of CR’ colonies emerging from ectopic *mvp*<sup>WT</sup> compared with native *mvp*<sup>WT</sup> (Fig 3C, *p* < 0.0001), and accounted for only 8% of CR’ colonies emerging from ectopic *mvp*<sup>WT</sup>. Instead, loss of *virF* was the most frequent event, accounting for 71% of CR’ derivatives (ectopic *mvp*<sup>WT</sup> vs. native *mvp*<sup>WT</sup>, *p* < 0.0001). However, differences in the proportion of genetic changes leading to loss of CR binding observed for native *mvp*<sup>WT</sup> and ectopic *mvp*<sup>WT</sup> need to be considered in the context of the rate of emergence of CR’ colonies in these strains. The absolute values for the loss of *virF* in ectopic *mvp*<sup>WT</sup> are comparable to those observed for native *mvp*<sup>WT</sup> (3.55% and 2% of all colonies, respectively). Thus, the position of *mvp* on pINV affects loss of *virB* but not loss of *virF*, demonstrating that *mvp* influences structural plasmid stability by acting to retain nearby sequences.

**IS-mediated deletions and plasmid insertion lead to CR’ colonies at environmental temperatures**

We also examined the emergence of CR’ colonies at 21˚C, the temperature the bacterium faces in the external environment during host-to-host transmission. Bacteria were grown in liquid media at 21˚C for approximately 50 generations, then aliquots were plated on solid media containing CR and incubated at 37˚C, as the CR binding is visible at 37˚C but not at 21˚C.
found previously that, although there is no difference in the stability of pINV in *S. flexneri* at 37˚C and at 21˚C [19], a large population of CR-bacteria emerges at the higher temperature from wild-type *S. flexneri* (Fig 1) because of the growth advantage of CR-colonies at 37˚C. As expected, when *S. flexneri* M90T was grown at 21˚C, fewer CR-colonies emerged after 50 generations than at 37˚C, and accounted for approximately 2% of the total population (Fig 4A); again results for native *mvp*WT were indistinguishable from *S. flexneri* M90T, confirming that the presence of the *cat* cassette does not affect plasmid stability (S1 Fig). Furthermore, there was no significant difference in the number of CR-bacteria emerging from ectopic *mvp*WT and native *mvp*WT (p = 0.6603). However, there was a slight but significant increase in the proportion of CR-colonies emerging from strains with the inactive *mvp*DTA allele compared with control strains (Fig 4A, native *mvp*WT vs. native *mvp*DTA, p = 0.0449; ectopic *mvp*WT vs. ectopic *mvp*DTA, p = 0.0397), consistent with our previous findings [19] that MvpAT possesses some residual activity at environmental temperatures.

Next, we examined the nature of CR-bacteria emerging at 21˚C by multiplex PCR (Fig 4B). Similar to 37˚C, loss of *virB* still prevailed in native *mvp*WT, and accounted for 80% of CR-colonies. Furthermore, the majority of CR-bacteria derived from native *mvp*DTA and ectopic *mvp*DTA had lost pINV, accounting for 73% and 55% of CR-colonies, respectively (Fig 4B;
native $mvp^{WT}$ vs. native $mvp^{D7A}$, $p <0.0001$; ectopic $mvp^{WT}$ vs. ectopic $mvp^{D7A}$, $p = 0.0454$), demonstrating that MvpAT prevents plasmid loss at 21°C as well as 37°C. However, we found that a considerable number of CR colonies emerging at 21°C from strains from wild-type $mvp^{AT}$ retained virB, virF, and ori (Fig 4B). More than 25% and 50% of all CR colonies derived from native $mvp^{WT}$ and ectopic $mvp^{WT}$, respectively, harboured all three pINV genes showing that alternative mechanisms might be responsible for the emergence of CR bacteria at 21°C. Illumina sequencing of nine CR colonies that emerged independently at 21°C from native $mvp^{WT}$ and ectopic $mvp^{WT}$, respectively, harboured all three pINV genes showing that alternative mechanisms might be responsible for the emergence of CR bacteria at 21°C.

Illumina sequencing of nine CR colonies that emerged independently at 21°C from native $mvp^{WT}$ and ectopic $mvp^{WT}$ confirmed that these strains harboured an intact plasmid with no detectable single nucleotide polymorphisms (S2 Fig).

pINV reversibly integrates into the chromosome

To identify the mechanisms by which CR bacteria emerge at 21°C, we analysed the nine CR colonies with intact pINV sequence which arose from native $mvp^{WT}$ (Fig 4B). It has previously been shown that pINV of S. flexneri M90T can integrate into the host chromosome [30], but only following exposure of bacteria to curing agents, or selecting for acquisition of a plasmid with an ori which is incompatible with pINV. Therefore, we hypothesised that plasmid integration could occur spontaneously, so analysed plasmid DNA extracted from the nine CR isolates. Results demonstrate that a distinct band for pINV was missing in two out of nine CR strains (Fig 5A, isolates 4 and 6). As previous Illumina sequence analysis confirmed that neither strain had lost any pINV sequence (S2 Fig), isolates 4 and 6 were subjected to PacBio sequencing. In both strains, we found an intact copy of pINV integrated into the chromosome via different copies of IS1294 (Fig 5B and S3 Fig); there are six copies of IS1294 on the plasmid and three copies on the chromosome. Integration in strain 4 involved the chromosomal copy of IS1294 at position 1,615,894–1,617,578 bp [31] and pINV IS1294 at 58,970–60,658 bp [9] (i.e. IS1294$_{pINV}$ and IS1294$_{ch,1}$, respectively, S3 Fig); for isolate 6, chromosomal IS1294 at position 1,872,287–1,874,073 bp [31] and plasmid IS1294 at position 205,186–206,599 bp are involved (IS1294$_{pINV}$ and IS1294$_{ch,3}$, respectively, S3 Fig) [9].

Furthermore, we investigated whether integration of pINV is a reversible event. We examined isolate 4 and 6 for the emergence of CR+ colonies, as phenotypic evidence for plasmid excision. The isolates were grown overnight in liquid media at 37°C and plated to solid media containing CR. CR+ colonies emerged from both isolates (Fig 6A), and we analysed plasmid DNA from two independent CR+ revertants (Fig 6B). Results confirm that pINV not only integrates into the host chromosome, but can also subsequently excise, restoring the CR+ phenotype (Fig 6A).

Chromosomal integration of pINV down-regulates expression of PAI genes

Next we examined the activity of the T3SS in isolates 4 and 6, and their corresponding revertants. Bacteria were grown to exponential phase at 37°C in liquid media and exposed to CR to induce secretion through the T3SS [32]. Silver staining of secreted proteins demonstrates that both isolates 4 and 6 fail to secrete the T3SS effectors, IpaA, IpaB, IpaC, IpaD and IpgD (Fig 6C). In contrast, the revertants secrete T3SS effectors at levels similar to S. flexneri M90T (Fig 6C), indicating that integration and excision of pINV provides a reversible mechanism that controls T3SS activity.

To determine the mechanisms underlying the lack of secretion through the T3SS in the pINV-integrated isolates, we examined mRNA levels of ipaB, virB and virF in strains 4 and 6, and wild-type S. flexneri M90T. mRNA levels were measured by qRT-PCR in bacteria during exponential growth at 37°C, and results were normalized to the expression of the chromosomal
gene, polA (Fig 7). Consistent with the secretion assays, mRNA levels of ipaB were significantly lower in the two strains with plasmid integration (Fig 7A, M90T vs. strain 4 or strain 6, p < 0.0001). A similar statistically significant trend was shown for mRNA levels of virB (Fig 7B, M90T vs. strain 4 or strain 6, p < 0.0001) and virF (Fig 7C, M90T vs. strain 4 p < 0.01; M90T vs. strain 6, p < 0.001), suggesting that the observed down-regulation of the T3SS PAI in plasmid-integrated isolates is correlated with reduced expression of virB and virF.

**Discussion**

For many pathogenic bacteria, large plasmids are critical for their virulence and/or the spread of antimicrobial resistance [33–36]. While plasmids confer beneficial traits to bacteria in
certain circumstances, they often impose a considerable metabolic cost on the host cell so have evolved dedicated mechanisms to ensure their maintenance within a bacterial population. Genetic plasticity is a fundamental feature of many large plasmids, and facilitates rearrangements that allow appropriate gene expression and the acquisition of novel traits such as antibiotic resistance, virulence or metabolic capabilities [37], facilitating adaptation to new ecological niches. *S. sonnei* pINV is an example of how genetic events contribute to the evolution of the plasmid, reflecting bacterial adaptation to new lifestyles [19]. For example, *S. sonnei* pINV has lost two TA systems and a partitioning system, and acquired an O-antigen gene cluster [19, 38], during its transition to a species undergoing predominant host-to-host transmission [19, 39].
Here we analysed the genetic changes associated with plasticity of pINV. The dynamic nature of S. flexneri pINV is a feature shared with plasmids in other Shigella spp., which contain a high proportion of ISs; ISs represent approximately 53% of the ORFs on pINV [5]. These elements have shaped the evolution of Shigella by acting as substrates for recombination, mediating inversions, translocations, insertions and deletions [37]. On the chromosome, similar processes have led to loss of co-linearity of genomes of four Shigella species, and gene loss associated with enhanced virulence [40, 41]. Nonetheless, the T3SS PAI and most other pINV encoded molecular effectors have been maintained in all Shigella spp., suggesting that selection pressure has preserved these sequences in disease-causing isolates.

We demonstrate that ISs influence the architecture of Shigella pINV and mediate a series of deletions that result in the loss of PAI-associated virulence genes and CR binding during growth in the laboratory. Others have found that the most frequent cause of loss of the CR+ phenotype during extended growth of S. flexneri 2a at 37˚C is loss of the entire plasmid [26]. In contrast, we show that the predominant event leading to loss of CR binding in S. flexneri 5a M90T is deletion of the T3SS PAI (Fig 1), consistent with other studies [27, 28, 42]. This discrepancy could be explained by the different MvpAs in the strains; for example, MvpA in M90T has a single amino acid difference (Glu70 instead of His70) compared with the corresponding protein in the S. flexneri 2a strain used in previous work [26, 43]. However, the consequences of this difference in MvpAT are unknown. We found that ISSfl4 is the most common IS involved in PAI deletion (Fig 2), even though there are only two copies of this IS on pINV. The plasmid harbours multiple copies of other ISs which occupy a far greater proportion of the plasmid sequence than ISSfl4. Therefore, the reason why this pair of ISs is prone to recombination is not clear. The frequency of recombination between pairs of ISs depends on several factors including the extent and length of homology, the coverage of sequence and local DNA topology (S1 Table). In fact, the percentage of coverage between the ISs involved in the less frequent deletions is lower than for the pair of ISs implicated in Variant 1 strains, which have the commonest rearrangement (S1 Table).

Similar to previous work [17, 19, 29], we found that MvpAT plays a critical role in segregational stability of pINV. Inactivation of this TA system led to increased loss of pINV at 37˚C. We also demonstrated that the location of mvpAT is critical in governing plasmid dynamics at a local level; repositioning mvpAT to near the PAI dramatically reduced the loss of this region (Fig 3). The likely reason for this is ‘post-recombinational killing’ whereby recombination
between flanking ISs causes loss of mvpAT and cell death in a manner analogous to PSK. To our knowledge, this is the first description of TA loci exerting local effects on plasmids, although they can prevent large-scale deletions of adjacent sequences on bacterial chromosomes [44, 45]. The localised effects of MvpAT and other TA systems might provide an explanation for the number and distribution of TA systems on pINV. Of note, mvpAT is located close to the ori of pINV in all Shigella species.

Following growth at 21°C, we identified CR− strains emerging from S. flexneri that retained all plasmid genes. By analysis of this population of strains, we demonstrate that pINV can spontaneously integrate into the Shigella chromosome. Previous reports of chromosomal integration of Shigella pINV and the related plasmid from enteroinvasive E. coli have only followed exposing bacteria to curing agents (such as rifampicin) or by introducing other plasmids with pINV-incompatible replicons [30]. Following these artificial treatments, pINV usually integrated into metB, leading to methionine auxotrophy. In contrast, we found that chromosomal and plasmid copies of IS1294 mediate intermolecular recombination, resulting in integration of pINV at different sites in the chromosome. At both chromosomal sites, integration led to loss of CR binding and prevented secretion through the T3SS (Fig 6B). Comparing pINV-integrated strains with the wild-type strain, integration was associated with reduced levels of mRNA for virF and virB, the two transcription factors involved in the activation cascade of T3SS PAI expression, and ipaB, which encodes a T3SS-secreted protein. We also demonstrate that integration is reversible and excision occurs during bacterial growth at 37°C (Fig 5).

The mechanisms responsible for reduced T3SS activity following pINV integration are unknown. H-NS is a transcriptional repressor of the genes in the PAI, which binds to the promoters of virF and virB in a temperature-dependent manner due to alterations in DNA topology [15]. When integration of pINV from S. flexneri and enteroinvasive E. coli was forced and occurred at metB, H-NS repression of the virB promoter was enhanced, probably due to alteration in the topology of pINV DNA when in the chromosome [46]. However, our qRT-PCR results showed that spontaneous integration via IS1294 also affects virF expression (Fig 7C). The E. coli and Salmonella chromosomes are organised into topologically distinct regions, containing structured DNA macrodomains [47–49]. As the architecture of the chromosome is not homogenous and other ISs could also mediate integration, the exact site and orientation of pINV integration are likely to influence gene expression [50]. Further experiments are underway to define the precise mechanisms modulating changes in T3SS expression following spontaneous integration of pINV through ISs.

Other examples of the integration of large plasmids in pathogens include the Salmonella typhimurium virulence plasmid [51]. Following plasmid integration, Salmonella becomes susceptible to complement-mediated killing probably through down-regulation of gene expression. Serum resistance can be restored by introduction of an autonomous plasmid harbouring a copy of rsk, a plasmid gene responsible for resistance to complement [51].

Our findings define the molecular rearrangements that affect S. flexneri pINV, and highlight the importance of IS elements in processes that are fundamental for plasmid remodelling and evolution. We found that intramolecular events between ISs on pINV lead to the emergence of CR− bacteria lacking the T3SS PAI during growth of S. flexneri. CR− bacteria have a significant growth advantage in rich media at 37°C compared with CR+ bacteria, and quickly proliferate and dominate a bacterial population [27]. While the loss of a T3SS could hinder survival of Shigella within the intestine, the fitness advantage upon deletion of the PAI could prove advantageous for Shigella outside the host. Indeed, a significant proportion of Shigella isolated from aquatic environments have lost key virulence genes on pINV [52], consistent with IS-mediated events. However, deletion of the PAI T3SS and loss of virulence, is a one-way process, that can only be reversed by acquisition of a new plasmid. Although S. flexneri pINV has an origin of
transfer, the plasmid is incapable of self-mobilisation [53]. Therefore, IS mediated deletions causing PAI loss are likely to be deleterious in the longer term as they will render the bacterium non-invasive.

In contrast, pINV integration follows an inter-molecular event that is completely reversible. We recovered virulent revertants in which pINV had excised and the T3SS was functional. Therefore, plasmid integration via ISs offers *S. flexneri* an alternative strategy to maintain the plasmid within dividing bacteria while circumventing the fitness costs imposed by expression of the T3SS. pINV integration/excision leads to bi-stable expression of virulence genes by *Shigella*, which provides a mechanism for the avoidance of host responses against immunogenic T3SS components [54], and for phenotypic heterogeneity without genetic loss which would promote the evolutionary stability of virulent *Shigella*.

**Materials and methods**

**Strains and growth media**

Bacterial strains used for this study are shown in S2 Table. *S. flexneri* was grown in Tryptic Soy Broth (TSB; Sigma Aldrich, St. Louis, USA) or on Tryptic Soy solid media containing 1.5% (w/v) agar (Oxoid, Basingstoke, UK) (TSA). Chloramphenicol was used as appropriate at a final concentration of 20 μg/mL. Congo red (Sigma Aldrich, St. Louis, USA) was added to TSA media at a final concentration of 0.01% w/v to make CR-TSA plates.

**Construction of strains**

DNA constructs were ligated into pUC19 using the NEBuilder HiFi master mix (New England Biolabs, NEB, Ipswich, MA), and PCR products were generated using primers shown in S3 Table. Resulting plasmids were transformed into *E. coli* DH5α and linear DNA constructs were amplified by PCR using plasmids as the template. Lambda Red recombination [55, 56] was used to introduce changes into pINV in *S. flexneri* M90T. Approximately 1 kb of homologous flanking sequence was used to allow integration into the plasmid. For strains native mvpAT<sup>Δ7A</sup> and ectopic mvpAT<sup>Δ7A</sup>, a point mutation was introduced into mvpAT by site-directed mutagenesis (S3 Table). Mutations were then transduced into *S. flexneri* using P1vir [19]. In ectopic mvpAT<sup>WT</sup> and ectopic mvpAT<sup>Δ7A</sup>, the native mvpAT locus was deleted and either the wild-type or the mutated mvpAT were positioned between virB and ipaJ at nt. 100,792 on pINV [9].

**CR binding assays**

*S. flexneri* was grown on CR-TSA plates overnight at the selected temperature to obtain single colonies. Three CR<sup>+</sup> colonies from each strain were re-suspended in 5mL TSB and incubated at 30°C, 180 r.p.m.; there is no detectable growth rate difference between CR<sup>+</sup> and CR<sup>−</sup> bacteria at this temperature [19], so CR<sup>−</sup> bacteria do not outcompete CR<sup>+</sup> bacteria. After this initial overnight growth, cultures were grown for approximately 50 generations at either 37°C or 21°C by sub-culture. Samples were diluted in PBS and plated onto CR-TSA, and incubated overnight at 37°C before CR<sup>+</sup> and CR<sup>−</sup> colonies were counted.

**Multiplex PCR for genes on pINV**

Multiplex PCR was performed with primers (S3 Table) to amplify virF, virB, mvpAT, and ori to generate products of distinct sizes; hns, a chromosomal gene, was included as a control (S4 Fig). Reactions included Taq polymerase (Sigma Aldrich, St. Louis, USA) with an annealing temperature of 51.2°C and extension time of 1.5 min. Eight CR<sup>−</sup> colonies emerging on six
independent occasions were analysed by multiplex PCR for each strain. Statistical analysis of results was performed using two-way ANOVA with Tukey’s post test comparison, evaluating loss of a single locus alone or in pairwise comparison with other loci.

Genome sequencing

For Illumina sequencing, genomic DNA was purified from bacteria grown overnight in 5 mL TSB using Charge Switch gDNA Mini Bacteria Kit (Thermo Fisher Scientific, MA, USA). Colonies were isolated after plating at 37°C or 21°C and grown at 37°C for 16 hrs or 21°C for 24 hrs, respectively, before DNA isolation. DNA was sequenced at The Wellcome Trust Centre for Human Genetics at the University of Oxford. Sequence data were analysed using Snap Gene for identifying single nucleotide polymorphisms and deletions, while BLAST Ring Image Generator (BRIG) was employed to align plasmid sequences [57]. For both approaches, S. flexneri M90T was used as the reference. For PacBio sequencing [58], DNA was recovered from a loop of colonies using Wizard Genomic DNA Purification kit (Promega, WI, USA). Sequencing was performed at The Earlham Institute, Norwich.

Purification and visualization of pINV, and protein secretion

Isolates were plated on CR-TSA from 15% glycerol stocks, and incubated at 37°C overnight. One colony from each isolate was streaked again onto CR-TSA and grown as above. A loop of each solid culture was then processed for pINV extraction as previously described [59]. DNA samples were run on 0.7% agarose gel at 1 V/cm voltage for approximately 16 hrs.

To activate secretion through the T3SS, bacteria were grown overnight at 37°C in TSB at 180 r.p.m., then sub-cultured to an OD manufacturer’s protocol.

RNA extraction and quantitative real-time PCR

Bacteria were grown at 37°C overnight in TSB then sub-cultured to OD600 = 0.05 into 25mL fresh TSB and grown at 37°C until the OD600 reached ~ 1. A 20mL aliquot of each culture medium was pelleted, resuspended in 460 μL Resuspension Solution (200 μL 20% glucose, 200 μL Tris 25mM pH7.6 10mM EDTA and 60 μL 0.5M EDTA) and lysed by using Lysing Matrix B Tubes (MP Biomedicals) in presence of 500 μL acid phenol. RNA was then purified by TRIzol-chloroform extraction, precipitated with isopropanol and washed in 75% ethanol. RNA samples were subjected to two DNase I treatments (TURBO DNase; Ambion), each followed by further phenol-chloroform purification. RNA quality was checked by gel electrophoresis and NanoDrop analysis. First-strand cDNA was synthesized from 5 μg total RNA by using polA-, ipaB-, virB- and virF-specific primers with a unique 5’ tag sequences not present in S. flexneri M90T (S3 Table). Reverse transcription was performed in presence of actinomycin and samples were then treated with RNase H (Life Technologies) at 37°C for 20 minutes. After purification, cDNA was analysed by quantitative real-time PCR with Power SYBR green PCR master mix (Applied Biosystems) using a gene-specific forward primer and a tag-specific reverse primer to ensure strand-specific cDNA amplification. StepOnePlus real-time PCR
system was used to monitor the reaction. Results represent the average of four biological replicates and were normalised to polA cDNA levels by using the 2-ΔΔCt method [60]. Values are shown in relation to wild type M90T 2-ΔCt levels, which are indicated as 1. Therefore, values less than 1 indicate decreased transcription levels of the target genes in the analysed samples. Statistical analysis was performed using one-way ANOVA with Tukey’s multiple comparison test.

Supporting information

S1 Table. Sequence identity and coverage of ISs involved in PAI deletions. Coverage and identity values were obtained by aligning the two IS copies from each variant by BLASTN. % GC content was calculated for IS and refers to the copy in order of their numbering according to Buchrieser et al. [9].

S2 Table. Strains used in this study.

S3 Table. Primers used in this study.

S1 Fig. CR binding loss in S. flexneri M90T and native mvpWT at 21˚C. (A) Proportion of CR− colonies in S. flexneri M90T and native mvpWT (reproduced from Fig 4A for statistical comparison) relative to total colonies after approximately 50 generations of growth at 21˚C. Solid line: mean of six biological replicates. (B) Multiplex PCR analysis was performed as described for Fig 1B. Results are shown as mean (n = 6 biological replicates). n.s., not significant; values analysed with one-way ANOVA, Tukey multiple comparisons test.

S2 Fig. Plasmid sequences from native mvpWT and ectopic mvpWT at 21˚C. Alignment of plasmid sequences of 10 independent CR− colonies emerging from native mvpWT (A) and ectopic mvpWT (B) at 21˚C which retained their virulence-related genes tested by multiplex PCR. Images were created as described in Fig 1C.

S3 Fig. IS1294 copies on pINV and the chromosome. The IS elements were numbered clock-wise, using the origin as a starting point [9].

S4 Fig. Example of multiplex PCR gel image of the positive and negative control. Paired adjacent columns correspond to a single assayed colony: pINV +, amplicons originating from S. flexneri M90T genomic DNA; pINV−, amplicons originating from S. flexneri BS176 genomic DNA. Each amplified locus is indicated in white letters. The sizes of a kb marker are shown.

Author Contributions

Conceptualization: Giulia Pilla, Gareth McVicker, Christoph M. Tang.
Data curation: Giulia Pilla.
Formal analysis: Giulia Pilla, Gareth McVicker, Christoph M. Tang.
Funding acquisition: Christoph M. Tang.
Investigation: Giulia Pilla.
Methodology: Giulia Pilla, Gareth McVicker.
Project administration: Christoph M. Tang.
Supervision: Gareth McVicker, Christoph M. Tang.
Validation: Christoph M. Tang.
Writing – original draft: Giulia Pilla, Christoph M. Tang.
Writing – review & editing: Giulia Pilla, Gareth McVicker, Christoph M. Tang.

References
1. Lima I.F., Havt A., and Lima A.A., Update on molecular epidemiology of Shigella infection. Curr Opin Gastroenterol, 2015. 31(1): p. 30–7. https://doi.org/10.1097/MOG.0000000000000136 PMID: 25394237
2. Kotloff K.L., et al., Global burden of Shigella infections: implications for vaccine development and implementation of control strategies. Bull World Health Organ, 1999. 77(8): p. 651–66. PMID: 10516787
3. Schroeder G.N. and Hilbi H., Molecular pathogenesis of Shigella spp.: controlling host cell signaling, invasion, and death by type III secretion. Clin Microbiol Rev, 2008. 21(1): p. 134–56. https://doi.org/10.1128/CMR.00032-07 PMID: 18202440
4. Lan R., et al., Molecular evolution of large virulence plasmid in Shigella clones and enteroinvasive Escherichia coli. Infect Immun, 2001. 69(10): p. 6303–9. https://doi.org/10.1128/IAI.69.10.6303-6309.2001 PMID: 11553574
5. Venkatesan M.M., et al., Complete DNA sequence and analysis of the large virulence plasmid of Shigella flexneri. Infect Immun, 2001. 69(5): p. 3271–85. https://doi.org/10.1128/IAI.69.5.3271-3285.2001 PMID: 11292750
6. Wei J., et al., Complete genome sequence and comparative genomics of Shigella flexneri serotype 2a strain 2457T. Infect Immun, 2003. 71(5): p. 2775–86. https://doi.org/10.1128/IAI.71.5.2775-2786.2003 PMID: 12704152
7. Suzuki T., et al., Differential regulation of caspase-1 activation, pyroptosis, and autophagy via Ipaf and ASC in Shigella-infected macrophages. PLoS Pathog, 2007. 3(8): p. e111. https://doi.org/10.1371/journal.ppat.0030111 PMID: 17696608
8. Zychlinsky A., Prevost M.C., and Sansonetti P.J., Shigella flexneri induces apoptosis in infected macrophages. Nature, 1992. 358(6382): p. 167–9. https://doi.org/10.1038/358167a0 PMID: 1614548
9. Buchrieser C., et al., The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of Shigella flexneri. Mol Microbiol, 2000. 38(4): p. 760–71. PMID: 11115111
10. Porter M.E. and Dorman C.J., A role for H-NS in the thermo-osmotic regulation of virulence gene expression in Shigella flexneri. J Bacteriol, 1994. 176(13): p. 4187–91. PMID: 8021202
11. Nakayama S. and Watanabe H., Involvement of cpxA, a sensor of a two-component regulatory system, in the pH-dependent regulation of expression of Shigella sonnei virF gene. J Bacteriol, 1995. 177(17): p. 5062–9. PMID: 7665485
12. Marteyn B., et al., Modulation of Shigella virulence in response to available oxygen in vivo. Nature, 2010. 465(7326): p. 355–8. https://doi.org/10.1038/nature08970 PMID: 20436458
13. Payne S.M., et al., Iron and pathogenesis of Shigella: iron acquisition in the intracellular environment. Biometals, 2006. 19(2): p. 173–80. https://doi.org/10.1007/s10534-005-4577-x PMID: 16718602
14. Maurelli A.T., Blackmon B., and Curtiss R. 3rd, Temperature-dependent expression of virulence genes in Shigella species. Infect Immun, 1984. 43(1): p. 195–201. PMID: 6360895
15. Falconi M., et al., Thermoregulation of Shigella and Escherichia coli EIEC pathogenicity. A temperature-dependent structural transition of DNA modulates accessibility of virF promoter to transcriptional repressor H-NS. EMBO J, 1998. 17(23): p. 7033–43. https://doi.org/10.1093/emboj/17.23.7033 PMID: 9843508
16. Dorman C.J., Virulence Gene Regulation in Shigella. EcoSal Plus, 2004. 1(1).
17. Sayeed S., et al., The stability region of the large virulence plasmid of Shigella flexneri encodes an efficient postsegregational killing system. J Bacteriol, 2000. 182(9): p. 2416–21. PMID: 10762240
18. Dienemann C., et al., Crystal structure of the VapBC toxin-antitoxin complex from Shigella flexneri reveals a hetero-octameric DNA-binding assembly. J Mol Biol, 2011. 414(5): p. 713–22. https://doi.org/10.1016/j.jmb.2011.10.024 PMID: 22037005
19. McVicker G. and Tang C.M., Deletion of toxin-antitoxin systems in the evolution of Shigella sonnei as a host-adapted pathogen. Nat Microbiol, 2016. 2: p. 16204. https://doi.org/10.1038/nmicrobiol.2016.204 PMID: 27819667

20. Brzozowska I. and Zielenkiewicz U., Regulation of toxin-antitoxin systems by proteolysis. Plasmid, 2013. 70(1): p. 33–41. https://doi.org/10.1016/j.plasmid.2013.01.007 PMID: 23396045

21. Gerdes K. and Maisonneuve E., Bacterial persistence and toxin-antitoxin loci. Annu Rev Microbiol, 2012. 66: p. 103–23. https://doi.org/10.1146/annurev-micro-092611-150159 PMID: 22994490

22. Gerdes K., Jacobsen J.S., and Franch T., Plasmid stabilization by post-segregational killing. Genet Eng (N Y), 1997. 19: p. 49–61.

23. Winther K.S. and Gerdes K., Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. Proc Natl Acad Sci U S A, 2011. 108(18): p. 7403–7. https://doi.org/10.1073/pnas.1019587108 PMID: 21502523

24. Payne S.M. and Finkelstein R.A., Detection and differentiation of iron-responsive avirulent mutants on Congo red agar. Infect Immun, 1977. 18(1): p. 94–8. PMID: 409688

25. Tobe T., et al., Temperature-regulated expression of invasion genes in Shigella flexneri is controlled through the transcriptional activation of the virB gene on the large plasmid. Mol Microbiol, 1991. 5(4): p. 887–93. PMID: 1857209

26. Schuch R. and Maurelli A.T., Virulence plasmid instability in Shigella flexneri 2a is induced by virulence gene expression. Infect Immun, 1997. 65(9): p. 3686–92. PMID: 9284138

27. Sasakawa C., et al., Isolation, characterization, and mapping of Tn5 insertions into the 140-megadalton invasion plasmid defective in the mouse Sereny test in Shigella flexneri 2a. Infect Immun, 1986. 54(1): p. 32–6. PMID: 3019891

28. Sasakawa C., et al., Molecular alteration of the 140-megadalton plasmid associated with loss of virulence and Congo red binding activity in Shigella flexneri. Infect Immun, 1986. 51(2): p. 470–5. PMID: 3002985

29. Sayeed S., et al., Surprising dependence on postsegregational killing of host cells for maintenance of the large virulence plasmid of Shigella flexneri. J Bacteriol, 2005. 187(8): p. 2766–73. https://doi.org/10.1128/JB.187.8.2766-2773.2005 PMID: 15805523

30. Zagaglia C., et al., Virulence plasmids of enteroinvasive Escherichia coli and Shigella flexneri integrate into a specific site on the host chromosome: integration greatly reduces expression of plasmid-carried virulence genes. Infect Immun, 1991. 59(3): p. 792–9. PMID: 1997431

31. Oncdera N.T., et al., Genome sequence of Shigella flexneri serotype 5a strain M90T Sm. J Bacteriol, 2012. 194(11): p. 3022. https://doi.org/10.1128/JB.00393-12 PMID: 22582379

32. Parsot C., et al., Enhanced secretion through the Shigella flexneri Mxi-Spa translocon leads to assembly of extracellular proteins into macromolecular structures. Mol Microbiol, 1995. 16(2): p. 291–300. PMID: 7665091

33. Porse A., et al., Survival and Evolution of a Large Multidrug Resistance Plasmid in New Clinical Bacterial Hosts. Mol Biol Evol, 2016. 33(11): p. 2860–2873. https://doi.org/10.1093/molbev/msw163 PMID: 27501945

34. Johnson T.J. and Nolan L.K., Pathogenomics of the virulence plasmids of Escherichia coli. Microbiol Mol Biol Rev, 2009. 73(4): p. 750–74. https://doi.org/10.1128/MMBR.00015-09 PMID: 19946140

35. Cornelis G.R., et al., The virulence plasmid of Yersinia, an antihost genome. Microbiol Mol Biol Rev, 1998. 62(4): p. 1315–52. PMID: 9841674

36. Rotger R. and Casadesús J., The virulence plasmids of Salmonella. Int Microbiol, 1999. 2(3): p. 177–84. PMID: 10943411

37. Touchon M. and Rocha E.P., Causes of insertion sequences abundance in prokaryotic genomes. Mol Biol Evol, 2007. 24(4): p. 969–81. https://doi.org/10.1093/molbev/msm014 PMID: 17251179

38. Shepherd J.G., Wang L., and Reeves P.R., Comparison of O-antigen gene clusters of Escherichia coli (Shigella) sonnei and Plesiomonas shigelloides O17: sonnei gained its current plasmid-borne O-antigen genes from P. shigelloides in a recent event. Infect Immun, 2000. 68(10): p. 6056–61. PMID: 10992522

39. The H.C., et al., The genomic signatures of Shigella evolution, adaptation and geographical spread. Nat Rev Microbiol, 2016. 14(4): p. 235–50. https://doi.org/10.1038/nrmicro.2016.10 PMID: 26923111

40. Yang F., et al., Genome dynamics and diversity of Shigella species, the etiologic agents of bacillary dysentery. Nucleic Acids Res, 2005. 33(19): p. 6445–58. https://doi.org/10.1093/nar/gki954 PMID: 16275786

41. Maurelli A.T., et al., “Black holes” and bacterial pathogenicity: a large genomic deletion that enhances the virulence of Shigella spp. and enteroinvasive Escherichia coli. Proc Natl Acad Sci U S A, 1998. 95(7): p. 3943–8. PMID: 9520472
42. Daskaleros P.A. and Payne S.M., Characterization of Shigella flexneri sequences encoding Congo red binding (crb): conservation of multiple crb sequences and role of IS1 in loss of the Crb+ phenotype. Infect Immun, 1986. 54(2): p. 435–43. PMID: 3021630

43. Jin Q., et al., Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157. Nucleic Acids Res, 2002. 30(20): p. 439–41. PMID: 12384590

44. Szekeres S., et al., Chromosomal toxin-antitoxin loci can diminish large-scale genome reductions in the absence of selection. Mol Microbiol, 2007. 63(6): p. 1588–605. https://doi.org/10.1111/j.1365-2958.2007.05613.x PMID: 17367382

45. Christensen-Dalsgaard M. and Gerdes K., Two highBA loci in the Vibrio cholerae superintegron encode mRNA cleaving enzymes and can stabilize plasmids. Mol Microbiol, 2006. 62(2): p. 397–411. https://doi.org/10.1111/j.1365-2958.2006.05385.x PMID: 17020579

46. Colonna B., et al., H-NS regulation of virulence gene expression in enteroinvasive Escherichia coli harboring the virulence plasmid integrated into the host chromosome. J Bacteriol, 1995. 177(16): p. 4703–12. PMID: 7642498

47. Postow L., et al., Topological domain structure of the Escherichia coli chromosome. Genes Dev, 2004. 18(14): p. 1766–79. https://doi.org/10.1101/gad.1207504 PMID: 15256503

48. Valens M., et al., Macrodomain organization of the Escherichia coli chromosome. EMBO J, 2004. 23(21): p. 4330–41. https://doi.org/10.1038/sj.emboj.7600434 PMID: 15470498

49. Espeli O., Mercier R., and Boccard F., DNA dynamics vary according to macrodomain topography in the E. coli chromosome. Mol Microbiol, 2008. 68(6): p. 1418–27. https://doi.org/10.1111/j.1365-2958.2008.06239.x PMID: 18410497

50. Bryant J., Chewapreecha C., and Bentley S.D., Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. Future Microbiol, 2012. 7(11): p. 1283–96. https://doi.org/10.2217/fmb.12.108 PMID: 23075447

51. Vandenbosch J.L., et al., Sequence analysis of rsk, a portion of the 95-kilobase plasmid of Salmonella typhimurium associated with resistance to the bactericidal activity of serum. Infect Immun, 1989. 57(3): p. 850–7. PMID: 2652123

52. Faruque S.M., et al., Isolation of Shigella dysenteriae type 1 and S. flexneri strains from surface waters in Bangladesh: comparative molecular analysis of environmental Shigella isolates versus clinical strains. Appl Environ Microbiol, 2002. 68(8): p. 3908–13. https://doi.org/10.1128/AEM.68.8.3908-3913.2002 PMID: 12147489

53. Sansonetti P.J., Kopecko D.J., and Formal S.B., Involvement of a plasmid in the invasive ability of Shigella flexneri. Infect Immun, 1982. 35(3): p. 852–60. PMID: 6279518

54. Oaks E.V., Hale T.L., and Formal S.B., Serum immune response to Shigella protein antigens in rhesus monkeys and humans infected with Shigella spp. Infect Immun, 1986. 53(1): p. 57–63. PMID: 3721580

55. Cherepanov P.P. and Wackernagel W., Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. Gene, 1995. 158(1): p. 9–14. PMID: 7789817

56. Datsenko K.A. and Wanner B.L., One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A, 2000. 97(12): p. 6640–5. https://doi.org/10.1073/pnas.120163297 PMID: 10829079

57. Alikhan N.F., et al., BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics, 2011. 12: p. 402. https://doi.org/10.1186/1471-2164-12-402 PMID: 21824423

58. Rhoads A. and Au K.F., PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics, 2015. 13(5): p. 278–89. https://doi.org/10.1016/j.gpb.2015.08.002 PMID: 26542840

59. Herina S. D., M.J.D., Herrick J. B., A simple, rapid method for extracting large plasmid DNA from bacteria. Nature Proceedings, 2007.

60. Livak K.J. and Schmittgen T.D., Analysis of relative gene expression data using real-time quantitative PCR and the 2 (-Delta Delta C(T)) Method. Methods, 2001. 25(4): p. 402–8. https://doi.org/10.1006/meth.2001.1262 PMID: 11846609