Deletion of toxin-antitoxin systems in the evolution of *Shigella sonnei* as a host-adapted pathogen

Gareth McVicker[a] and Christoph M Tang[a]*

[a]Sir William Dunn School of Pathology
University of Oxford
South Parks Road
Oxford OX1 3RE
United Kingdom

*for correspondence
tel: 44-1865-275560
email: christoph.tang@path.ox.ac.uk

Short title: Host adaptation of *Shigella* virulence plasmid

Abstract

Pathogenic *Shigella* spp. are the leading cause of bacterial dysentery, with *Shigella flexneri* and *Shigella sonnei* accounting for around 90 % of cases worldwide. While *S. flexneri* causes most disease in low-income countries (following ingestion of contaminated food and/or water), *S. sonnei* predominates in wealthy countries and is mainly spread from person-to-person. Although both species contain a large virulence plasmid, pINV, that is essential for the organism to cause disease, little is known about its maintenance. Here, using a counterselectable marker within the virulence-encoding region of pINV, we show that the *S. flexneri* plasmid is less stable than that of *S. sonnei*, especially at environmental temperatures. GmvAT, a toxin-antitoxin system, is responsible for the difference in stability, and is present in pINV from *S. flexneri* but is absent in *S. sonnei* pINV; GmvT is an acetyltransferase toxin that inhibits protein translation. Loss of GmvAT and a second toxin-antitoxin system, CcdAB, from pINV reduces *S. sonnei* plasmid stability outside the host, reflecting the host-adapted lifestyle and person-to-person transmission of this species, and hence the striking differences in its epidemiology.

Many critical functions in bacteria, including antibiotic resistance and virulence, are encoded on plasmids[1] which establish long-standing associations with particular bacterial lineages, even though they can impose fitness costs on their bacterial host. Human pathogens such as *Shigella* spp., *Salmonella enterica* and *Yersinia* spp. harbour plasmids which are essential for pathogenesis[1-4], although little is known about the mechanisms of plasmid maintenance that lead to the retention of a functional plasmid, and how they are adapted to the lifecycle of these important human pathogens.

Shigella spp. are an important cause of bacillary dysentery that have emerged from *Escherichia coli* on several occasions following the acquisition of a single copy 210 kb plasmid, pINV[5]. The plasmid contains a pathogenicity island (PAI) of ~ 30 kb that encodes a Type III Secretion System (T3SS) essential for virulence, as well as effector molecules and regulatory proteins[6,7]. The *Shigella* T3SS acts as a molecular syringe that delivers effectors into host cells[8,9] and mediates entry of the bacterium[3,8]. During *in vitro* growth, *Shigella flexneri* can lose the PAI or the entire plasmid, which results in increased growth rate[10], highlighting the fitness cost of pINV in *S. flexneri* (pINV$_{Sf}$).

S. sonnei and *S. flexneri* cause the overwhelming majority of shigellosis worldwide[11,12]. Within less wealthy countries, *S. flexneri* infections are more common than *S. sonnei*[11,13], and are typically acquired from contaminated food or water[14]. In contrast, *S. sonnei* accounts for over 70 % of disease in wealthy countries[11]. In this setting, *S. sonnei* is mainly spread by direct person-to-person contact between individuals in close proximity, such as children in playgroups, men having sex with men (MSM), and those living in institutions[15-19]. As countries have improved their sanitation and water supply, *S. sonnei* has often displaced *S. flexneri* as the main cause of shigellosis[12,20-22].

Since pINV is non-conjugative and present in single copy[6], maintenance systems must exist to prevent its loss from *Shigella*. To date, these have only been studied in *S. flexneri* pINV$_{Sf}$ which encodes MvpAT, a VapBC-family toxin-antitoxin (TA) system[7,23,24]. TA systems consist of a toxin and a cognate antitoxin which prevents its activity. The toxin is relatively stable, whereas the antitoxin is specifically degraded by cellular proteases such as Lon[25,26]. If the genetic element encoding a TA system is lost during cell division, the pool of antitoxins is

degraded, leading the toxin to inhibit bacterial viability. MvpT is an RNAse that targets tRNA^{fMet} and stalls translation[24], promoting post-segregational killing of bacteria which have lost pINV$_{Sf}$[23,27].

Little is known about the maintenance of pINV in other species of *Shigella,* even though they have different modes of spread and capacity to cause disease while sharing the same virulence plasmid. Here we sought to define the mechanisms of pINV stability in *S. flexneri* and *S. sonnei* at temperatures found within and outside the human host. Here we examined the loss of the PAI and/or the entire plasmid, as both have a profound effect on host:pathogen interactions (through absence of the T3SS) and both are examples of plasmid instability that occur during growth[10]. To achieve this, we introduced a counterselectable marker into the PAI to report the loss of *Shigella* virulence. We show that discrete deletions in *S. sonnei* pINV impair its stability especially in environmental temperatures and reflect its recent emergence as a host-adapted pathogen.

Results

pINV stability differs in S. flexneri *and* S. sonnei

Plasmid stability (*i.e.* maintenance of the PAI and/or the entire plasmid) in *S. flexneri* and *S. sonnei* can be detected on solid media containing Congo red dye. When the T3SS is functional, *S. flexneri* colonies appear red, while avirulent bacteria, which lack either the PAI, its activator VirF (which is encoded outside the PAI[6]) or pINV entirely, form large white colonies[28] (**Fig. 1a**). Similarly, "Phase II" colonies of *S. sonnei* which do not express the T3SS are larger and paler than "Phase I" *S. sonnei* containing pINV$_{Ss}$; loss of the O antigen capsule

96 locus from *S. sonnei* pINV$_{Ss}$ also changes the colony morphology and leads to a 'rough'
97 phenotype[29,30] which can be detected as sectored colonies (**Fig. 1a**).
98
99 Initially we used colony morphology to compare pINV stability in *S. flexneri* M90T and *S.*
100 *sonnei* 53G. Bacteria were grown on solid media for ~ 25 generations at selected
101 temperatures, and individual colonies were then resuspended and plated to media
102 containing Congo red (**Fig. 1b**). At 37°C, pINV stability was over two orders of magnitude
103 higher in *S. flexneri* than *S. sonnei,* consistent with previous reports[29,31]. We also observed
104 no difference in the stability of pINV in *S. flexneri* at 37°C and 21°C, temperatures found in
105 the mammalian host and external environment, respectively (**Fig. 1b**; $p = 0.7819$). This
106 finding is in contrast to work that measured the emergence of PAI-negative bacteria during
107 liquid growth over 40 generations, although results may have been confounded by the
108 increased growth rate of avirulent bacteria at higher temperatures[27,32], which could
109 outgrow wild-type bacteria in these circumstances. In contrast to *S. flexneri*, pINV stability in
110 *S. sonnei* was markedly influenced by the ambient temperature, and was over an order of
111 magnitude more stable at 37°C than 21°C (**Fig. 1b**, $p = 0.0015$). Similar results were observed
112 with *S. sonnei* CS20, which belongs to a genetic lineage that has shown enhanced worldwide
113 spread in the last hundred years and is phylogenetically distinct from *S. sonnei* strain
114 53G[33,34] (**Supplementary Fig. 1**, $p < 0.0001$ for stability between 37°C and 21°C).
115
116 The use of colony morphology to assess pINV stability can be confounded by the rapid
117 replication of bacteria lacking the PAI (which can outgrow wild-type bacteria in mixed
118 populations[32]), and the difficulty detecting rare instances of plasmid instability. To
119 circumvent these issues, we introduced a *sacB-neo*^{R} marker into *mxiH*, which is located on

120 the PAI and encodes the needle subunit of the T3SS[35], in pINV$_{Sf}$ and pINV$_{Ss}$. The marker
121 allows selection of bacteria retaining the PAI (by kanamycin resistance), and those which
122 have lost the PAI (growth on media containing sucrose). In pilot experiments, *sacB* and *virB*
123 could not be amplified by PCR from 100/100 sucrose-resistant *S. flexneri* and *S. sonnei*
124 colonies, indicating a false positive rate of < 1 %. We also found that the *mxiH*::*sacB-neo*^{R}
125 strains have the same growth rate as their wild-type parental strains (**Supplementary Fig. 2,**
126 p = 0.7552 and 0.6368 for *S. flexneri* and *S. sonnei*, respectively). Furthermore, levels of pINV
127 stability in *S. flexneri* and *S. sonnei* measured using loss of *mxiH*::*sacB-neo*^{R} were comparable
128 to those detected on Congo red (**Fig. 1c-d**), except the limit of detection for PAI loss was
129 lower (~ 10^{-4} and 10^{-7} for assays using Congo red and *mxiH*::*sacB-neo*^{R}, respectively).
130 Consistent with Congo red assays, temperature does not influence pINV$_{Sf}$ stability in assays
131 with strains containing the *mxiH::sacB-neo*^{R} marker (**Fig. 1c**, p = 0.9465). Furthermore,
132 pINV$_{Ss}$ is significantly more stable at 37°C compared with lower temperatures (**Fig. 1b & 1d**,
133 p < 0.001) and colony sectoring in *S. sonnei* significantly increases as growth temperature
134 decreases in both *S. sonnei* 53G and its *mxiH::sacB-neo*^{R} derivative (**Supplementary Fig. 3**).
135 These data validate the use of *mxiH::sacB-neo*^{R} to accurately quantify PAI stability, and
136 confirm the effect of temperature on pINV stability in *S. flexneri* and *S. sonnei*.
137
138 S. flexneri *pINV*$_{Sf}$ *harbours three functional TA systems*
139 Previous work demonstrated that the TA system MvpAT encoded on pINV$_{Sf}$ contributes to
140 plasmid stability at 37°C[23,27]. Consistent with this, we found that deletion of *mvpAT* leads to
141 a marked reduction in pINV$_{Sf}$ stability at 37°C (**Fig. 1c,** 16-fold reduction, p < 0.0001). At
142 30°C, the effect of deleting *mvpAT* was less marked but was still significant (p = 0.0015),
143 while loss of *mvpAT* does not influence pINV$_{Sf}$ stability at 21°C (p = 0.9995).

144

145 To determine the mechanisms conferring pINV stability at environmental temperatures, we
146 analysed the sequence of pINV$_{Sf}$, and found four putative TA systems in addition to MvpAT.
147 *ccdAB* is closely related to a TA system on the *E. coli* F plasmid (90.9 % nucleotide identity)[36],
148 while ORFs *0062-3*, *0111-2* (gene numbers from pWR501[7]) and *yacAB* are previously
149 uncharacterised. *0062* and *0111* are both predicted to encode GCN5-related *N*-
150 acetyltransferases (GNATs), which use acetyl-CoA as a donor to acetylate target molecules,
151 while the adjacent genes, *0063* and *0112*, encode potential antitoxins with predicted DNA-
152 binding domains, which could mediate autoregulation as seen with other TA systems[37]. The
153 *yacA* product is a member of the RelB/DinJ-like antitoxin family, although *yacB* has a
154 frameshift mutation that leads to a premature stop codon. We also identified a gene (*0205*)
155 encoding a potential RelB-like antitoxin.
156
157 To determine whether these TA systems are functional, each gene was expressed in *S.*
158 *flexneri* BS176 (which lacks pINV[38]) using an arabinose-inducible promoter. As expected,
159 none of the predicted antitoxins (*mvpA*, *ccdA*, *yacA*, *0063*, *0112* and *0205*) affected bacterial
160 survival when expressed (**Supplementary Fig. 4**) while expression of *mvpT* dramatically
161 impaired bacterial viability (**Fig. 2a**). In addition, induction of *ccdB* or *0062* significantly
162 reduced bacterial survival at 21°C and 37°C (**Fig. 2b, 2c and Supplementary Fig. 5**, two-way
163 ANOVA $p < 0.0001$), while expression of *yacB* and *0111* did not affect viability
164 (**Supplementary Fig. 4**). We next introduced amino acid substitutions into MvpT (D7A[24]),
165 CcdB (W99S, G100E and I101K[39], generating CcdB^{SEK}) and 0062 (G103D) to impair their
166 known/predicted activities; the modification in 0062 is predicted to abolish acetyl-CoA
167 binding, and hence its ability to acetylate its target(s)[40]. In each instance, the substitutions

abolished the toxicity of the proteins (**Fig. 2**). Additionally, bacterial death was prevented by co-expression of *mvpA*, *ccdA* and *0063* with their corresponding toxins (**Fig. 2**). Taken together these results confirm that 0062/0063 is a functional TA system (like *mvpAT* and *ccdAB*) and was designated as GmvAT (for GNAT maintenance of virulence antitoxin/toxin) due to its predicted biochemical function.

GmvAT stabilises pINV$_{Sf}$ at environmental temperatures

To determine the contribution of *mvpAT*, *ccdAB* and *gmvAT* to pINV$_{Sf}$ stability, we constructed strains lacking these TA systems either individually or in combination in *S. flexneri*, and assessed plasmid stability at 21°C and 37°C. At 37°C, *mvpAT* is necessary and sufficient for pINV$_{Sf}$ stability **(Fig. 3a**), as shown previously[27]; deletion of *ccdAB* and *gmvAT* alone or together has no impact on pINV$_{Sf}$ stability at this temperature (each deletion *vs.* wild-type, $p > 0.9999$). In contrast, at 21°C, GmvAT was sufficient for plasmid stability (**Fig. 3f**, GmvAT alone *vs.* wild-type, $p > 0.9999$). However, MvpAT and CcdAB in combination can compensate for the absence of *gmvAT* at 21°C ($p = 0.5596$), demonstrating redundancy between the TA systems at this temperature. CcdAB only affects pINV$_{Sf}$ stability at 37°C in the absence of MvpAT ($p < 0.02$) or at 21°C in the absence of both other TA systems ($p = 0.0049$). qRT-PCR analysis of *gmvA* and *mvpA* indicated that there was no significant difference between expression of these two TA systems at 21°C relative to 37°C, so the effect of temperature is likely to operate at a post-transcriptional level (**Supplementary Fig. 6**, $p = 0.4783$). Overall, these data demonstrate that MvpAT is responsible for the stability of *S. flexneri* pINV$_{Sf}$ at 37°C, while GmvAT is the only TA system that is sufficient for pINV$_{Sf}$ stability at environmental temperatures.

GmvT is a novel acetyl transferase that inhibits protein translation

In *E. coli*, GNATs modify a variety of targets including RNA polymerase and ribosomal proteins[41] using acetyl-CoA as a donor, while many TA system toxins affect protein synthesis[24,42,43]. To examine the mechanism of action of GmvT and whether it affects transcription and/or translation, we generated linear DNA templates of His-tagged GmvT and GmvT^{G103} as templates for *in vitro* transcription/translation assays (**Fig. 4a**). The production of GmvT was not detected either by Coomassie staining or Western blot analysis in reactions containing the template for the wild-type protein and acetyl-CoA, but was detected in reactions lacking the acetyl donor. In contrast, GmvT^{G103D} was generated in reactions with or without acetyl-CoA (**Fig. 4a**). To define the step inhibited by GmvT, we measured levels of *gmvT*-specific mRNA produced in the reactions (**Fig. 4a**); *gmvT* mRNA was detected in all reactions regardless of the nature of the template or the presence of acetyl-CoA. Additionally we purified mRNA generated from *in vitro* transcription reactions using *gmvT* as the DNA template, then added a fixed amount of mRNA (2.5 µg) to translation reactions with or without acetyl-CoA (**Fig. 4b**). GmvT was produced only in reactions lacking acetyl-CoA. Finally, we examined the effect of GmvT on a second DNA template (encoding sfGFP) in *in vitro* transcription/translation reactions. The band corresponding to GFP was markedly reduced only in reactions that contained both acetyl-CoA and functional GmvT (**Fig. 4c**). Taken together, these results demonstrate that GmvT is an acetyl-CoA dependent GNAT that blocks protein translation.

Distinct mechanisms of pINV maintenance in S. sonnei *and* S. flexneri

To define the mechanisms underlying the distinct behaviours of pINV in *S. flexneri* and *S. sonnei*, we examined the plasmid sequences. Of note, pINV$_{Ss}$ from *S. sonnei* 53G entirely

lacks *gmvAT* (as part of a 4.9 kb deletion) and *ccdB*, although 23 nucleotides of *ccdA* remain (**Fig. 5**). To determine whether the absence of these TA systems accounts for the difference in pINV stability between *S. flexneri* and *S sonnei*, we introduced *ccdAB* and *gmvAT* individually or together into *S. sonnei* pINV$_{Ss}$ at sites corresponding to their positions in *S. flexneri* (generating *S. sonnei ccd*^{WT}*gmv*^{WT}, **Fig. 5b**); strains with non-functional toxins in the same loci were also constructed (generating *S. sonnei ccd*^{SEK}*gmv*^{G103D}). As expected, the non-functional TA systems had no impact on pINV$_{Ss}$ stability at all temperatures tested (**Supplementary Fig. 7**). Of note, we found that insertion of functional *gmvAT* into the *S. sonnei* pINV$_{Ss}$ is sufficient to stabilise pINV$_{Ss}$ at 37°C to pINV$_{Sf}$ levels (**Fig. 5c**, $p < 0.0001$); *ccdAB* does not increase stability at this temperature. At 21°C, *gmvAT* has an even more dramatic effect, stabilising *S. sonnei* pINV$_{Ss}$ by over 200-fold ($p < 0.0001$), while *ccdAB* has a minor contribution only in the presence of *gmvAT* ($p = 0.0176$). Furthermore, introduction of both functional TA systems stabilised pINV$_{Ss}$ to levels that are not statistically different from pINV$_{Sf}$ at 37°C and 21°C (**Fig. 5c** and **Supplementary Fig. 8**, $p = 0.4677$ and 0.3958, respectively). PAI retention could result from increased plasmid stability in a sub-population of bacteria. Therefore we measured PAI-loss in cells that had retained the PAI in an initial assay; there was no significant difference in the rate of PAI-loss between the initial and subsequent assays (**Supplementary Fig. 9;** $p > 0.08$). Furthermore, deletion of TA systems from *S. flexneri* or their introduction into *S. sonnei* did not significantly affect the fitness of bacteria (**Supplementary Fig. 10**). These results demonstrate that the absence of *gmvAT* (and to a lesser extent *ccdAB*) in *S. sonnei* 53G accounts for the reduced stability of pINV$_{Ss}$ at environmental temperatures, and the difference in plasmid stability between *S. flexneri* and *S. sonnei*.

The role of the TA systems in maintenance of virulence was confirmed by introducing GmvAT and CcdAB into a wild-type *S. sonnei* strain, then using Congo red, rather than sucrose selection, to detect the emergence of avirulent (PAI-negative) bacteria during growth at 21°C. Over 48 generations, the number of virulent bacteria in *S. sonnei* 53G cultures fell to approximately 40 %, whereas it remained above 98 % in the TA^+ mutant (**Fig. 6a**, $p < 0.0001$). We also analysed the ability of bacteria to form plaques in epithelial cell monolayers which measures their invasive capacity, intracellular replication, and cell-to-cell spread[44]. *S. sonnei* 53G had a significantly reduced ability to form plaques following passage at 21°C compared (**Fig. 6b**, $p = 0.0009$), whereas the number of plaques formed by *S. sonnei* $ccd^{WT}gmv^{WT}$ strain did not change over time ($p = 0.9442$), demonstrating that the loss of virulence during growth at 21°C was impaired by insertion of the TA systems.

gmvAT *and* ccdAB *were lost from* $pINV_{Ss}$ *soon after the emergence of* S. sonnei

Given their role in plasmid maintenance, we examined available *Shigella* genome sequences for the presence of *ccdAB*, *gmvAT* and *mvpAT*. In *Shigella boydii* (strains Sb227 and CDC 3083-94) and *Shigella dysenteriae* (Sd197 and 1617), all three TA systems are present. Similarly, in a collection of *S. flexneri* isolates[45], we found that all strains with ≥ 45% sequence coverage of $pINV_{Sf}$ (over 200 isolates) possess *ccdAB*, *gmvAT* and *mvpAT*. In contrast, *ccdAB* or *gmvAT* were not detected by BLAST in the genome of any of 132 *S. sonnei* strains isolated across four continents worldwide between 1948 and 2008[33]. However only 46 of 132 isolates have ≥ 90% sequence coverage of pINV. Of these 46 strains, 41 contain the identical *gmvAT* deletion as in *S. sonnei* 53G, and 38 isolates have the same 23 bp *ccdA* fragment. The remaining strains either have larger deletions or insufficient sequence for analysis.

The *ccd* deletion is seen across all three phylogenetic lineages of *S. sonnei*, including the oldest lineage, lineage I. Lineage II and III strains harbour the same *gmv* deletion although there is insufficient available sequence to define the nature of the loss in lineage I strains. Therefore, there are specific, discrete deletions of *ccdAB* and *gmvAT* present across the entire phylogenetic tree of *S. sonnei*, indicating that these TA systems were each lost from pINV on a single occasion soon after *S. sonnei* diverged from *E. coli*.

Discussion

Acquisition of pINV was a critical step in the evolution of *Shigella* spp., enabling bacteria to invade epithelial cells using the plasmid-encoded T3SS[3,46]. *Shigella* spp. have subsequently adapted to their pathogenic lifestyle by a series of chromosomal deletions, resulting in the loss of catabolic functions and motility[5,46]. *S. sonnei* differs from other *Shigella* spp. as it is largely responsible for disease in industrialised countries with the bacterium usually acquired direct from another infected individual rather than from an environmental source[15-18]. We show that two unique genetic events have led to the loss of TA systems from pINV in *S. sonnei* and its reduced capacity to maintain the plasmid and/or the PAI while in the external environment. These discrete deletions reflect the ongoing transition of *S. sonnei* into an obligate pathogen that is less dependent on survival outside a mammalian host than other species of *Shigella*. The presence of genes encoding an O-antigen capsule[30,47] on *S. sonnei* pINV is a further adaptive change in the plasmid of a bacterium that maintains an intimate relationship with its host, as the capsule promotes resistance to complement-mediated killing and lowers the inflammatory response[47].

GNATs are a highly abundant family of proteins found in all domains of life[41] and GmvT is the first example of a GNAT that stabilises a mobile genetic element through inhibition of protein translation. GmvT might have similar activity to TacT, which is a chromosomally-encoded GNAT that mediates persister formation in *Salmonella* by acetylation of aminoacyl-tRNA[48]. We show that GmvAT is sufficient for retention of the *S. flexneri* invasion plasmid and/or the PAI at temperatures that would be encountered outside of the human host. Critical virulence functions conferred by pINV are under precise regulatory control, and triggered by environmental cues such a temperature[49] and oxygen[50]. Our work demonstrates that the mechanisms of maintenance of pINV are also influenced by specific signals, and are tailored to the distinct lifestyles of different *Shigella* spp.

Methods

Strains and growth media

Bacterial strains and plasmids used in this study are shown in **Supplementary Table 1**. Unless otherwise stated, *E. coli* and *Shigella* were grown in lysogeny broth (LB; Invitrogen, Waltham, MA), or on solid media containing 1.5 % (w/v) agar (Oxoid, Basingstoke, UK). Antibiotics were used at the following concentrations: carbenicillin, 50 μg/ml; chloramphenicol, 20 μg/ml; kanamycin, 50 μg/ml. Congo red dye (0.01 % w/v, final concentration) was added to tryptic soy broth (Fluka, Buchs, Switzerland) plus agar. For selection on sucrose, 1 % (w/v) tryptone (Fluka), 0.5 % (w/v) yeast extract (Fluka) and agar as above were autoclaved in water, and sucrose was added to a final concentration of 10 % (w/v).

Genetic manipulation of S. flexneri *and* S. sonnei

DNA constructs were assembled into plasmids (**Supplementary Table 1**) in *E. coli* DH5α using NEBuilder HiFi DNA Assembly master mix (New England Biolabs (NEB), Ipswich, MA); oligonucleotide primers are given in **Supplementary Table 2**. λ Red recombination[51,52] was employed to construct mutations in *Shigella*. Approximately 1 kb of flanking sequence was used to integrate linear fragments into *Shigella*. For TA restoration in *S. sonnei*, *ccdAB* was positioned at nucleotide position 150,672 of the 53G pINV$_{Ss}$, while *gmvAT* was inserted at nucleotide position 40,777 (GenBank HE616529.1). The *sacB-neo*^{R} cassette was amplified from pIB279[53].

323 *Transductions*

324 P1*vir* bacteriophage transductions were performed in *S. flexneri* M90T by adapting a
325 previously described protocol[54]. Donor cells were grown overnight in LB at 37°C with
326 shaking at 180 rpm, then diluted 1/1000 into 10 ml fresh LB plus 0.2 % (w/v) glucose, 5 mM
327 $CaCl_2$ and grown for 1 hour. P1*vir* lysate (200 µl) was added to the bacteria and incubated
328 for 1-4 hours. The clarified lysate was treated with 200 µl chloroform, then the cell debris
329 was pelleted at 4,500 rpm for 10 minutes. Lastly the supernatant was passed through a 0.45
330 µm filter and stored at 4°C. Recipient strains were grown overnight in LB, pelleted then
331 resuspended in 0.5 volumes phage buffer (10 mM $CaCl_2$, 20 mM $MgSO_4$). Phage lysate was
332 added to 0.1 ml cell suspension at a range of volume ratios (1:100 to 1:1). Mixtures were
333 incubated without shaking at 37°C for not longer than 30 minutes, after which 1 ml LB
334 containing 0.1 M sodium citrate was added. Cells were then incubated at 37°C, 180 rpm for
335 1 hour prior to being pelleted and were then re-suspended in fresh LB + 0.1 M citrate.
336 Aliquots were spread onto LB agar plates containing appropriate antibiotics. Mutations were
337 verified by PCR and sequencing.

338

339 *Toxicity assays*

340 To assess toxicity of TA-related proteins encoded by pINV$_{Sf}$, strains containing pBAD33[55]
341 derivatives encoding inducible toxins and pGM101 derivatives encoding antitoxins were
342 grown at selected temperatures to OD = 0.1 in LB + 0.2 % (w/v) glucose, then pelleted and
343 resuspended in LB + 1 % (w/v) arabinose. A serial dilution of each culture in PBS was plated
344 on LB agar + 0.2 % (w/v) glucose at various timepoints post induction and bacteria were
345 grown overnight at 37°C to obtain viable counts. Appropriate antibiotics were included in all
346 media during the assay to maintain the integrity of expression vectors.

Virulence plasmid stability assays

pINV stability was measured in two ways: i) analysis of colonies on Congo red and ii) selection on kanamycin or sucrose in bacteria harbouring *mxiH*::*sacB-neo*^{R}. Cells from frozen stocks were plated on LB agar at selected temperatures and grown for ~ 25 generations. Individual colonies were suspended in PBS, serially diluted, and plated onto either Congo red agar, or both to LB plates containing kanamycin and to sucrose plates (to quantify bacteria without the *sacB-neo*^{R} cassette). Colonies were discarded from analysis if they were found to contain ~ 100 % sucrose resistance, indicating a PAI-negative founder. To assess plasmid instability during longer growth periods in liquid media, cultures were repeatedly subcultured into fresh LB after reaching stationary phase (*e.g.* every 24 hours at 21°C).

Analysis of gene expression and protein synthesis

In vitro transcription was carried out using the MEGAscript T7 kit (Ambion, Basingstoke, UK) and combined transcription/translation reactions were performed using the PURExpress kit (NEB), according to manufacturers' instructions. RNA was extracted from *in vitro* transcription reactions using the MEGAclear kit (Ambion) or from exponentially-growing bacterial cultures using Trizol (Ambion) and ethanol precipitation. Western blots were carried out using antibodies ab18184 (Abcam, Cambridge, UK) and JL-8 (Clontech, Mountain View, CA) in phosphate buffered saline containing 5 % (w/v) milk powder and 0.1 % (v/v) Tween-20. Quantitative reverse-transcription polymerase chain reaction (qRT-PCR) was conducted according to standard protocols using Power SYBR Green master mix in a StepOnePlus thermocycler (Applied Biosystems, Foster City, CA).

371 *Plaque assay*

372 HeLa cell monolayers were infected with *S. sonnei* strains by adapting an established
373 protocol[44]. HeLa cells were originally obtained from the European Collection of
374 Authenticated Cell Cultures and were not further authenticated or tested for *Mycoplasma*
375 contamination in-house during routine assays. HeLa cells were grown to confluency in high-
376 glucose Dulbecco's Modified Eagle Medium (DMEM; Sigma-Aldrich, Dorset, UK) containing
377 10 % heat-inactivated fetal bovine serum (FBS; Gibco, Paisley, UK) in 6-well plates at 37°C, 5
378 % CO_2. Passaged and ancestral *S. sonnei* strains were subcultured and grown at 37°C with
379 aeration to mid-exponential phase in brain heart infusion medium (BHI; Oxoid). Bacteria
380 were added at an appropriate dilution in 200 µl BHI to each well to give a final multiplicity of
381 infection of 10:1. Plates were then incubated for 90 minutes at 37°C in 5 % CO_2 with gentle
382 rocking every 30 minutes. During this time, 1 % (w/v) low-melt agarose (AGTC Bioproducts,
383 Fleet, UK) was mixed at a 1:1 ratio with warm DMEM + 10 % FBS, then gentamycin was
384 added at a final concentration of 20 µg/ml to prevent extracellular bacterial replication. 2 ml
385 of this mixture was overlaid onto to each well and allowed to set for approximately 45
386 minutes at room temperature. Plates were incubated at 37°C in 5 % CO_2 for 72 hours, after
387 which plaques were counted under magnification.

388

389 *Statistical and computational methods*

390 Log-transformed (normally-distributed) data were analysed using either an unpaired t-test,
391 linear regression or one-way or two-way ANOVA with Tukey's multiple comparisons tests as
392 appropriate (see figure legends). Plasmid alignment diagrams were created using BRIG
393 v0.95[56] and BLAST (blastn) v2.2.29.

394

Data availability

The data that support the findings of this study are available from the corresponding author upon request.

References

1 Johnson, T. J. & Nolan, L. K. Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiology and molecular biology reviews : MMBR* **73**, 750-774, doi:10.1128/mmbr.00015-09 (2009).

2 Rychlik, I., Gregorova, D. & Hradecka, H. Distribution and function of plasmids in *Salmonella enterica*. *Veterinary microbiology* **112**, 1-10, doi:10.1016/j.vetmic.2005.10.030 (2006).

3 Sansonetti, P. J. Genetic and molecular basis of epithelial cell invasion by *Shigella* species. *Reviews of infectious diseases* **13 Suppl 4**, S285-292 (1991).

4 Cornelis, G. R. *et al.* The virulence plasmid of *Yersinia*, an antihost genome. *Microbiology and molecular biology reviews : MMBR* **62**, 1315-1352 (1998).

5 Pupo, G. M., Lan, R. & Reeves, P. R. Multiple independent origins of Shigella clones of Escherichia coli and convergent evolution of many of their characteristics. *Proc Natl Acad Sci U S A* **97**, 10567-10572 (2000).

6 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* **38**, 760-771. (2000).

7 Venkatesan, M. M. *et al.* Complete DNA sequence and analysis of the large virulence plasmid of *Shigella flexneri*. *Infection and immunity* **69**, 3271-3285, doi:10.1128/iai.69.5.3271-3285.2001 (2001).

8 Nhieu, G. T. & Sansonetti, P. J. Mechanism of Shigella entry into epithelial cells. *Curr Opin Microbiol* **2**, 51-55 (1999).

9 Parsot, C. *Shigella* spp. and enteroinvasive *Escherichia coli* pathogenicity factors. *FEMS microbiology letters* **252**, 11-18, doi:10.1016/j.femsle.2005.08.046 (2005).

10 Sasakawa, C. *et al.* Molecular alteration of the 140-megadalton plasmid associated with loss of virulence and Congo red binding activity in *Shigella flexneri*. *Infection and immunity* **51**, 470-475 (1986).

11 Kotloff, K. L. *et al.* Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bulletin of the World Health Organization* **77**, 651-666 (1999).

12 Thompson, C. N., Duy, P. T. & Baker, S. The Rising Dominance of *Shigella sonnei*: An Intercontinental Shift in the Etiology of Bacillary Dysentery. *PLoS neglected tropical diseases* **9**, e0003708, doi:10.1371/journal.pntd.0003708 (2015).

13 Livio, S. *et al. Shigella* isolates from the global enteric multicenter study inform vaccine development. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **59**, 933-941, doi:10.1093/cid/ciu468 (2014).

14 DuPont, H. L., Levine, M. M., Hornick, R. B. & Formal, S. B. Inoculum size in shigellosis and implications for expected mode of transmission. *J Infect Dis* **159**, 1126-1128 (1989).

15 Baker, K. S. *et al.* Intercontinental dissemination of azithromycin-resistant shigellosis through sexual transmission: a cross-sectional study. *The Lancet. Infectious diseases* **15**, 913-921, doi:10.1016/s1473-3099(15)00002-x (2015).

16 Genobile, D. *et al.* An outbreak of shigellosis in a child care centre. *Communicable diseases intelligence quarterly report* **28**, 225-229 (2004).
17 Simms, I. *et al.* Intensified shigellosis epidemic associated with sexual transmission in men who have sex with men--*Shigella flexneri* and *S. sonnei* in England, 2004 to end of February 2015. *Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin* **20** (2015).
18 Mahoney, F. J., Farley, T. A., Burbank, D. F., Leslie, N. H. & McFarland, L. M. Evaluation of an intervention program for the control of an outbreak of shigellosis among institutionalized persons. *The Journal of infectious diseases* **168**, 1177-1180 (1993).
19 Cohen, D. *et al.* Recent trends in the epidemiology of shigellosis in Israel. *Epidemiol Infect* **142**, 2583-2594, doi:10.1017/s0950268814000260 (2014).
20 Kotloff, K. L. *et al.* Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet (London, England)* **382**, 209-222, doi:10.1016/s0140-6736(13)60844-2 (2013).
21 Vinh, H. *et al.* A changing picture of shigellosis in southern Vietnam: shifting species dominance, antimicrobial susceptibility and clinical presentation. *BMC infectious diseases* **9**, 204, doi:10.1186/1471-2334-9-204 (2009).
22 Lima, I. F., Havt, A. & Lima, A. A. Update on molecular epidemiology of *Shigella* infection. *Current opinion in gastroenterology* **31**, 30-37, doi:10.1097/mog.0000000000000136 (2015).
23 Sayeed, S., Reaves, L., Radnedge, L. & Austin, S. The stability region of the large virulence plasmid of *Shigella flexneri* encodes an efficient postsegregational killing system. *Journal of bacteriology* **182**, 2416-2421 (2000).
24 Winther, K. S. & Gerdes, K. Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7403-7407, doi:10.1073/pnas.1019587108 (2011).
25 Winther, K. S. & Gerdes, K. Regulation of Enteric *vapBC* Transcription: Induction by VapC Toxin Dimer-Breaking. *Nucleic Acids Research* **40**, 4347-4357, doi:10.1093/nar/gks029 (2012).
26 Van Melderen, L. *et al.* ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions. *The Journal of biological chemistry* **271**, 27730-27738 (1996).
27 Sayeed, S., Brendler, T., Davis, M., Reaves, L. & Austin, S. Surprising dependence on postsegregational killing of host cells for maintenance of the large virulence plasmid of *Shigella flexneri*. *Journal of bacteriology* **187**, 2768-2773, doi:10.1128/jb.187.8.2768-2773.2005 (2005).
28 Payne, S. M. & Finkelstein, R. A. Detection and differentiation of iron-responsive avirulent mutants on Congo red agar. *Infection and immunity* **18**, 94-98 (1977).
29 Kopecko, D. J., Washington, O. & Formal, S. B. Genetic and physical evidence for plasmid control of *Shigella sonnei* form I cell surface antigen. *Infection and immunity* **29**, 207-214 (1980).
30 Shepherd, J. G., Wang, L. & 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. *Infection and immunity* **68**, 6056-6061 (2000).
31 Sansonetti, P. J., Kopecko, D. J. & Formal, S. B. Involvement of a plasmid in the invasive ability of Shigella flexneri. *Infect Immun* **35**, 852-860. (1982).
32 Schuch, R. & Maurelli, A. T. Virulence plasmid instability in *Shigella flexneri* 2a is induced by virulence gene expression. *Infection and immunity* **65**, 3686-3692 (1997).

33 Holt, K. E. *et al. Shigella sonnei* genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. *Nature genetics* **44**, 1056-1059, doi:10.1038/ng.2369 (2012).

34 Angelini, M., Stehling, E. G., Moretti, M. L. & da Silveira, W. D. Molecular epidemiology of *Shigella* spp strains isolated in two different metropolitam areas of southeast Brazil. *Braz J Microbiol* **40**, 685-692, doi:10.1590/s1517-838220090003000034 (2009).

35 Blocker, A. *et al.* Structure and composition of the Shigella flexneri "needle complex", a part of its type III secreton. *Mol Microbiol* **39**, 652-663. (2001).

36 Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *Journal of Molecular Biology* **226**, 735-745 (1992).

37 Chan, W. T., Espinosa, M. & Yeo, C. C. Keeping the Wolves at Bay: Antitoxins of Prokaryotic Type II Toxin-Antitoxin Systems. *Frontiers in molecular biosciences* **3**, 9, doi:10.3389/fmolb.2016.00009 (2016).

38 Zychlinsky, A., Prevost, M. C. & Sansonetti, P. J. *Shigella flexneri* induces apoptosis in infected macrophages. *Nature* **358**, 167-169, doi:10.1038/358167a0 (1992).

39 Bahassi, E. M., Salmon, M. A., Van Melderen, L., Bernard, P. & Couturier, M. F plasmid CcdB killer protein: *ccdB* gene mutants coding for non-cytotoxic proteins which retain their regulatory functions. *Molecular microbiology* **15**, 1031-1037 (1995).

40 Lu, L., Berkey, K. A. & Casero, R. A., Jr. RGFGIGS is an amino acid sequence required for acetyl coenzyme A binding and activity of human spermidine/spermine N1acetyltransferase. *The Journal of biological chemistry* **271**, 18920-18924 (1996).

41 Vetting, M. W. *et al.* Structure and functions of the GNAT superfamily of acetyltransferases. *Archives of biochemistry and biophysics* **433**, 212-226, doi:10.1016/j.abb.2004.09.003 (2005).

42 Christensen, S. K. *et al.* Overproduction of the Lon protease triggers inhibition of translation in *Escherichia coli*: involvement of the *yefM-yoeB* toxin-antitoxin system. *Molecular microbiology* **51**, 1705-1717 (2004).

43 Christensen, S. K., Mikkelsen, M., Pedersen, K. & Gerdes, K. RelE, a global inhibitor of translation, is activated during nutritional stress. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 14328-14333, doi:10.1073/pnas.251327898 (2001).

44 Oaks, E. V., Wingfield, M. E. & Formal, S. B. Plaque formation by virulent *Shigella flexneri*. *Infection and immunity* **48**, 124-129 (1985).

45 Connor, T. R. *et al.* Species-wide whole genome sequencing reveals historical global spread and recent local persistence in *Shigella flexneri*. *eLife* **4**, e07335, doi:10.7554/eLife.07335 (2015).

46 Lan, R. & Reeves, P. R. *Escherichia coli* in disguise: molecular origins of *Shigella*. *Microbes and infection / Institut Pasteur* **4**, 1125-1132 (2002).

47 Caboni, M. *et al.* An O antigen capsule modulates bacterial pathogenesis in *Shigella sonnei*. *PLoS pathogens* **11**, e1004749, doi:10.1371/journal.ppat.1004749 (2015).

48 Cheverton, A. M. *et al.* A *Salmonella* Toxin Promotes Persister Formation through Acetylation of tRNA. *Molecular cell* **63**, 86-96, doi:10.1016/j.molcel.2016.05.002 (2016).

49 Falconi, M., Colonna, B., Prosseda, G., Micheli, G. & Gualerzi, C. O. 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. *The EMBO journal* **17**, 7033-7043, doi:10.1093/emboj/17.23.7033 (1998).

50 Marteyn, B. *et al.* Modulation of *Shigella* virulence in response to available oxygen in vivo. *Nature* **465**, 355-358, doi:10.1038/nature08970 (2010).

51 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**, 6640-6645. (2000).

52 Cherepanov, P. P. & Wackernagel, W. Gene disruption in Escherichia coli: TcR and KmR cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9-14 (1995).
53 Blomfield, I. C., Vaughn, V., Rest, R. F. & Eisenstein, B. I. Allelic exchange in Escherichia coli using the Bacillus subtilis sacB gene and a temperature-sensitive pSC101 replicon. *Mol Microbiol* **5**, 1447-1457 (1991).
54 Ramadoss, N. S., Zhou, X. & Keiler, K. C. tmRNA is essential in *Shigella flexneri*. *PloS one* **8**, e57537, doi:10.1371/journal.pone.0057537 (2013).
55 Guzman, L. M., Belin, D., Carson, M. J. & Beckwith, J. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of bacteriology* **177**, 4121-4130 (1995).
56 Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC genomics* **12**, 402, doi:10.1186/1471-2164-12-402 (2011).

Acknowledgements

We thank Ian Blomfield for the gift of P1*vir* bacteriophage and pIB279 (*sacB-neo*^{R}), Wanderley Dias da Silveira for *S. sonnei* strain CS20, and Kate Baker and Nicholas Thomson for assistance with genomic analyses. We also thank the Kirkhouse Trust (Scottish Charity No. 030508) for donation of laboratory equipment. This work was funded by Stopenterics EU grant no. 261472, and a Wellcome Trust Senior Investigator award to C.M.T.

Author Contributions

G.M. performed experiments and analysed data; G.M. and C.M.T. designed experiments, interpreted data and wrote the manuscript; C.M.T. secured funding.

The authors declare no competing financial interests.

Figure Legends

Figure 1. Loss of virulence in *S. sonnei* is temperature-dependent and higher than in *S. flexneri*. a: *S. flexneri* M90T and *S. sonnei* 53G grown on Congo red agar showing virulent bacteria (black arrows), avirulent bacteria (white arrows) and a sectored colony (yellow arrow). Images are representative of at least three independent experiments. **b**: Proportion of avirulent *S. flexneri* and *S. sonnei* after ~ 25 generations at the indicated temperatures (n = 3 cultures in independent experiments): solid line, mean; dotted line, limit of detection (l.o.d.). Pathogenicity island (PAI) loss in *S. flexneri mxiH::sacB-neo*^{R} and its Δ*mvpAT* derivative (**c**), and in *S. sonnei mxiH::sacB-neo*^{R} (**d**), detected by plating on sucrose and kanamycin; n = 9 colonies from three independent experiments, except *ΔmvpAT* at 21°C, n = 7 as two PAI-negative colonies were discarded. **** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$, * $p < 0.05$; n.s. not significant, one- or two-way ANOVA. Individual bacterial colonies were grown on nonselective LB agar plates for the stated number of generations, then resuspended in PBS and plated onto the appropriate media for quantification of either red/white (**b**) or kanamycin-/sucrose-resistant colonies (**c**, **d**).

Figure 2. *S. flexneri* pINV has three functional TA systems. Viability of *S. flexneri* BS176 following expression of MvpT (**a**), CcdB (**b**) or GmvT (**c**), or modified toxins, ± cognate antitoxins at 37°C; data shows mean ± s.e.m. (n = 3 cultures in independent experiments). pControl: empty vector. **** $p < 0.0001$ by two-way ANOVA.

Figure 3. *S. flexneri* pINV TA systems have temperature-dependent effects on plasmid maintenance. pINV$_{Sf}$ stability in mutants lacking TA systems at 37°C (**a, b, c**) and 21°C (**d, e, f**) assessed by loss of *mxiH::sacB-neo*^{R}. Colonies were grown for ~ 25 generations as

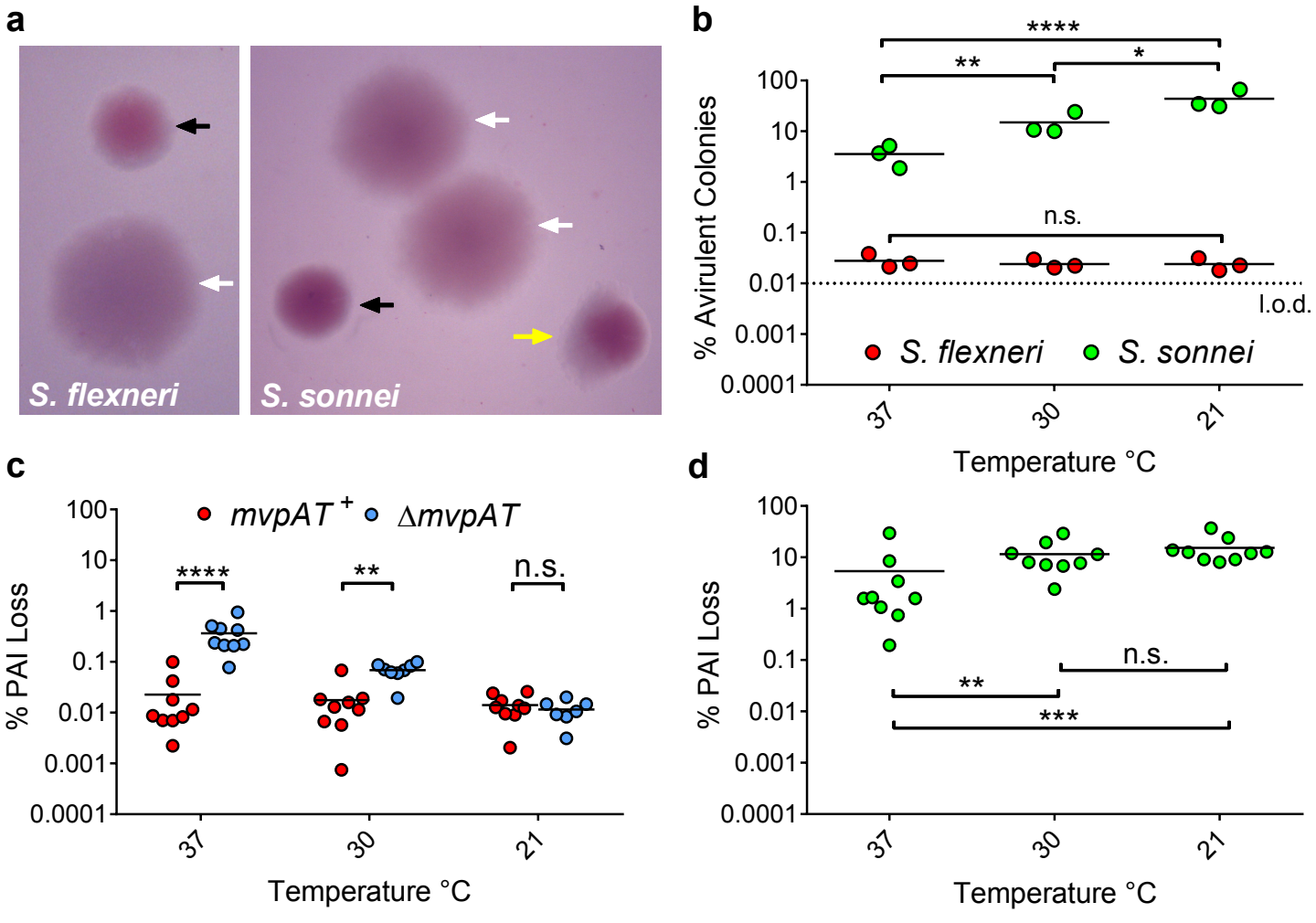
described previously. **** $p < 0.0001$; *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$; n.s. not significant. Statistical comparisons by one-way ANOVA with Tukey's multiple comparisons test (n = 9 colonies from three independent experiments, except *ΔmvpAT* at 21°C, n = 7 as two PAI-negative colonies were discarded).

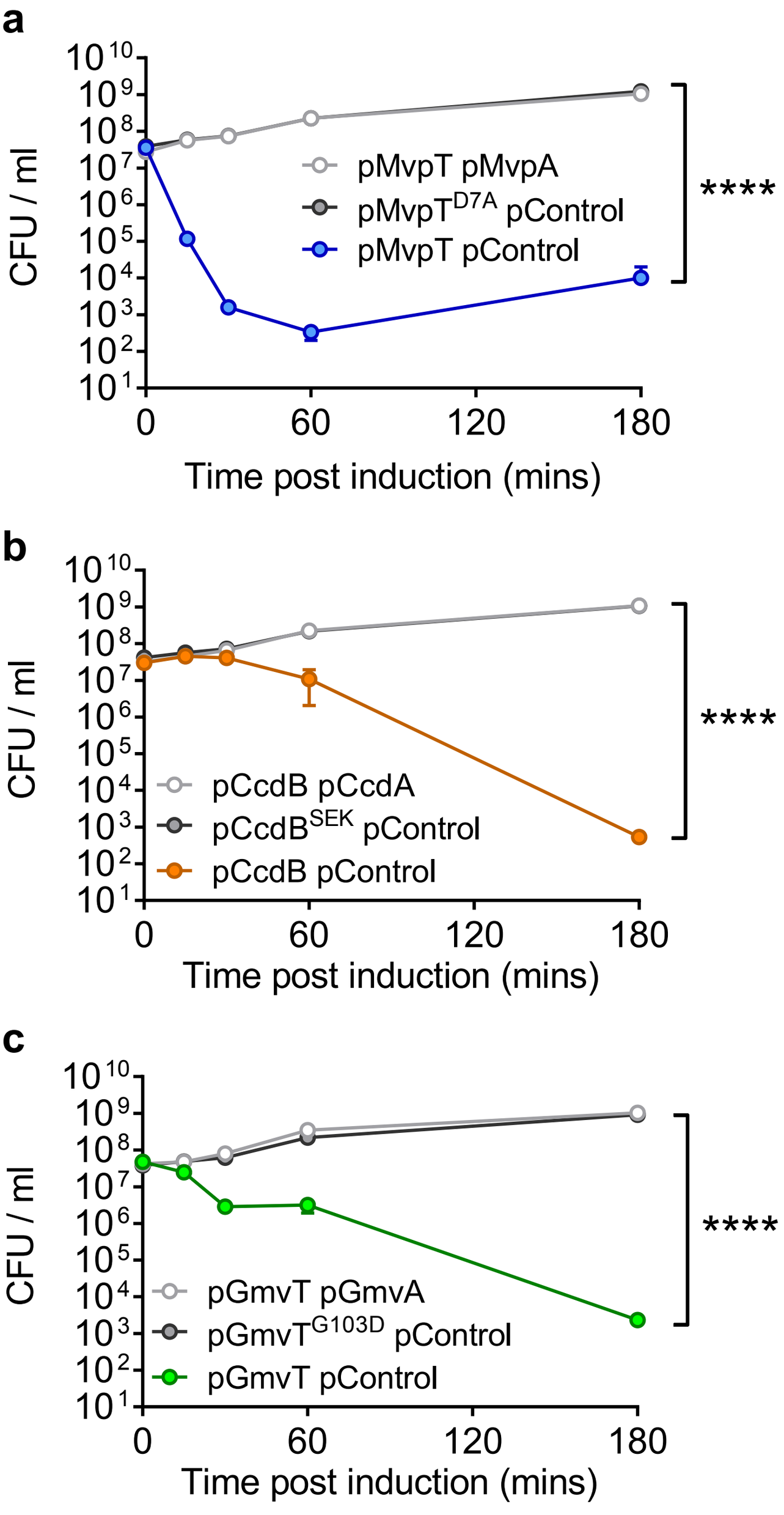
Figure 4. GmvT blocks translation in an acetyl-CoA dependent manner. a: Products of *in vitro* transcription/translation reactions using DNA templates encoding GmvT or $GmvT^{G103D}$, or with no template, with or without 100 µM acetyl-CoA (ac-CoA), analysed by SDS-PAGE and Coomassie staining: GmvT arrowed (upper panel); Western blot analysis with α-His pAbs to detect GmvT (middle panel); qRT-PCR analysis of mRNA generated in the reactions above, showing mean ± s.e.m. (n = 2 independent transcription assays, lower panel). **b**: Products of *in vitro* translation reactions following addition of 2.5 µg of purified *gmvT* mRNA, analysed by SDS-PAGE and Coomassie staining: GmvT arrowed (upper panel); Western blot analysis with α-His pAbs (lower panel). **c**: Products of *in vitro* transcription/translation reactions using DNA templates encoding sfGFP and either GmvT, $GmvT^{G103D}$ or no second protein, with or without 100 µM acetyl-CoA, analysed by Western blot analysis with α-GFP mAbs. Gels and blots are representative of at least 2 independent experiments.

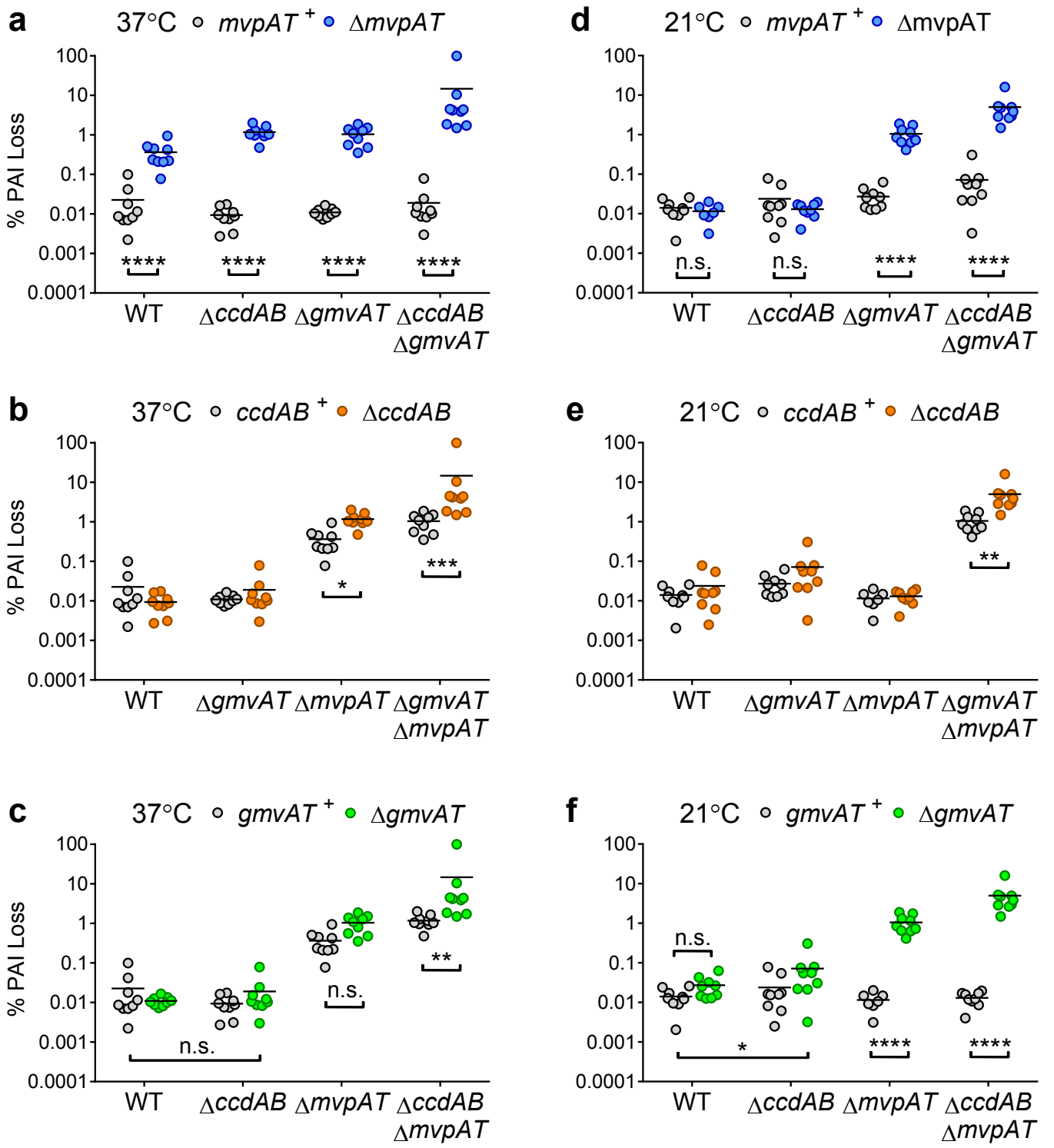
Figure 5. Absence of GmvAT and CcdAB in *S. sonnei* leads to increased pINV loss. a: Comparison between *S. flexneri* M90T $pINV_{Sf}$ (AL391753) and *S. sonnei* 53G $pINV_{Ss}$ (NC_016833). Inner black ring: $pINV_{Sf}$. Selected T3SS- and TA-related ORFs on $pINV_{Sf}$ are indicated as black arrows/boxes. Green outer ring: regions of ≥ 95 % identity of $pINV_{Ss}$ to $pINV_{Sf}$. **b**: Regions where TA systems were introduced to generate *S. sonnei gmvAT* (top) and

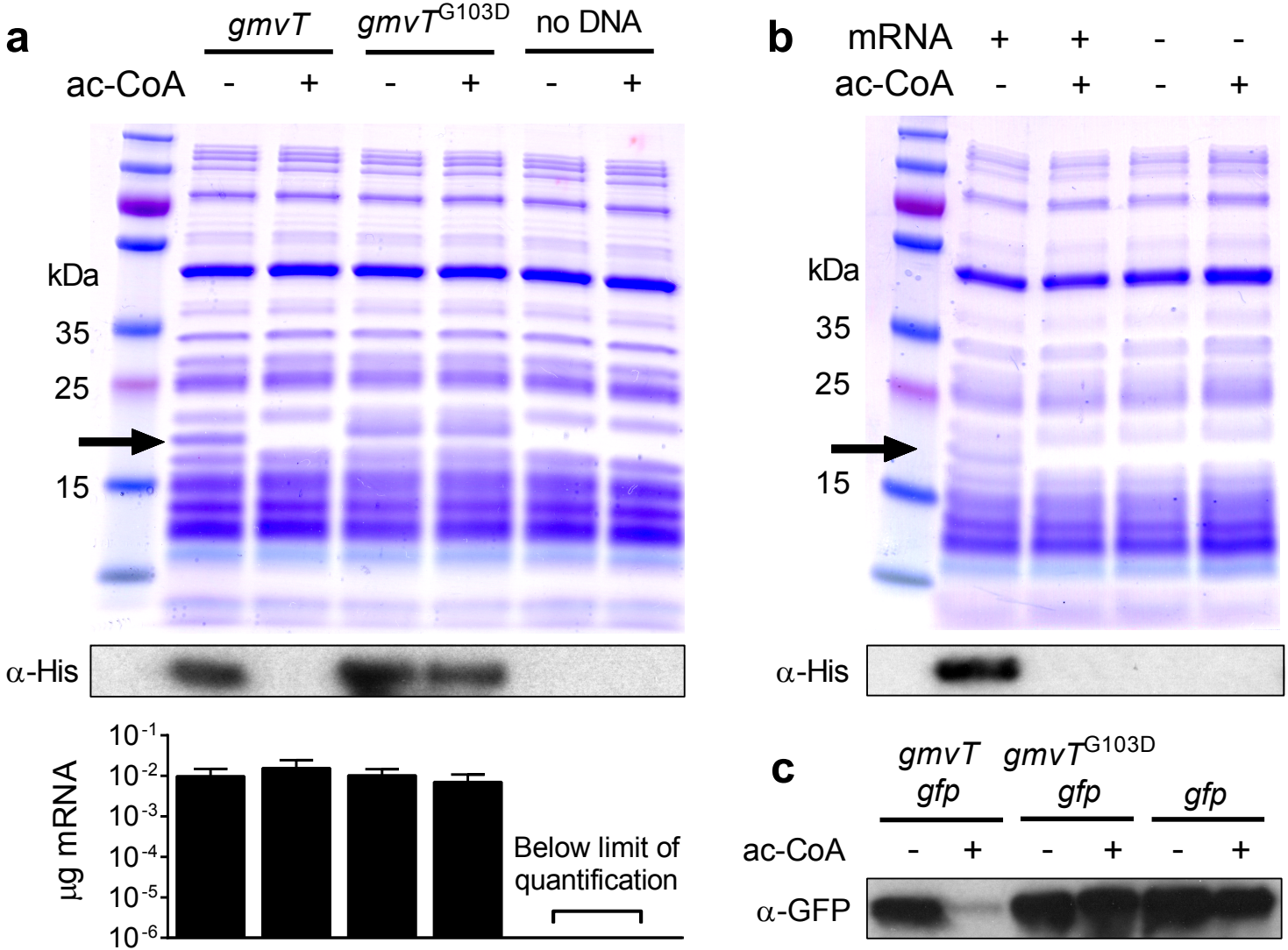
ccdAB (bottom) insertion mutants. **c**: PAI loss in strains with TA insertions grown at 37°C and 21°C for ~ 25 generations as described previously. Insertion of sequences encoding the toxin (WT) or toxoid (G103D/SEK) are indicated. **** $p < 0.0001$; * $p < 0.05$; n.s. not significant by one-way ANOVA with Tukey multiple comparisons test (n = 9 colonies from three independent experiments).

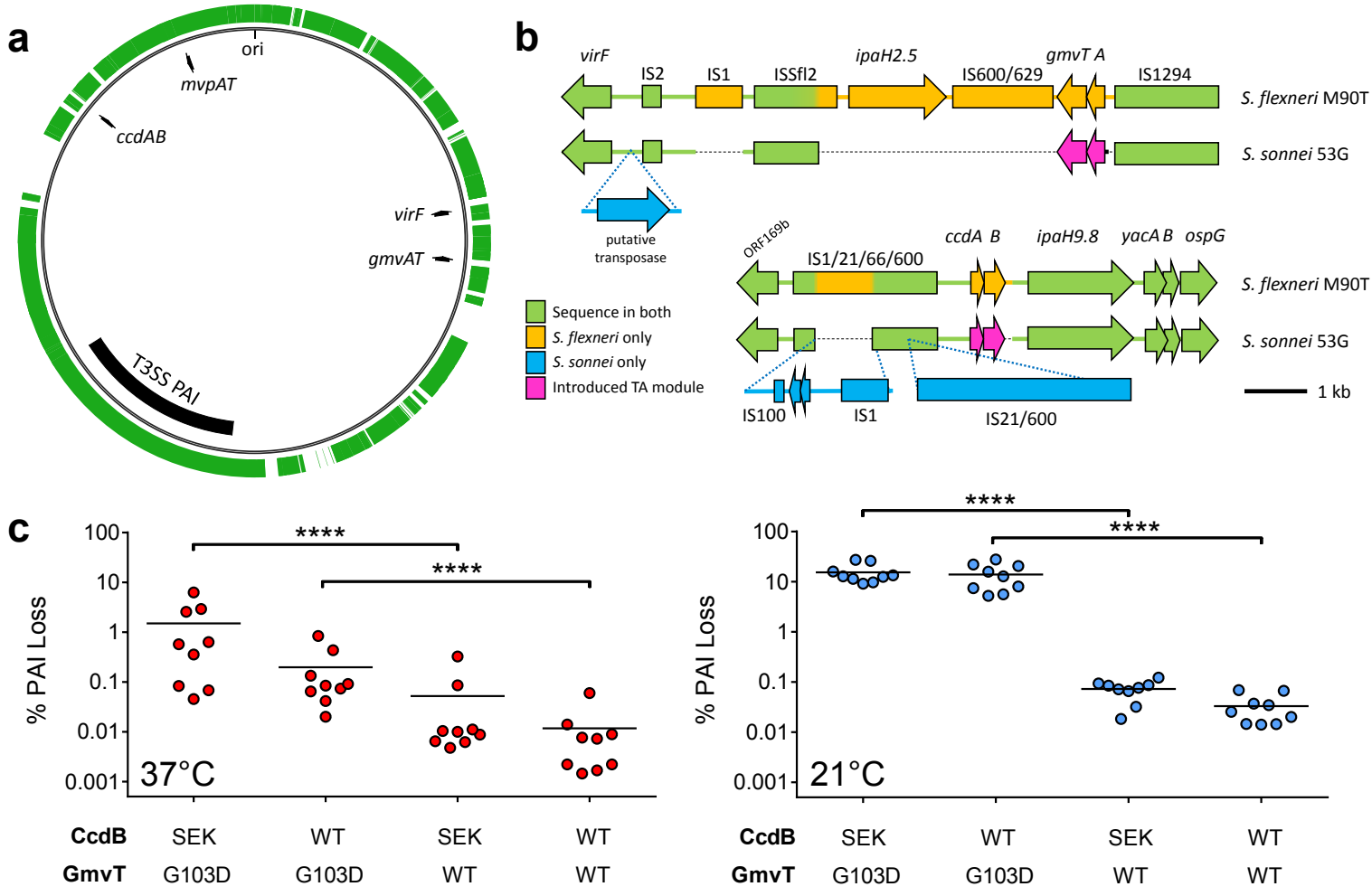
Figure 6. Insertion of CcdAB and GmvAT into wild-type *S. sonnei* stabilises virulence at environmental temperatures. *S. sonnei* 53G, a pINV-cured derivative and a strain carrying pINV encoding CcdAB and GmvAT were grown for approximately 48 generations at 21°C. **a:** Virulent bacteria (forming red colonies) were detected in each culture by plating to Congo red agar. Data show mean ± s.e.m. (n = 3 cultures in independent experiments). **** $p < 0.0001$ by two-way ANOVA. **b:** Epithelial cell monolayers were infected with strains after no or 48 generations of bacterial growth at 21°C. Plaques (indicating lysis of epithelial cells) were counted 72 hours post infection. Data show the mean number of plaques per well ± s.e.m. (n = 4 independent experiments of two wells per strain at each time point). *** $p < 0.001$; n.s. not significant by two-way ANOVA with Sidak multiple comparisons test.

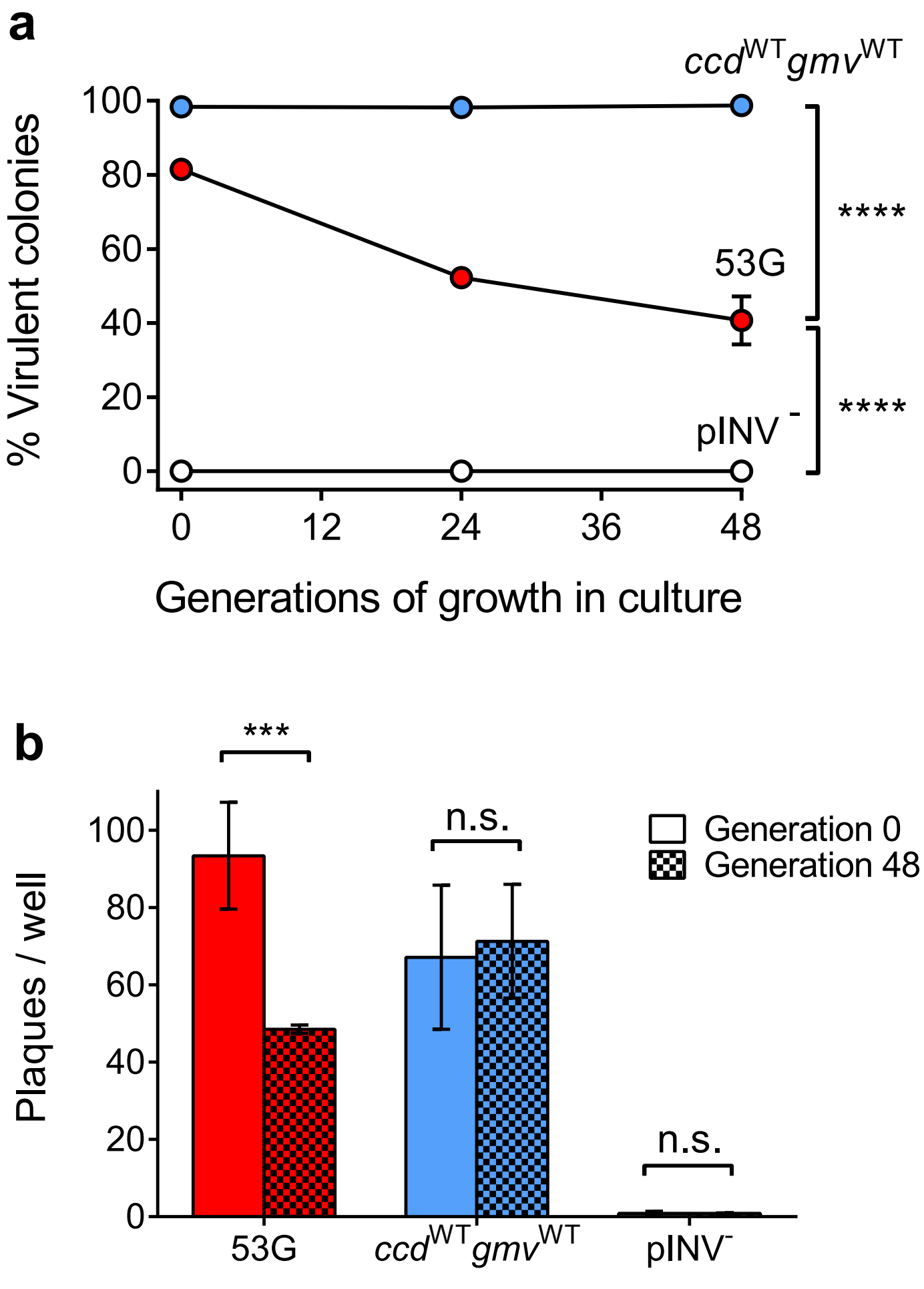

a
S. flexneri
S. sonnei
b
****
**
*
n.s.
100
10
1
0.1
0.01
0.001
0.0001
% Avirulent Colonies
l.o.d.
S. flexneri
S. sonnei
37
30
21
Temperature °C
c
mvpAT+
ΔmvpAT
****
**
n.s.
100
10
1
0.1
0.01
0.001
0.0001
% PAI Loss
37
30
21
Temperature °C
d
100
10
1
0.1
0.01
0.001
0.0001
% PAI Loss
**
n.s.
***
37
30
21
Temperature °C

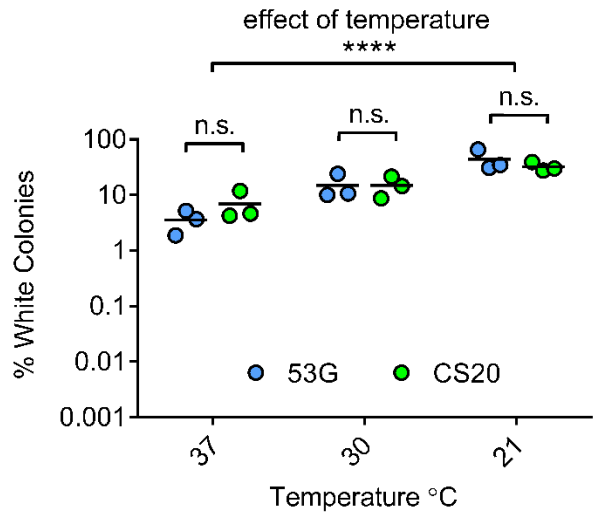

a
CFU / ml
10^{10}
10^{9}
10^{8}
10^{7}
10^{6}
10^{5}
10^{4}
10^{3}
10^{2}
10^{1}
pMvpT pMvpA
pMvpT^{D7A} pControl
pMvpT pControl
****
0
60
120
180
Time post induction (mins)
b
CFU / ml
pCcdB pCcdA
pCcdB^{SEK} pControl
pCcdB pControl
****
0
60
120
180
Time post induction (mins)
c
CFU / ml
pGmvT pGmvA
pGmvT^{G103D} pControl
pGmvT pControl
****
0
60
120
180
Time post induction (mins)


a
37°C
mvpAT +
ΔmvpAT
% PAI Loss
100
10
1
0.1
0.01
0.001
0.0001
****
****
****
****
WT
ΔccdAB
ΔgmvAT
ΔccdAB ΔgmvAT
d
21°C
mvpAT +
ΔmvpAT
n.s.
n.s.
****
****
WT
ΔccdAB
ΔgmvAT
ΔccdAB ΔgmvAT
b
37°C
ccdAB +
ΔccdAB
*
***
WT
ΔgmvAT
ΔmvpAT
ΔgmvAT ΔmvpAT
e
21°C
ccdAB +
ΔccdAB
**
WT
ΔgmvAT
ΔmvpAT
ΔgmvAT ΔmvpAT
c
37°C
gmvAT +
ΔgmvAT
n.s.
n.s.
**
WT
ΔccdAB
ΔmvpAT
ΔccdAB ΔmvpAT
f
21°C
gmvAT +
ΔgmvAT
n.s.
*
****
****
WT
ΔccdAB
ΔmvpAT
ΔccdAB ΔmvpAT

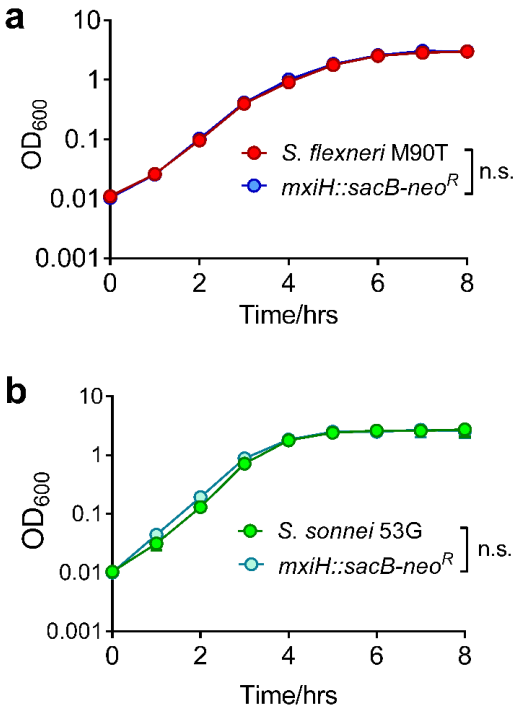

a
gmvT
gmvT^G103D
no DNA
ac-CoA - + - + - +
kDa
35
25
15
α-His
μg mRNA
10^-1
10^-2
10^-3
10^-4
10^-5
10^-6
Below limit of quantification
b
mRNA + + - -
ac-CoA - + - +
kDa
35
25
15
α-His
c
gmvT gfp
gmvT^G103D gfp
gfp
ac-CoA - + - + - +
α-GFP


a
ori
mvpAT
ccdAB
virF
gmvAT
T3SS PAI
b
virF
IS2
IS1
ISSfl2
ipaH2.5
IS600/629
gmvT A
IS1294
S. flexneri M90T
S. sonnei 53G
putative transposase
ORF169b
IS1/21/66/600
ccdA B
ipaH9.8
yacA B
ospG
S. flexneri M90T
S. sonnei 53G
1 kb
IS100
IS1
IS21/600
Sequence in both
S. flexneri only
S. sonnei only
Introduced TA module
c
% PAI Loss
100
10
1
0.1
0.01
0.001
****
****
37°C
CcdB SEK WT SEK WT
GmvT G103D G103D WT WT
****
****
21°C
CcdB SEK WT SEK WT
GmvT G103D G103D WT WT

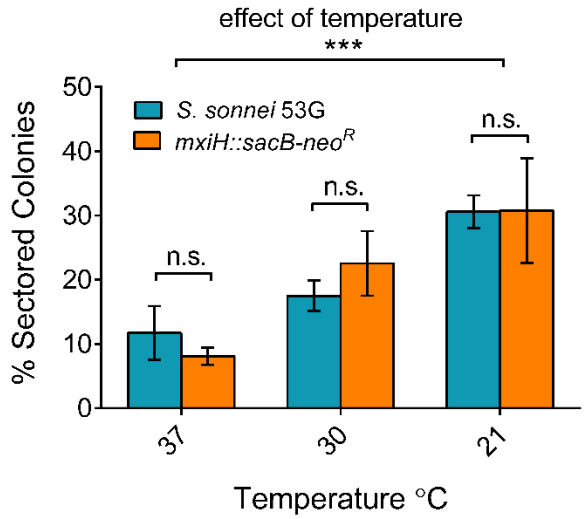

a
ccd^WT gmv^WT
100
80
60
40
20
0
% Virulent colonies
****
53G
****
pINV -
0
12
24
36
48
Generations of growth in culture
b
***
n.s.
n.s.
Generation 0
Generation 48
100
80
60
40
20
0
Plaques / well
53G
ccd^WT gmv^WT
pINV-



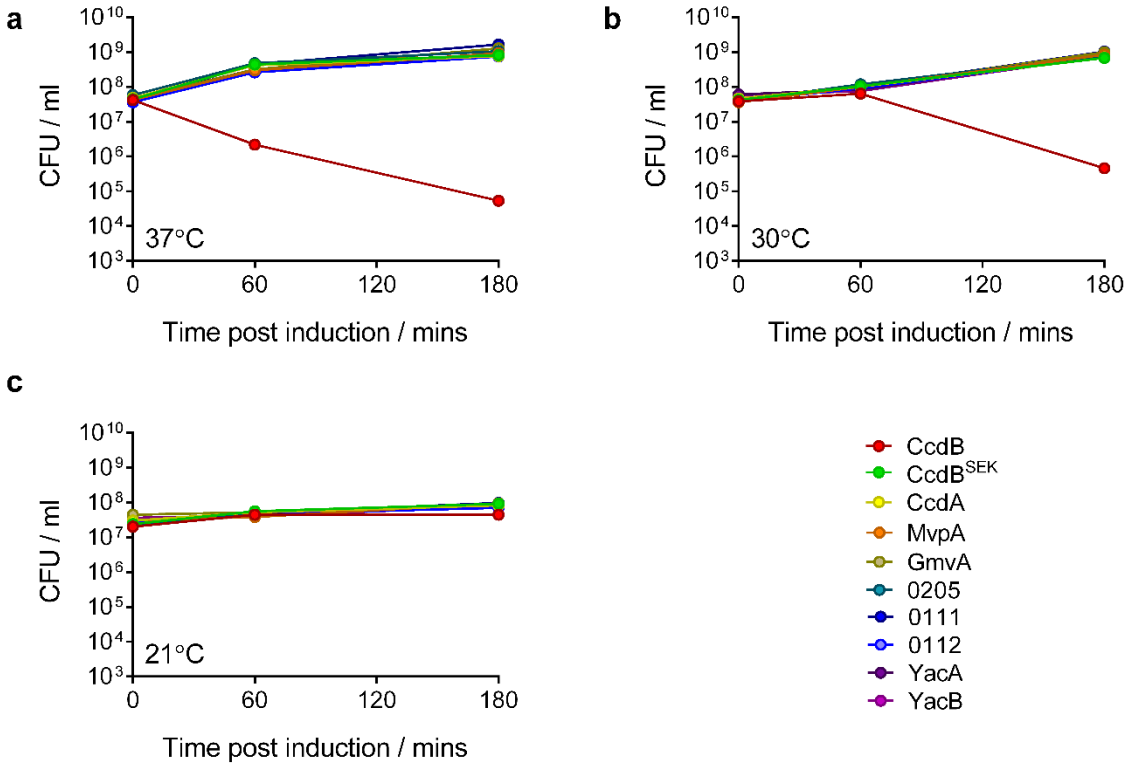
Supplementary Figure 1. Stability of pINV$_{Ss}$ is temperature-dependent in multiple strains. Proportion of avirulent *S. sonnei* 53G (lineage II, data from Fig. 1B) and *S. sonnei* CS20 (lineage III) after growth for ~ 25 generations at the indicated temperatures; solid line, mean (n = 3 independent experiments). Individual bacterial colonies were grown on nonselective LB agar plates for the stated number of generations, then were resuspended in PBS and plated onto the appropriate media for quantification of red/white colonies. Statistical analysis by two-way ANOVA.



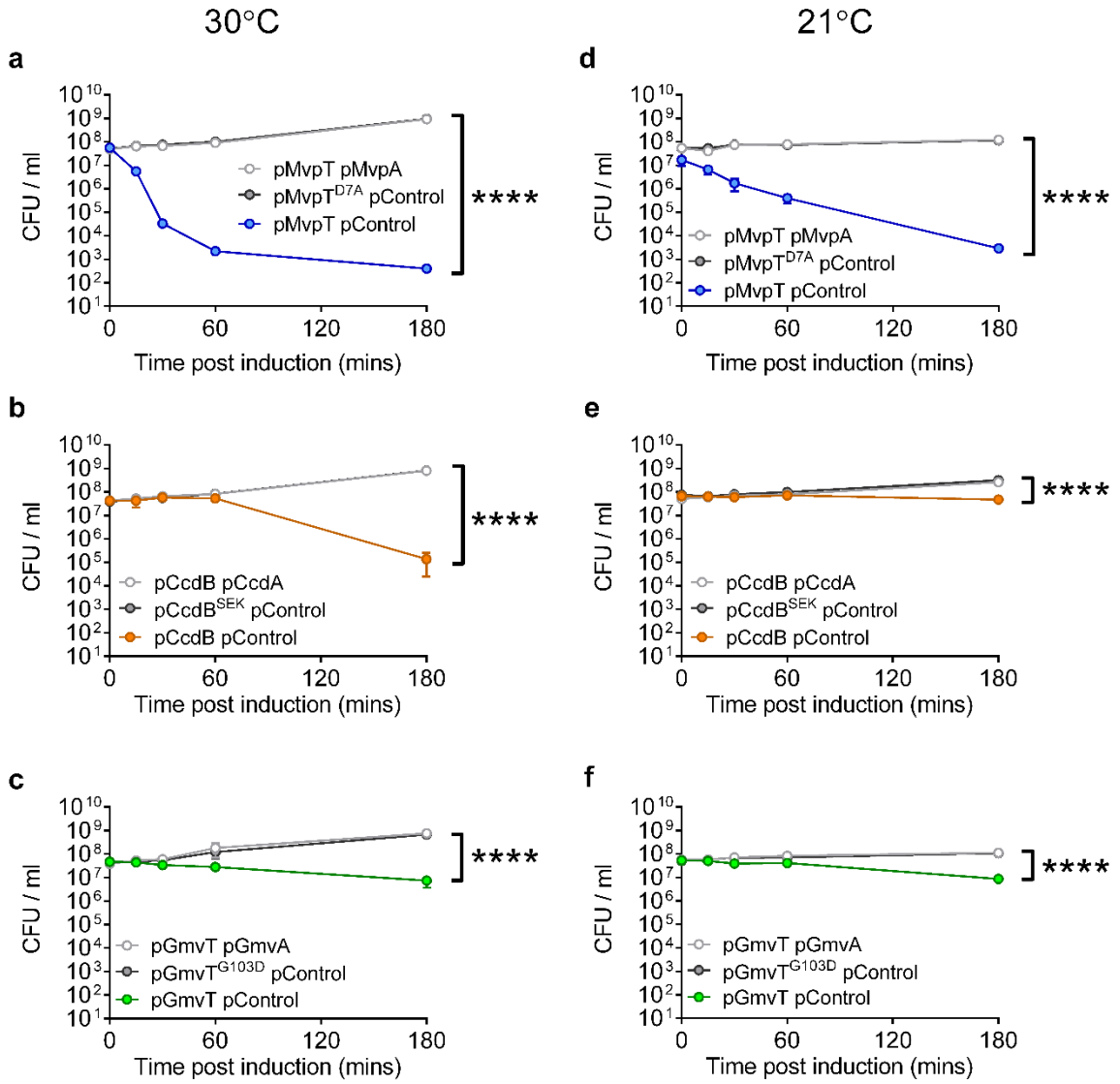
Supplementary Figure 2. The *mxiH*::*sacB-neo*^{R} marker does not significantly affect *Shigella* growth. Growth of *S. flexneri* M90T (A), *S. sonnei* 53G (B) and their *mxiH*::*sacB-neo*^{R} derivatives. Interaction *p* value calculated by two-way ANOVA (n ≥ 3 from at least three independent experiments); n.s. not significant.



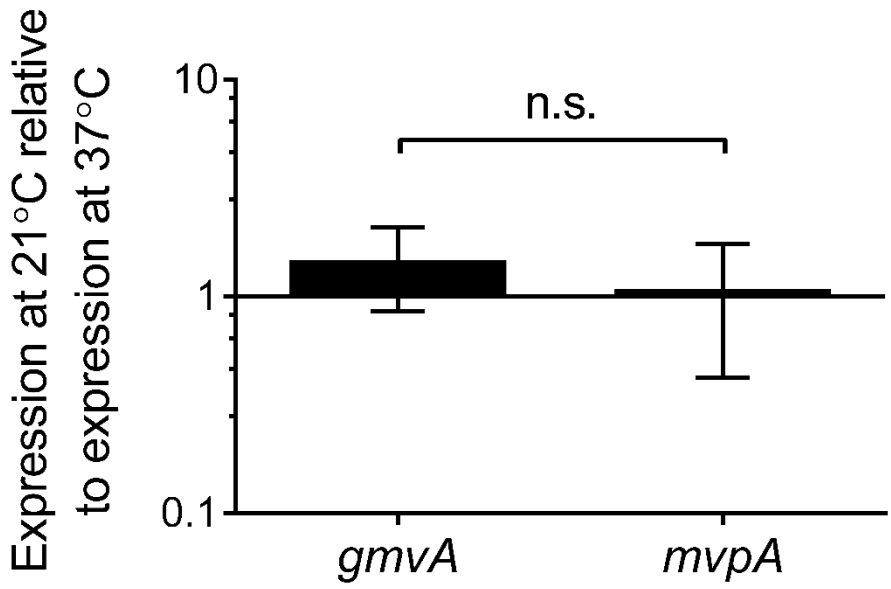
Supplementary Figure 3. Colony sectoring in S. sonnei is temperature-dependent. Proportion of sectored colonies of *S. sonnei* 53G and its *mxiH::sacB-neo*^{R} derivative after growth for ~ 25 generations at the indicated temperatures; mean ± SEM. **** $p < 0.0001$; *** $p < 0.001$; n.s. not significant by two-way ANOVA with Tukey's multiple comparisons test (n = 3 independent experiments).



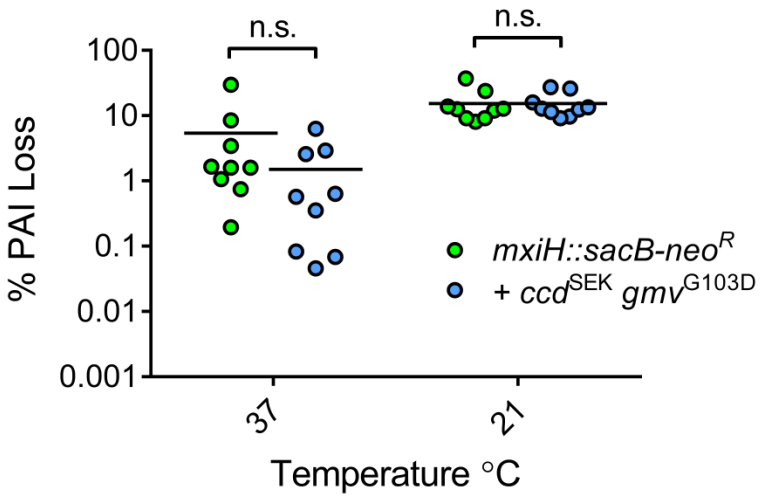
Supplementary Figure 4. Several putative TA-related genes on pINV$_{Sf}$ encode non-toxic products. Viability of *S. flexneri* BS176 following expression of putative TA genes at 37°C (A), 30°C (B) or 21°C (C). CcdB- and CcdB^{SEK}-expressing plasmids were included as controls. Data (n = 1) are representative of experiments carried out under a range of inducer concentrations in various strain/species backgrounds.



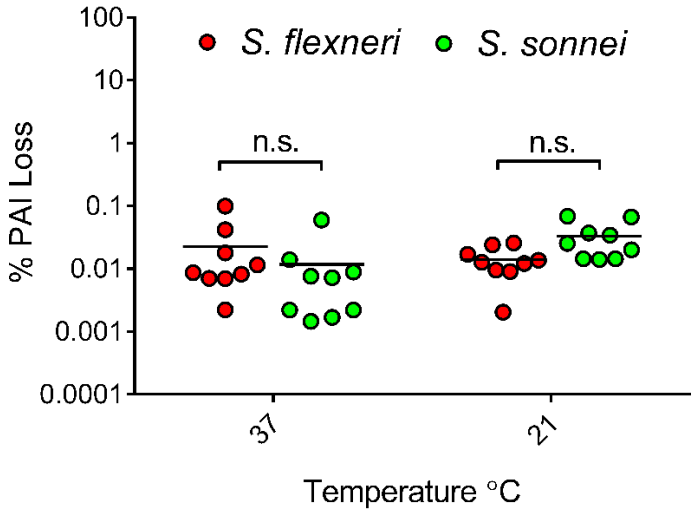
Supplementary Figure 5. CcdAB, GmvAT and MvpAT toxin-antitoxin pairs function at 30°C and 21°C. Strains and experimental conditions are identical to those described in Figure 2A-C, except for growth temperature. A-C, 37°C; D-F, 21°C. Data shows mean ± SEM (n = 3). Two-way ANOVA interaction $p < 0.0001$ for each data set.



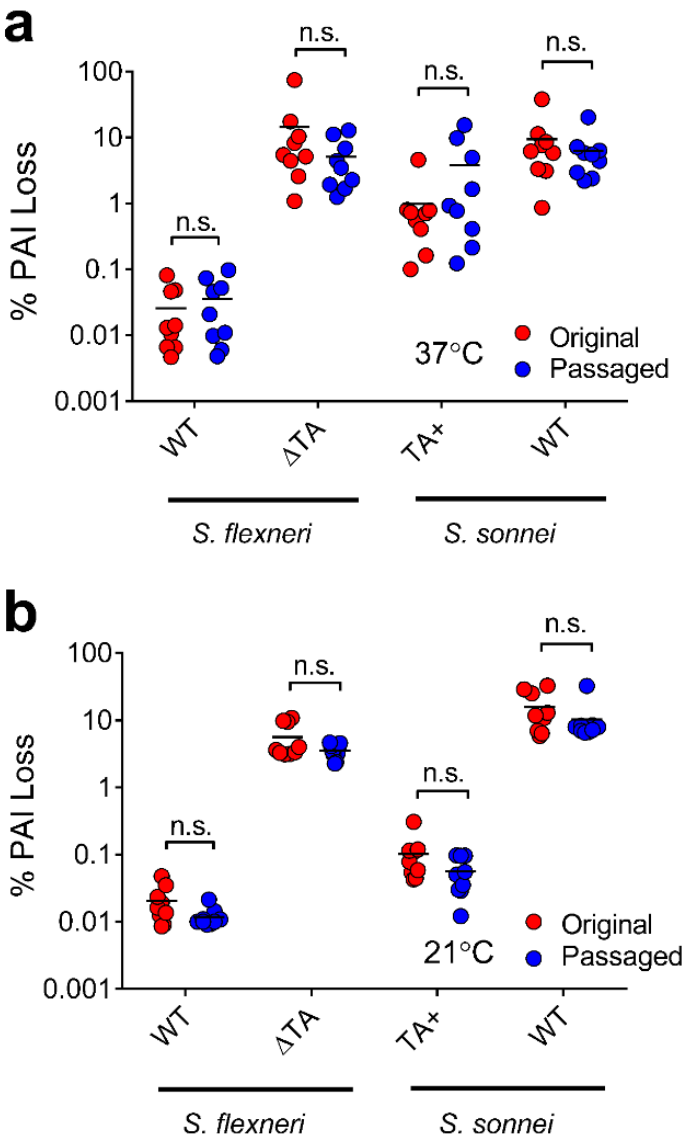
Supplementary Figure 6. Expression of *gmvAT* and *mvpAT* is not temperature-dependent. qRT-PCR analysis of antitoxin gene expression, as a measure of TA operon expression, in cells grown at the indicated temperatures. Mean ± SEM shown (n = 3 independent cultures).



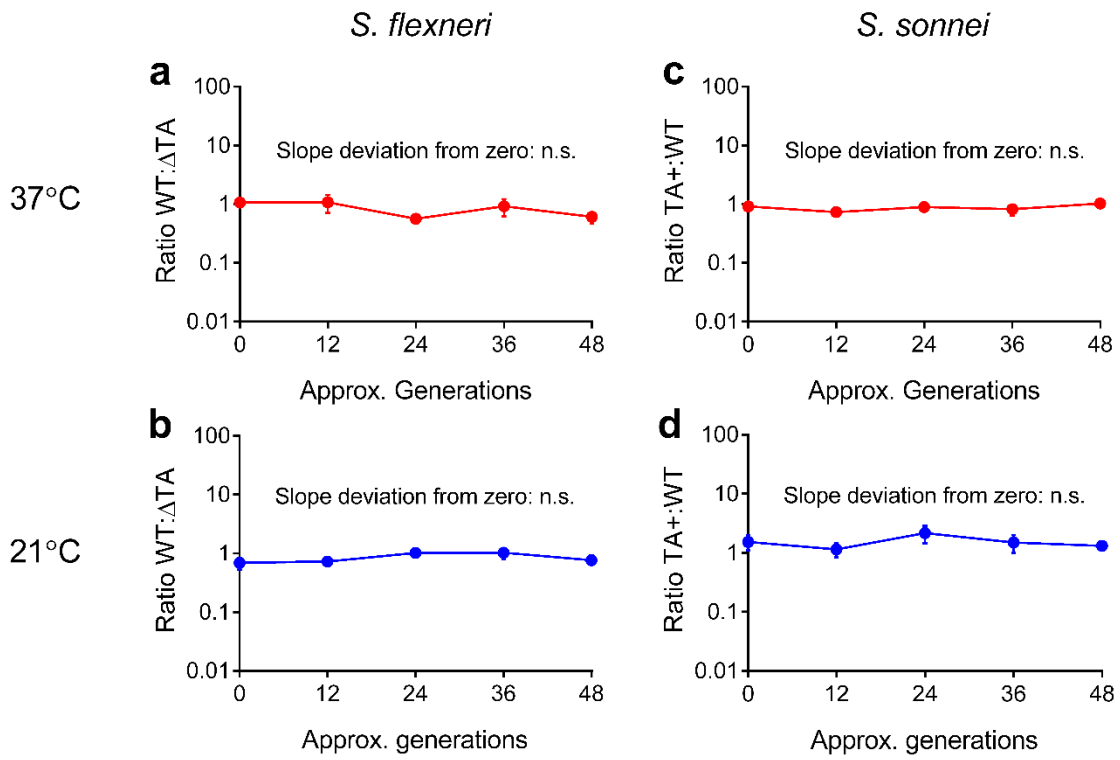
Supplementary Figure 7. Insertion of non-functional TA operons into pINV$_{Ss}$ does not affect plasmid stability. PAI loss in *S. sonnei mxiH::sacB-neo*^{R} and *S. sonnei mxiH::sacB-neo*^{R} *ccd*^{SEK} *gmv*^{G103D}; n.s. not significant by two-way ANOVA with Tukey's multiple comparisons tests (n = 9).



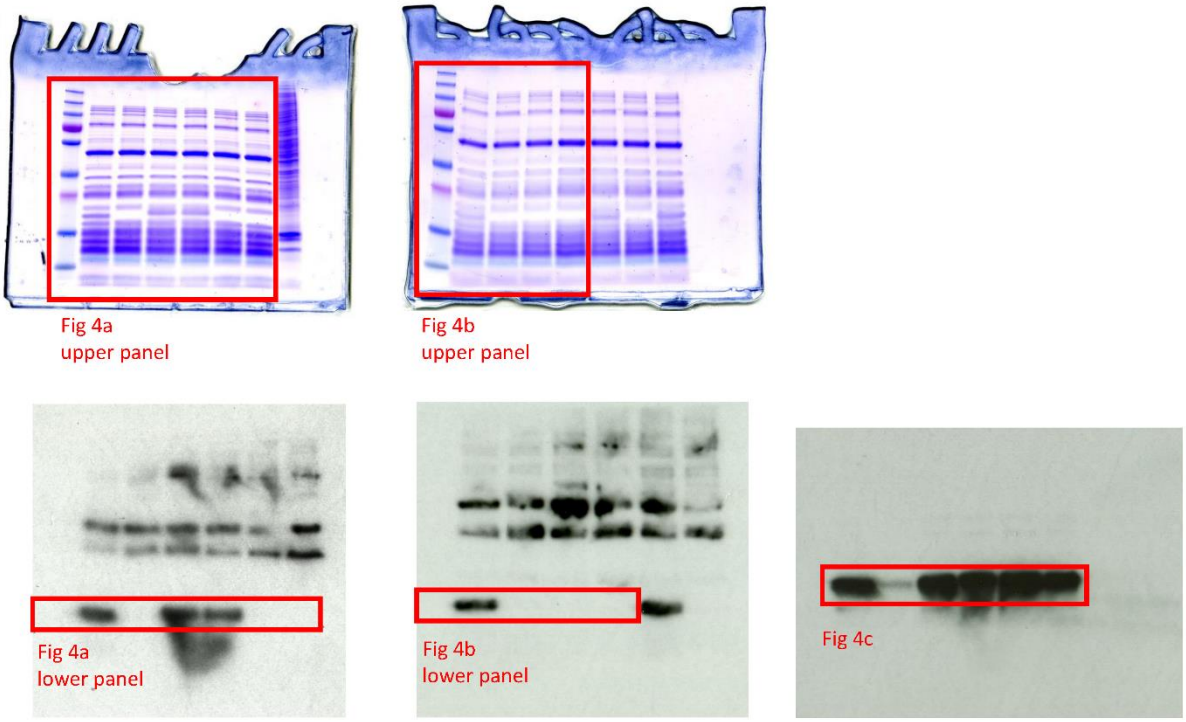
Supplementary Figure 8. Inserting *S. flexneri ccdAB* and *gmvAT* into *S. sonnei* 53G pINV is sufficient to abolish the difference in plasmid stability between the species. Data reproduced from Fig. 1C and 4B for clarity of comparison. PAI loss in *S. flexneri mxiH::sacB-neo*^{R} and *S. sonnei mxiH::sacB-neo*^{R} *ccd*^{WT} *gmv*^{WT}, detected by plating on sucrose and kanamycin; n.s. not significant by two-way ANOVA with Tukey's multiple comparisons tests (n = 9).



Supplementary Figure 9. Retention of the PAI is not due to a sub-population of bacteria. Strains were subjected to a standard PAI loss assay (indicated Original, shown as red circles) at 37°C (a) or 21°C (b), then a second assay carried out was carried out with PAI^{+} bacteria from the first assay (indicated Passaged, blue circles). *S. flexneri* ΔTA lacks MvpAT, CcdAB and GmvAT, while *S. sonnei* TA+ has *ccd*^{WT}*gmv*^{WT} on pINV. n.s. not significant by two-way ANOVA with Tukey's multiple comparisons tests (n = 9).



Supplementary Figure 10. The TA systems do not affect fitness. *S. flexneri* (**a, b**) and *S. sonnei* (**c, d**) *mxiH::sacB-neo*^{R} strains were competed in co-culture with chloramphenicol-resistant *S. flexneri* ΔTA or *S. sonnei* TA+, in the presence of kanamycin to ensure plasmid retention. *S. flexneri* ΔTA lacks MvpAT, CcdAB and GmvAT, while *S. sonnei* TA+ has *ccd*^{WT}*gmv*^{WT} on pINV. After the indicated number of generations at 37°C (**a, c**) or 21°C (**b, d**), bacteria were plated onto selective media to quantify the ratio of strains. n.s. not significant by linear regression analysis (n = 3 independent cultures; error bars show SEM).



Supplementary Figure 11. Complete gel and blot images from Figure 4.

Supplementary Table 1: Bacterial strains and plasmids.

Strain Name	Genotype	Reference
53G	*S. sonnei* lineage II wild-type	(29)
BS176	pINV-cured *S. flexneri* M90T	(38)
GMCT189	M90T *mxiH::sacB-neo*^{R}	This work
GMCT197	M90T *mxiH::sacB-neo*^{R} Δ*mvpAT*	This work
GMCT198	M90T *mxiH::sacB-neo*^{R} Δ*gmvAT*	This work
GMCT199	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB*	This work
GMCT208	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB* Δ*gmvAT*	This work
GMCT209	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB* Δ*mvpAT*	This work
GMCT210	M90T *mxiH::sacB-neo*^{R} Δ*gmvAT* Δ*mvpAT*	This work
GMCT213	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB* Δ*gmvAT* Δ*mvpAT*	This work
GMCT215	53G *mxiH::sacB-neo*^{R}	This work
GMCT232	53G *mxiH::sacB-neo*^{R} *ccdAB*^{WT}*-cat gmvAT*^{WT}	This work
GMCT233	53G *mxiH::sacB-neo*^{R} *ccdAB*^{SEK}*-cat gmvAT*^{WT}	This work
GMCT234	53G *mxiH::sacB-neo*^{R} *ccdAB*^{WT} *gmvAT*^{G103D}*-cat*	This work
GMCT236	53G *mxiH::sacB-neo*^{R} *ccdAB*^{SEK} *gmvAT*^{G103D}*-cat*	This work
GMCT253	M90T *mxiH::sacB-neo*^{R} Δ*clpP::cat*	This work
GMCT254	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB* Δ*gmvAT* Δ*mvpAT* Δ*clpP::cat*	This work
GMCT255	M90T *mxiH::sacB-neo*^{R} Δ*lon::cat*	This work
GMCT256	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB* Δ*gmvAT* Δ*mvpAT* Δ*lon::cat*	This work
GMCT266	M90T *mxiH::sacB-neo*^{R} Δ*mvpAT* Δ*lon::cat*	This work
GMCT267	M90T *mxiH::sacB-neo*^{R} Δ*gmvAT* Δ*lon::cat*	This work
GMCT268	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB* Δ*gmvAT* Δ*lon::cat*	This work
GMCT269	M90T *mxiH::sacB-neo*^{R} Δ*ccdAB* Δ*mvpAT* Δ*lon::cat*	This work
GMCT301	pINV-cured *S. sonnei* 53G	This work
GMCT364	53G *ccdAB*^{WT}*-cat gmvAT*^{WT}	This work
M90T	*S. flexneri* 5a wild-type	(31)
Plasmid	**Purpose/Genotype**	**Reference**
pBAD33	Arabinose-inducible expression vector; *cat*	(55)
pCP20	FLP recombinase for λ Red cassette excision	(52)
pGM085	pBAD33-*ccdB*	This work
pGM087	pGM101-*ccdA*	This work

pGM099	pBAD33-*ccdB*^{SEK}	This work
pGM101	pTrc99-derived vector compatible with pBAD33; *bla*	This work
pGM103	pGM101-*mvpA*	This work
pGM105	pBAD33-*ccdA*	This work
pGM107	pBAD33-*gmvT*	This work
pGM109	pBAD33-*gmvA*	This work
pGM111	pBAD33-*pWR501_0111*	This work
pGM113	pBAD33-*pWR501_0112*	This work
pGM115	pBAD33-*pWR501_0205*	This work
pGM117	pBAD33-*pWR501_yacA*	This work
pGM119	pBAD33-*pWR501_yacB*	This work
pGM121	pBAD33-*mvpA*	This work
pGM123	pBAD33-*mvpT*	This work
pGM129	pGM101-*gmvA*	This work
pGM131	pBAD33-*gmvT*^{G103D}	This work
pGM153	pGM101-*pWR501_0112*	This work
pGM163	pBAD33-*mvpT*^{D7A}	This work
pGM176	pET28a-*gmvT-gmvA*	This work
pGM178	pET28a-*gmvT*^{G103D}-*gmvA*	This work
pIB279	Source of *sacB-neo*^{R} cassette	(53)
pKD3	Source of *camR* cassette for λ Red recombination	(51)
pKD46	Helper plasmid for λ Red recombination	(51)

Supplementary Table 2: Oligonucleotide primers.

Primer	5'-3' Sequence	Construct
GM069	CTGAGGCAAAGAACTCAAATCTTGC	Δ*mxiH*::*sacB-neo*^{R}
GM070	CCGGATCCTTTATCCTCACTTATTTTTATC	Δ*mxiH*::*sacB-neo*^{R}
GM071	CCCATATGTAATAGGGAGCATTCATG	Δ*mxiH*::*sacB-neo*^{R}
GM072	GTTTCTCTTGAACCACCCTGTTTTG	Δ*mxiH*::*sacB-neo*^{R}
GM073	CCGGATCCTTTTTAACCCATCAC	*sacB-neo*^{R} cassette
GM074	CCCATATGTGCAAGCCTCGTCGT	*sacB-neo*^{R} cassette
GM075	CTAGCGAATTCGAGCTTTGCCGATGAGAACAGGGACTG	pGM085
GM076	CCGCCAAAACAGCCATTATATTCCCCAGAACATCAGG	pGM085
GM077	AGGCGAAGCGGCATGGGTACCATTCTCTGTGGG	pGM087
GM078	CCGCCAAAACAGCCATCACCAGTCCCTGTTCTC	pGM087
GM084	CCGCCAAAACAGCCATCGGAAAAGTAAGAACATCAGGTTAATGGCG	pGM099
GM086	GGGAGGCTGTTGGAGCCTATAATG	*mxiH*$^+$
GM089	CGCCATTTTCTAACAGTTCAGAAGG	*mxiH*$^+$
GM101	CTAGCGAATTCGAGCTAACATTAATTGCATAGCAAATTG	pGM109
GM102	CCGCCAAAACAGCCATTATTTCCATTCAGGCTTTAC	pGM109
GM104	CCGCCAAAACAGCCATCAGACTTTATAAAACAAGGTATTAGGTG	pGM107
GM105	AGGCGAAGCGGCATGCGATGCCTGTTACTGCCA	pGM103
GM106	CCGCCAAAACAGCCATCAGAATGACTCCCTTTCC	pGM103
GM107	CTAGCGAATTCGAGCTAAGCTCCTGTGCAGAATG	pGM107
GM108	CTAGCGAATTCGAGCTCATAAGTATGTTTTTGAGGGC	pGM105
GM109	CCGCCAAAACAGCCAAGCTTCACCAGTCCCTGTTCTC	pGM105
GM110	CTAGCGAATTCGAGCTCTCCACAGAGGTAAATGCC	pGM113
GM111	CCGCCAAAACAGCCAAGCTTACACATCCCATTGAGG	pGM113
GM112	CTAGCGAATTCGAGCTCGAAAAATTGCTGGCAAG	pGM111
GM113	CCGCCAAAACAGCCAAGCTTCTATTGAGGGAGTTTAAGAAAC	pGM111
GM114	CTAGCGAATTCGAGCTCGTAATGCAAATATGCATTATTGTC	pGM115
GM115	CCGCCAAAACAGCCAAGCTTAAGGGACACAGCCTAG	pGM115
GM116	CTAGCGAATTCGAGCTCTATACTGTGTATATACGTGGTAATG	pGM117
GM117	CCGCCAAAACAGCCAAGCTTCATTTATCAGCAATCTTCCTG	pGM117

GM118	CTAGCGAATTCGAGCTCATGAAATGAGCAGGAAGATTG	pGM119
GM119	CCGCCAAAACAGCCAAGCTTCACAAGAGATTTTCACACAAAATAAC	pGM119
GM120	CTAGCGAATTCGAGCTCAGACATATCCACATAAGGAG	pGM121
GM121	CCGCCAAAACAGCCAAGCTTCAGAATGACTCCCTTTC	pGM121
GM122	CTAGCGAATTCGAGCTCATGCAGGAAAGGGAGTC	pGM123
GM123	CCGCCAAAACAGCCAAGCTTCAGCTCCAGTCTTCAGTTC	pGM123
GM124	CCTGCGCGATACTCATCATAAACGTATATCCCTTTGACATATCCCGGTATCAATCCCACAATAGATAT ACACAAGACATATCCACATAAGGAGGCAAATATGTGTAGGCTGGAGCTGCTTC	M90T Δ*mvpAT*
GM125	CACCTGGCAGCAGATGCAGCGTCGTGAAGAGGTGAACATCAATGTGCACCGGGAACGCGGGGAGG ATGTTGAGCCGGGAGATGATTTCTGATGAACAGGATGGGAATTAGCCATGGTCC	M90T Δ*mvpAT*
GM128	AGGCGAAGCGGCATGCGCATTTCATTATCAAAATCACATTAAAC	pGM129
GM129	CCGCCAAAACAGCCAAGCTTATTTCCATTCAGGCTTTAC	pGM129
GM130	ACCGAAATCATGCCCCTGGGTACAAAC	pGM131
GM131	AGGGGCATGATTTCGGTAAATGGCTACTAAG	pGM131
GM165	AGGCGAAGCGGCATGCTGGTTCAACGACAAGAGC	pGM153
GM166	CCGCCAAAACAGCCAAGCTTACACATCCCATTGAGG	pGM153
GM174	TGTGTAGGCTGGAGCTGCTT	*cat* cassette
GM175	ATGGGAATTAGCCATGGTCC	*cat* cassette
GM176	GGTCTTTACCGCATTGAG	M90T Δ*ccdAB*
GM177	AAGCAGCTCCAGCCTACACAGATAAGCCCTCAAAAACATAC	M90T Δ*ccdAB*
GM179	GGACCATGGCTAATTCCCATATGTCAGGCTAGGGTCAAAAATC	M90T Δ*ccdAB*
GM180	AGATTAAGAATGCTTTCCGGG	M90T Δ*ccdAB*
GM183	CTCCACAGGTTACAGACAC	M90T Δ*gmvAT*
GM184	AAGCAGCTCCAGCCTACACATAACAGAACGCTATACAGAATC	M90T Δ*gmvAT*
GM186	GGACCATGGCTAATTCCCATTTGTTCAGTCAGTTGTGG	M90T Δ*gmvAT*
GM187	CTCAGTATGTATCGCTGG	M90T Δ*gmvAT*
GM201	GTTTATGCTCGCTACCAACATCTG	pGM163
GM202	CAGATGTTGGTAGCGAGCATAAAC	pGM163
GM208	ACGGCCAGTGAATTCGAGCTCGTGAATGACCTGTTTGCC	53G *ccdAB* insertion
GM209	TACGCTGCTTCATGATAAGCCCTCAAAAACATAC	53G *ccdAB* insertion
GM210	AGGGCTTATCATGAAGCAGCGTATTACTG	53G *ccdAB* insertion
GM211	CCATGGCTAATTCCCATTTATATTCCCCAGAACATCAG	53G *ccdAB* insertion
GM212	AGCCTACACAATACGGTGCAAACAGGCC	53G *ccdAB* insertion

GM213	CTATGACCATGATTACGCCAAGCTTATGCCATATCTGTGATACATCAG	53G *ccdAB* insertion
GM214	TTCCCATTTACTTTTCCGAGAACATCAGGTTAATGGC	53G *ccdAB*^{SEK} insertion
GM215	CGGAAAAGTAAATGGGAATTAGCCATGGTCC	53G *ccdAB*^{SEK} insertion
GM216	ACGGCCAGTGAATTCGAGCTCGGCTTCTCCACTGTGATTCCGTCGTG	53G *gmvAT* insertion
GM217	AATGAAATGCCACGGCAGCCTGGCGGCT	53G *gmvAT* insertion
GM218	GGCTGCCGTGGCATTTCATTATCAAAATCACATTAAAC	53G *gmvAT* insertion
GM219	CCATGGCTAATTCCCATTCAGACTTTATAAAACAAGGTATTAG	53G *gmvAT* insertion
GM220	AGCCTACACAGTTCAGGTAAACGCCTGTTC	53G *gmvAT* insertion
GM221	CTATGACCATGATTACGCCAAGCTTGCGGGTGTAGTAAGCCCT	53G *gmvAT* insertion
GM251	CTGGTGCCGCGCGGCAGCCATATGGAAATAAATGTCACCGCG	pGM176/8
GM252	TTTACCGAAATCATGCCCCTGGGTACAAAC	pGM178
GM253	CAGGGGCATGATTTCGGTAAATGGCTACTAAG	pGM178
GM256	AGTGGTGGTGGTGGTGGTGCTTATTTCCATTCAGGCTTTACCG	pGM176/8
GM257	ACATGGTATATCTCCTTCTCAGACTTTATAAAACAAGGTATTAG	pGM176/8
GM258	AAAGTCTGAGAAGGAGATATACCATGTCTACAGCTGCAAGC	pGM176/8
GM261	ACGGCCAGTGAATTCGAGCTCGAAGGCGGTAAAGCGTCTG	M90T Δ*clpP*
GM262	CCATGGCTAATTCCCATTTCCGTCTCCTGGATAAAATTG	M90T Δ*clpP*
GM263	AGCCTACACACATCGTAATTGATGCCAGAGG	M90T Δ*clpP*
GM264	CTATGACCATGATTACGCCAAGCTTCGCTTGCTTTGTCGGACT	M90T Δ*clpP*
GM267	ACGGCCAGTGAATTCGAGCTCATCATTCAGAAGCTGTTG	M90T Δ*lon*
GM268	CCATGGCTAATTCCCATAGAGCTCTCTCTTAGTTTAATTTC	M90T Δ*lon*
GM269	AGCCTACACATGACCTCGCGCAAAATGC	M90T Δ*lon*
GM270	CTATGACCATGATTACGCCAAGCTTCCCAGTTTCAGCTCACGAG	M90T Δ*lon*
GM302	GCGGTCAGACTGCCAAAAG	qRT-PCR *mvpA*
GM303	CCGTCTAGCTCTCTCTAATCGGTGTCCGTCGAACCATTCG	qRT-PCR *mvpA*
GM304	AAACACCGCGAGAACATCAG	qRT-PCR *gmvA*
GM305	CCGTCTAGCTCTCTCTAATCGGTACGCCAGTTCCATGATGA	qRT-PCR *gmvA*
GM319	AAAACAGACAGCGCAAGAAG	qRT-PCR *repA*
GM320	CCGTCTAGCTCTCTCTAATCGCTCTGTCTGGTAACTGCGGA	qRT-PCR *repA*
GM312	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATATGGGCAGCAGCCATCATCATCATCAT	PURExpress

		GmvT
GM313	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTATCAGACTTTATAAAACAAGGTATT	PURExpress GmvT
GM314	AGGGGCATGGTTTCGGTAAATGGCTACTAAG	PURExpress $GmvT^{G103D}$
GM315	ACCGAAACCATGCCCCTGGGTACAAAC	PURExpress $GmvT^{G103D}$
GM332	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATATGAGCAAAGGAGAAG AACTTTTCACT	PURExpress sfGFP
GM333	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTATCATTTGTAGAGCTCATCCATGCC	PURExpress sfGFP
GM336	CGCACGTTCGTTATTTGCTT	qRT-PCR *gmvT*
GM337	CCGTCTAGCTCTCTCTAATCGAACAACAGGGACAGGATTGG	qRT-PCR *gmvT*
GM338	GCGAATTAATACGACTCACTATAGGGCTTAAGTATAAGGAGGAAAAAATATGCATCATCATCATCAT CACAGCAGCGGCATGGGATGTGTAACTGCACCAGAACCT	PURExpress pWR501_0111
GM339	AAACCCCTCCGTTTAGAGAGGGGTTATGCTAGTTACTATTGAGGGAGTTTAAGAAACAA	PURExpress pWR501_0111
qRT_Tag	CCGTCTAGCTCTCTCTAATCG	qRT-PCR