Going around in circles: virulence plasmids in enteric pathogens

Giulia Pilla and Christoph M. Tang*

Sir William Dunn School of Pathology,

South Parks Road,

University of Oxford,

Oxford OX1 3RE,

United Kingdom

for correspondence: christoph.tang@path.ox.ac.uk

tel: 44- 1865- 275560

Abstract

Plasmids play a major role in the development of disease caused by enteric pathogens. Virulence plasmids are usually large (> 40 kb) low copy elements, and carry genes which promote host:pathogen interactions. Although virulence plasmids provide advantages to the bacteria in specific conditions, they often impose fitness costs on their host. Here we review virulence plasmids in Enterobacteriaceae that are important causes of diarrhoea in humans, *Shigella* spp., pathovars of *Escherichia coli*, *Salmonella* spp. and *Yersinia* spp.. We contrast these plasmids with those which are routinely used in the laboratory, and outline the mechanisms by which virulence plasmids are maintained in bacterial populations. We highlight examples of virulence plasmids which contain multiple mechanisms for their maintenance (*e.g.* toxin:antitoxin systems, partitioning systems), and speculate how these might contribute their propagation and success.

Bullet points

- Large, low copy number plasmids in Enterobacteriaceae contribute to host:pathogen interactions
- Plasmid copy number is determined by the replicon and can vary during intestinal colonisation
- Virulence plasmids carry partitioning systems which segregate copies of the plasmid into daughter cells
- Toxin:antitoxin systems on plasmids mediate post-segregational killing, and can have localised effects
- Several plasmids have multiple maintenance systems

Introduction

Several important properties in bacteria, including antibiotic resistance and virulence, are encoded on plasmids. Indeed, several pathogens have arisen following the acquisition of a single plasmid. Plasmids establish long-standing associations with certain bacterial lineages, and can spread by horizontal transfer through transformation and/or conjugation. However, the expression of virulence genes or resistance mechanisms encoded on these extrachromosomal elements often imposes significant fitness costs on their bacterial host.

The biology of plasmids was intensively studied from the 1960s, initially as mediators of anti-microbial resistance (*i.e.* R factors [1]) then as tools for molecular biology. This work provided a wealth of understanding about the replication and maintenance of plasmids, and led to the construction of numerous plasmids for cloning and gene expression. Laboratory plasmids are available with various copy number (often > 50 per cell), and are usually small (≤ 5 kb) to facilitate their manipulation. They also carry antibiotic resistance genes for selection, with engineered sequences to allow introduction of heterologous genes, together with promoters to enable precise gene regulation (**Figure 1A**). In the laboratory, plasmids are usually introduced into bacteria by transformation. Virulence plasmids are distinct as they tend to occur in low and often single copy number (<10 per cell), and are generally larger than 40 kb (**Figure 1B**). Furthermore, virulence plasmids are often transferred between bacteria by conjugation either through systems encoded on the plasmid itself or by another plasmid (helper plasmid) in the same cell.

In recent years, there has been little interest in plasmid biology as many vectors are now available with the properties required for molecular biology. However, whole genome

sequencing of pathogens has revealed the extent and diversity of virulence and resistance plasmids, and there has been a renaissance in research to define how these selfish elements contribute to host:pathogen interactions and the mechanisms that ensure their propagation.

The purpose of this review is to outline the contribution of plasmids to the virulence of pathogens, and highlight recent advances in our understanding of mechanisms that are necessary for their maintenance in bacterial populations. We focus on the pathogenic members of *Enterobacteriaceae* [**Box 1**], which contain related bacteria that occupy the same niche in the body, the intestinal tract, and harbour plasmids with similar features. Some **enteropathogens**, such as *Salmonella enterica* serovar Typhi, can invade the intestinal epithelium and cause systemic disease; we will not consider these bacteria here, and instead focus on bacteria that cause disease at the mucosal surface of the intestine, and, thence, share similar habitats. We discuss how plasmid copy number is controlled, and how plasmids are segregated during cell division. Furthermore, we describe how plasmids guarantee their maintenance by post-segregational killing (PSK), which occurs when a daughter cell fails to receive a plasmid and is usually mediated by Toxin:Antitoxin (TA) systems. However, in the past few years, novel functions have been discovered for TA systems on plasmids that contribute to bacterial fitness and virulence, and these will be discussed here.

The contribution of plasmids to the virulence of enteric pathogens

Virulence plasmids are abundant in enteric pathogens

As well as residing in the same niche, enteric pathogens also share the common feature of harbouring virulence plasmids, which, depending on the species and even the strain, differ

in their composition as well as in their contribution to pathogenesis. For example, *Shigella* spp. possess a highly conserved large virulence plasmid (pINV), [2] which enables the bacterium to invade into epithelial cells, and, thus, is essential for virulence. Virulence plasmids are also widely distributed in different ***Escherichia coli*** **pathotypes** (**Box 1**), conferring a broad range of characteristics (**Table 1**) [3] but contribute to host:pathogen interactions to varying degrees. The large virulence plasmid of Enteroinvasive *E. coli* (EIEC) carries virulence genes which are closely related to those on *Shigella* pINV, pINV [4] but is usually significantly larger. Indeed, the EIEC and *Shigella* plasmids appear to have evolved through independent insertion/deletion events, while retaining a high level of homology in their virulence genes [4-6], suggesting they were acquired from a common ancestor [5, 7]. For Enteropathogenic *E. coli* (EPEC) and Enterohaemorrhagic *E. coli* (EHEC), plasmids appear to be less critical to virulence (**Table 1**) and strains belonging to the same species and serotype can harbour different plasmids. EHEC O157:H7 is the commonest serotype of EHEC causing bloody diarrhoea and **haemolytic uraemic syndrome** (HUS), and many strains of this serotype harbour the well characterised virulence plasmid pO157 (~ 92 kb) [8, 9], which contributes to (HUS) and adherence to intestinal epithelial cells[7]. In EPEC, large virulence plasmids are also abundant, with pB171-8 (~ 69 kb) the prototypical EPEC plasmid conferring localised adherence to intestinal epithelial cells (**Table 1**) [7, 10]. In Entertoxigenic *E. coli* (ETEC) the impact of plasmids on virulence is highly strain-dependent. Plasmids in ETEC vary in the colonisation factors, toxins, and virulence-related factors they encode [7], with pCoo (~ 98 kb) being long considered the prototypic ETEC plasmid [11]. Similar to ETEC, the composition of plasmids in EAEC (pAA) varies greatly [12], with a high diversity in plasmid-encoded fimbriae [7, 12] and toxins [13, 14] (**Table 1**).

Many serotypes of *Salmonella enterica* subsp. I, the predominant cause of human infection with *Salmonella* [15], carry plasmids that vary in size from 50 to 100 kb, all of which have conserved virulence genes [16]. Additionally, a plasmid, referred to as pYV, is critical for the virulence of *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*, two leading agents of foodborne and zoonotic gastroenteritis (**Table 1**). This ~ 70 kb element mediates growth restriction of bacteria at 37°C and is necessary for the expression of several virulence factors [17, 18]. Although the composition and overall architecture of pYV differs between species and strains [18-20], the regions conferring virulence are highly conserved, and are closely related to those on plasmids from *Shigella* and EIEC, suggesting that they were acquired from a common ancestor [18, 21]. However, in contrast to the invasive properties conferred by pINV, pYV encodes factors which impede the uptake of *Yersinia* by professional phagocytes, such as macrophages [22].

Virulence plasmids of *Shigella* and *Salmonella:* cell invasion and intracellular survival

Plasmid-encoded genes can influence different steps in pathogenesis, such as adhesion, invasion, colonisation, and modulation of host immune responses (**Figure 1**). Therefore, because of their high variability, we will not attempt to list all functions conferred by every plasmid. Instead, we highlight two selected examples**,** the virulence plasmids in *Shigella* and *Salmonella* (**Figure 2**) because their contribution to pathogenesis have been well characterised and are distinct.

Shigella is a human-specific pathogen and the leading cause of bacillary dysentery [23], estimated to be responsible for 164,300 deaths in 2015 [24]. *Shigella flexneri* is the main cause of shigellosis in low income countries, while *Shigella sonnei* is prevalent in more developed

societies [25, 26]. *Shigella boydii* and *Shigella dysenteriae* infections are less frequent overall but are important causes of disease in particular settings and geographic regions [25, 26].

Although most of our knowledge about *Shigella* and its virulence plasmid is derived from studies of *S. flexneri*, plasmid sequences from other *Shigella* species are now available, allowing understanding of their distinctive epidemiological and pathological features. pINV confers on *Shigella* the ability to invade into intestinal epithelial cells [27]. A 32 kb **pathogenicity island** (PAI), also known as *ipa-mxi-spa* locus, lies at the core of the plasmid and encodes for the assembly of a Type three secretion system (T3SS) [28]. The PAI contains 38 genes, half of which encode for the T3SS apparatus, while the remainder encode for chaperones, regulators and effectors. The effectors are proteins secreted through the T3SS into the host cell cytoplasm and enable *Shigella* to penetrate and survive inside epithelial cells, as well as induce macrophage killing via **pyroptosis** [29-32]. Therefore, T3SS PAI is necessary for fundamental steps in *Shigella* virulence.

Comparative analysis of plasmid sequences demonstrates that the *ipa-mxi-spa* locus is highly conserved among *Shigella*. The presence of the **insertion sequences**, IS*100* and *IS600*, flanking the PAI and the absence of IS elements within the PAI indicates that it was acquired by *Shigella* species from a common ancestor [28]. Moreover, loss of the T3SS PAI occurs spontaneously and frequently in *S. flexneri* [33], and is mediated by intra-molecular recombination events between ISs flanking the PAI [34]. Interestingly, through a similar IS-mediated recombination process, pINV can integrate into the chromosome. This leads to downregulation of expression of genes in the T3SS PAI, due to a reduction in the expression of the two plasmid-encoded transcription factors *virB* and *virF* [34] and spontaneous excision of pINV from the chromosome restores T3SS

function [34]. These events potentially provide an alternative reversible mechanism that *Shigella* employs to modulate expression of the T3SS.

Virulence-associated genes outside the T3SS PAI are mainly implicated in post-invasion events, such as the modulation of host immune responses (*e.g osp* and *ipaH* family genes, and *sepA*) and those (*i.e. virA*, *virG*, *sopA* and *phoN2*) which mediate actin-directed movement of bacteria in the host cell cytosol [31, 35, 36]. The ShET2 enterotoxin is also encoded outside the T3SS PAI and is responsible for the characteristic watery diarrhoeal phase of shigellosis [36]. Like the T3SS PAI, these genes are often flanked by ISs, suggesting that they have been acquired horizontally in blocks, resulting in the mosaic structure of pINV [28].

Salmonella is a major foodborne pathogen, accounting for 93.8 million of cases and 155,000 deaths per year worldwide [15]. Among all *Salmonella enterica* subspecies, *Salmonella enterica* subsp. I is the predominant cause of human infection [15]. Many *Salmonella* strains carry plasmids containing the *spvRABCD* (*Salmonella* plasmid virulence) locus [16] (**Table 1**). The Spv proteins enhance bacterial replication within the intracellular environment of host cells, in particular macrophages [37, 38], and are involved in infection of systemic organs [39]. For instance, SpvC is a phosphothreonine lyase which modulates host inflammatory responses *in vivo* during early stages of infection by irreversibly inactivating MAPK signalling in infected cells [40]. This results in an increased bacterial dissemination which is essential for a systemic infection [41]. Although the composition of the plasmid differs between serovars even among the same subspecies, the *spv* locus is conserved [16]. In *Salmonella* strains that do not harbour the plasmid (*e.g.* subspecies II, IIIa, IV and VII), the *spv* locus is found on the chromosome, highlighting its importance and suggesting it has been acquired independently on several occasions,

maintained through positive selection [42] and may have been transferred from the plasmid to the chromosome .

Some *Salmonella* plasmids contain additional genes that are involved in pathogenesis, such as the *pef* (plasmid-encoded fimbriae) locus, *rck* and *rsk* (**Table 1**). The *pef* locus encodes for fimbrial structures that enhance adhesion to host tissues [16]. In particular, Pef mediates adhesion to the murine small intestine and is responsible for fluid secretion during intestinal salmonellosis [43]. However, Pef does not contribute to adhesion to some human cell lines, suggesting that these fimbriae may confer host and cell specificity to the bacterium [44]. Rck and Rsk mediate resistance against host complement. Rck is an outer membrane protein highly homologous to Ail from *Yersinia* spp. [45], and is able to inhibit polymerization of the complement component C9 on the bacterial cell surface [46]. Recent studies have also confirmed a significant role for Rck during bacterial internalization, in particular by inducing entry [47]. Rsk (reduced serum killing) is also involved in resistance to the bactericidal activity of human complement, but the mechanism for this remains unknown. Plasmid integration events have also been identified in *Salmonella*, and following integration, expression of plasmid-encoded *rsk* is down-regulated [48].

Controlling numbers: less is more for virulence plasmids

As well as carrying genes involved in pathogenesis, virulence plasmids contain systems for their replication and maintenance, which ensure their transmission to daughter cells. These systems are highly regulated, interconnected, and tightly coordinated. Because large plasmids represent an increased DNA load during replication and usually impose a considerable metabolic burden, they are generally present in low copy number [49]. The copy number,

together with **plasmid incompatibility**, are determined by the plasmid **replicon**, which includes a replication origin and systems for replication control.

Antisense RNA-regulated and iteron-regulated replicons in prototypic plasmids

Initial studies on plasmid replication were conducted on the **prototypic plasmids** F, R and R100 in *E. coli*, which are paradigms for the function and mechanisms of plasmid maintenance and the classification of plasmid replicons. Plasmid replication systems can be divided into two main groups: antisense RNA-regulated and iteron-regulated replicons [18, 50, 51] (**Figure 3**). Both types of replicons contain a Rep protein and an origin (*ori*), where DNA replication is initiated, then proceeds unidirectionally according to the **theta replication** model [52] . Binding of Rep to *ori* is essential to render the DNA in this region single stranded and, allow recruitment of the DNA replication machinery. However, the systems these two replicons employ to control plasmid replication are different.

In the first group, plasmid replication is controlled by a short antisense transcript which binds the 5' end of the longer of two Rep transcripts, preventing its translation (**Figure 3A**). Because the antisense transcript is unstable, low concentrations of the antisense are not sufficient to prevent replication. R1 and R100 plasmid replicons are well characterised antisense-regulated replicons [53, 54], sharing 98% nucleotide identity of their replication regions [55, 56], with the R1 RepFIIA used as a model for antisense RNA-controlled plasmid replication (**Figure 3A**). It contains the replication origin (*ori*) and the gene encoding the replication initiation protein, RepA. Upstream of *repA* is *tap*, which encodes a short leader peptide whose translation is coupled with and required for *repA* translation [54]. Plasmid copy number is mainly controlled by the antisense transcript *copA*, which inhibits RepA translation by preventing translation of

Tap [54]. The *copA* antisense RNA determines the incompatibility group to which the plasmid belongs, thus, *copA* is also referred to as *incRNA* [18]. The second element that controls plasmid copy number of RepFIIA is CopB, a protein that represses *repA* transcription by binding its promoter, *PrepA*. Transcription of *copB* is constitutive, so high CopB concentrations in the cell from high plasmid copy number inhibits plasmid replication [54].

In iteron-regulated replicons, plasmid replication is controlled by iterons, directly repeated sequences, specifically bound by Rep. There are two of sets iterons, one located at the origin and one located downstream *rep* (**Figure 3B**). With low plasmid copy number and low concentrations of Rep, Rep binds to the iterons at the origin, allowing the initiation of replication. However, when Rep concentrations increase due to a rise in plasmid copy number, Rep can also bind the second set of iterons, resulting in sequestration of Rep (through the Titration Model) or formation of Rep-DNA complexes that interact with each other (the Handcuffing Model) [18, 51, 57, 58], preventing DNA replication.

In contrast to plasmid R1 and R100, the F plasmid harbours three replicons, two of which, RepFIA and RepFIB, are functional and representative of the iteron-regulated replication systems [59, 60]. RepFIA has an origin of replication, *oriS*, which comprises tandem **DnaA boxes** at one end, and iterons that can be bound by monomeric RepE, the replication initiation protein [50] (**Figure 3B**). RepE, together with DnaA, unwinds the DNA duplex allowing DNA replication [61, 62]. The *repE* gene is downstream *oriS* and is subject to autoregulation, due to the iteron-like sequences in its promoter [63]. The second set of iterons, located downstream *repE,* is termed *incC*. The RepFIB replicon has a similar genetic organisation as RepFIA (**Figure 3B**), and is sufficient for F plasmid replication [59, 64].

256

257 **Replicons in virulence plasmids**

258 All large virulence plasmids harboured by enteropathogens are duplicated by theta
259 replication [65] and fall in the IncF incompatibility group (**Table 1**). The *Shigella* invasion
260 plasmids pINV possess replicons which are closely related to the Rep FIIA replicon from R100
261 [2]. In particular, the *S. flexneri* pINV replicon shares approximately 75% identity with the R100
262 replicon at the nucleotide and amino acid levels. However, *inc* from pINV differs from *inc*
263 genes of IncFII plasmids, suggesting that pINV might fall in a different incompatibility group
264 [28]. Furthermore, there is significant homology between *Shigella* and EIEC replicons [66].

265

266 The presence of more than one replication region on a plasmid is seen in the F plasmid and
267 also in some virulence plasmids. For instance, both EPEC pB171 and EHEC pO157 possess
268 multiple replication regions which belong to different replicon groups. pB171 has replication
269 origins related to RepFIIA and RepFIB origins. The sequence of pB171 RepFIIA is 93% identical
270 to the corresponding sequence of R100 and the RepFIIA replicon of pO157 [10]. In contrast, the
271 negative regulator *copB* shows higher homology with the corresponding gene of IncFVII
272 plasmids, indicating that the RepFII replicon of pB171 is likely to be a mosaic of replication
273 regions from IncFII and IncFVII plasmids [10]. Interestingly, the second pB171 replicon has high
274 identity with the RepFIB replicon of *S.* Enteriditis and, is likely to be functional based on
275 sequence analysis [10]. Besides the RepFIIA replicon, two other replication regions are present
276 on pO157. Like pB171, the second pO157 replicon belongs to the RepFIB family [8], while the
277 third falls into RepFIA group although it is probably inactive due to an insertion sequence [8].
278 ETEC pCoo has two functional origins of replication [11]; one is homologous to the replication
279 region of the conjugative IncI1 plasmid R64, while the second is highly related to the R100

RepFIIA replicon. Analysis of the sequence that separates the two replicons suggests that pCoo has emerged following **co-integration** of two independent replicons, as this region contains direct repeats that might have been involved in the recombination event [11].

Distinct from the plasmids in other *E. coli* pathotypes, Enteroaggregative *E. coli* (EAEC) pAA plasmids have a variety of replicons, that are single or multiple, and belong to RepFIIA, RepFIB or RepFIC replicon groups [7, 12]. This suggests that EAEC plasmids have been subject to extensive rearrangements/exchange or have had different ancestors and acquired the virulence genes on different occasions [7]. Different replicons are also found on *Salmonella* virulence plasmids. In particular, two replicons have been identified, *repB* and *repC*, which resemble RepFIIA and RepFIB replicons, respectively, with lower degree of identity compared with other plasmids. Therefore, it is likely that following acquisition, the systems have been subject to diversifying selection, leading to their divergence from the IncFII and IncFI incompatibility groups [16, 67, 68].

Recombination is frequent during plasmid evolution and can involve the replication region. For example, most *Yersinia* pYV carry an IncFII-related replicon, although *Y. enterocolitica* pYVe8081 (pYV of serotype 0:8) has a replicon with a mosaic structure in which RepA is related to the incompatibility group IncL/M. This has been suggested to result from recombination events between pYVe8081 and these plasmids [18] and is an example of the high genetic variation found in *Yersinia* plasmids. Several lineages have been identified, probably as a result of independent gene gain, gene loss and genetic rearrangements occurred on different occasions [69, 70].

The control of plasmid replication in *Y. pseudotuberculosis* is crucial for pathogenesis as it influences virulence gene expression. The plasmid copy number increases from one to four *in vitro* to up to twelve *in vivo,* in response to an increase in temperature. The change in copy number leads to an increase in expression of genes on the *yscM-yopD* PAI, enabling *Yersinia* to increase its virulence in response to a rise in temperature [71]. It will be interesting to see whether other virulence plasmids share similar mechanisms of regulation.

Separate pathways: the complexity of partitioning systems in virulence plasmids

Genetic organisation and mechanisms of partitioning systems

Partitioning systems promote the stability of plasmids during cell division by directly segregating plasmids into daughter cells. Partitioning systems are characterised by an **NTPase** (either an ATPase or GTPase), necessary for the process of segregation, DNA *par* sites, which act as centromere-like elements, and a **centromere-binding protein (CBP)** that binds the **DNA *par* sites**, forming a partitioning complex. Two types of partitioning systems, type I and II, are found on virulence plasmids (**Table 1**). Type I loci (*e.g. parABS*) contain **Walker-type ATPases**, while type II loci (*e.g. parMRC* and *stbAB*) utilise an **actin-like ATPase** [72]. Type I systems can be divided into type Ia, of which ParAB and SopAB are representatives, and type Ib, of which ParAB from pB171 is an exemplar (**Figure 4A)** [72]. Type Ib systems differ from type Ia in their genetic organisation, regulation, the absence of a Helix-Turn-Helix (HTH) motif in ParA, and the structure of ParB.

Type I and II systems employ fundamentally different modes for partitioning, designated as "pulling" and "pushing", respectively [73]. According to the "pulling" model, the ParA NTPase polymerises on nucleoid DNA, forming extended filaments. Contact with the partitioning complex stimulates depolymerisation of ParA filaments, pulling the plasmid in the direction of ParA depolymerisation (**Figure 4B**). Recent *in vitro* studies have supported an alternative view of type I partitioning, termed the 'diffusion ratchet' model [74] [75] (**Figure 4B**). It is hypothesised that ParA binds dynamically and non-specifically to nucleoid DNA, and is released when the partitioning complex interacts with ParA. This generates an area of ParA-free nucleoid DNA around the partitioning complex. The slow re-association and the fast de-association of ParA from DNA forms a local ParA concentration gradient on DNA, driving plasmid segregation [76] (**Figure 4B**).

In contrast, in the type II "pushing" model, the partitioning complex serves as nucleation point for the polymerisation of the ATPase (ParM), which forms bundles of actin-like filaments, that push the partitioning complex-bound plasmid towards the cell poles [77] (**Figure 4B**).

Segregating virulence plasmids

Most virulence plasmids harbour a type I, ParAB-type partitioning system which generally resembles the P1 or F plasmid system (**Table 1**) [72]. For example, *Shigella* pINV contains a ParAB-like element that is highly conserved across all *Shigella* spp. [29]. However, in *S. flexneri* a second potential partitioning system, named StbAB, is found, approximately 90 kb from *parAB* [28]. StbAB is similar to the ParMR-like system of the same name found on R100 [29, 55, 78]. Interestingly, this element is not present *S. sonnei* pINV, yet found in plasmids from *S. dysenteriae* and *S. boydii.* Two *par* loci have been described on pB171. In contrast with the *S.*

flexneri partitioning systems, the pB171 *par* loci, named *par1* (type II) and *par2* (type I), are adjacent to each other and share a common *par* site but are oriented in opposite directions (**Figure 5**) [79]. Interestingly each system is sufficient for plasmid segregation although both systems are needed for optimal plasmid maintenance. There is regulatory cross-talk between the two systems, where the ParB from *par2* represses the transcription of *par1* in addition to regulating its own expression (**Figure 5**) [80]. Similarly, the *S.* Typhimurium virulence plasmid harbours a locus which appears to contain two partitioning systems (**Table 1**), one of which is closely related to the P1 plasmid ParABS system. The second potential system might be formed by ParA and a DNA sequence *incR* [16, 81].

In contrast, the ETEC plasmid pCoo has two partitioning systems (a ParAB-like and a StbAB-like system) which probably result from the origin of the plasmid as co-integrate of two plasmids (**Table 1**). The ParAB-like system is located in the R100 region of pCoo, while the StbAB-like system is derived from R64 [11]. In contrast, EHEC pO157, *Yersinia* pYV, and EAEC pAA each possess a single partitioning system similar to the F plasmid type Ia system (pO157 and pYV), and ParMR type II system (pAA) (**Table 1**) [8, 82].

Toxin-Antitoxin Systems: multiplicity and alternative functions

Post-segregational killing (PSK) provides a further mechanism that plasmids employ to ensure their vertical transmission to daughter cells. PSK involves plasmid-encoded Toxin:Antitoxin (TA) systems which are usually composed of a protein toxin and a cognate “antidote” (antitoxin). The antitoxin is less stable than the toxin as it is specifically targeted by cellular proteases, such as Lon or Clp [83]. However, high levels of antitoxin expression ensure its

continual replenishment [84]. When the plasmid is inherited by a daughter cell, the antitoxin expressed from the plasmid counteracts the action of the stable toxin. If the plasmid is lost, any residual antitoxin in the daughter cell is degraded by proteases and no longer replaced, so the toxin arrests cellular growth, leading to PSK.

TA systems are classified into six groups depending on the nature of the antitoxin and its mechanism of action. Type II TA systems, which comprise a protein toxin and protein antitoxin [85], are the most abundant on virulence plasmids, with multiple systems often found on a single plasmid (**Table 4**). The VapBC system is one of the most characterised family of type II TA systems, and consists of the toxin VapC, a site-specific endonuclease that cleaves $tRNA^{fMet}$, and the VapB antitoxin, which forms a hetero-octameric complex with VapC, leading to its inactivation [86, 87]. The same system has been named MvpAT (for maintenance virulence plasmid) in *Shigella*, and is highly conserved in all *Shigella* pINVs [2] and essential for pINV maintenance at 37°C [88]. VapBC also contributes to plasmid maintenance in *S.* Typhimurium [89].

Interestingly recent studies have indicated that TA systems might have roles aside from mediating PSK. An example is the plasmid-encoded VapBC system in *Salmonella*, in which expression of the toxin VapC increases significantly when the bacterium is in the intracellular compartment [90], and increases bacterial fitness in this environment [90]. As VapBC is only found in pathogenic *Salmonella* serovars, the acquisition of this system, together with other TA modules, may contribute to the pathogenic lifestyle of *Salmonella* [90].

CcdAB is another widespread plasmid-encoded type II TA system. This TA system is composed of the toxin CcdB, which targets **DNA gyrase** by binding its GyrA subunit, and the antitoxin

CcdA. which binds the toxin, blocking its interaction with GyrA [91, 92]. Homologues of this system are present on pINV, pB171, pCoo, pO157 and *Salmonella* pSLT. Interestingly, the *ccdAB* locus is conserved in all *Shigella* species except *S. sonnei*, which contains remnant of *ccdA* [88]. The *ccdAB* locus is less widely distributed than *vapBC* in *S. enterica* serovars, and frequently absent in less virulent strains [93]. Furthermore, in *S*. Typhimurium an amino acid substitution in the pSLT-encoded *ccdB* renders it non-functional [90], so this TA system does not contribute directly to stability [89]. However, *ccdAB* is still transcribed and part of a **poly-cistronic operon** composed of other four downstream genes, including one that encodes a putative **resolvase** possibly involved in the resolution of plasmid multimers [89]. Therefore, this version of *ccdAB* may act as a regulatory system, and exert indirect functions that confer plasmid stability by modulating transcription of downstream genes. The *ccdAB* loci in pB171, pO157 and pCoo share high sequence identity with the locus in the F plasmid, suggesting that this element retains its function in these plasmids [10, 82].

In general, virulence plasmids have multiple TA systems. For example, aside from *ccdAB*, pB171 harbours *relBE*, which encodes a **mRNA interferase**, RelE, that cleaves mRNA, and its cognate antitoxin, RelB [94 95]; pAA of EAEC 042 also possesses a second TA module belonging to the *relBE* family [82]. EHEC pO157 has a second TA module, related to the F plasmid *flmAB*, which is, in turn, homologous to the R1 plasmid *hok*/*sok* system and is a type I TA system (**Table 4**). The toxin Hok is a trans-membrane protein that causes membrane depolarisation [96]; its mRNA is targeted by the antitoxin *sok*, a small antisense RNA that prevents Hok translation by interfering with ribosome binding [97, 98 99]. pCoo and pAA also have the same TA system [11 82].

Recently, another family of type II TA systems has been identified on *S. flexneri* pINV, which harbours a toxin that is a Gcn5 N-acetyl transferase (GNAT). Only two homologues of this TA system has been characterised previously, in *Salmonella* and *E. coli* EHEC O157:H7, but they are chromosomally-encoded [100, 101]. Named GmvAT (GNAT Maintenance of Virulence) in *Shigella*, where GmvA is the antitoxin and GmvT is the toxin, this TA system is absent in *S. sonnei* pINV but found in *S. boydii* and *S. dysenteriae* pINV [88]. The toxin GmvT arrests bacterial growth by blocking translation in an acetyl-CoA-dependent manner. However, the mode of action and the target of the toxin are not known in *Shigella*. A study on the *Salmonella* GmvAT homologue suggested that the toxin acetylates elongator tRNAs [101], while the related toxin in EHEC N-acetylates the initiator $tRNA^{fMET}$ [100]. Although the toxins might have different targets, they both cause impaired translation.

Conversely, plasmid-encoded TA systems have not been well characterised in *Yersinia* pYV. Sequencing database analysis has revealed that a ParDE family TA system [102] is commonly present among some *Yersinia* plasmids [103], however, not much has been investigated in terms of its biological role and contribution to the plasmid maintenance of pYV.

Conclusions

Virulence plasmids are critical elements in many enteropathogens. Although the structure and the contribution of these plasmids to virulence depend on the species and even strain, the plasmids share similar features (**Figure 1**). Maintenance is a fundamental aspect of plasmid biology that is highly regulated through their replicons, partitioning systems, and TA modules. Comparative analysis of maintenance systems among different enteropathogens affords a better understanding of their conservation and, therefore, their importance for plasmid stability. In many cases, maintenance loci on different plasmids are closely related. However, this is not the case in some enteropathogens, which have plasmids equipped with unique maintenance systems. In these cases, many questions remain about how they contribute to the plasmid success, and modulate host:pathogen interactions.

Interestingly, plasmids in several enteropathogens have multiple replicons, partitioning systems and TA modules. The reasons for this remain obscure. However, combinations of maintenance systems might allow efficient plasmid maintenance in distinct environments, such as the TA systems on *S. flexneri* pINV, which act at different temperatures [88]. Alternatively, multiple systems might reflect an evolutionary intermediate, with one system in the process of replacing another [80] or simply be remnants of past recombination events. Additionally, cross-talk could occur between different systems, similar to the regulatory cross-talk between *par1* and *par2* in pB171[80], or the cross-talk between chromosomally and plasmid-encoded TA systems [104, 105].

Alternative functions of maintenance systems could also explain the presence of multiple elements on a single plasmid **(Figure 6**). Multiple replicons, which might allow plasmids to

circumvent incompatibility, in a similar way as partitioning systems do [106], or provide a means to alter plasmid copy number for regulating plasmid gene expression [71]. As chromosomally-encoded TA systems can enhance biofilm formation and stress response [107], so plasmid-encoded TA systems might offer alternative functions, by promoting bacterial adaptation to different conditions. Otherwise, they might control the expression of nearby genes [89] or stabilise different regions on a plasmid [34]. For example, *Shigella* pINV frequently loses its T3SS PAI via IS-mediated rearrangements, and a TA system on this element has recently been shown to exert localised effects through a process termed post-recombinational killing, when loss of part of a plasmid containing a TA system leads to bacterial death. Interestingly, MvpAT is located near the ori of pINV from *Shigella* spp., indicating that its location might be critical for its role in maintaining the entire plasmid [34]. The TA system, GmvAT, is found near to the key regulator *virF* on *S. flexneri* pINV, with both genes flanked by copies of *IS*1294 that could mediate recombination. Thus, GmvAT might select against this event.

Another factor shared between virulence plasmids is genetic plasticity. This is reflected in the diversity of plasmids recently circulating in current populations of bacteria, indicating that new combinations will occur and emerge in the future. Co-integration by recombination events between plasmids belonging to different incompatibility group is often observed and results in chimeric megaplasmids [108, 109]. Whole genome sequencing should reveal and track the evolution of virulence plasmids during the spread of infections, in the same way it was applied to document the emergence of resistance.

Recombination can also occur between plasmids and the chromosome. In the case of *Salmonella* and *Shigella* plasmids, recombination events can result in integration of the

plasmid into the chromosome [34, 48]. This offers a strategy by which bacteria avoid the fitness costs often associated with carrying a plasmid, as integration typically leads to reduced gene expression and subsequent excision of the plasmid restores virulence in both species. A similar bi-stable situation occurs in *Yersinia* spp but is achieved by regulation of the copy number of pYV in response to the temperature [71].

Many questions about the biology of virulence plasmids remain unanswered, and not just in Enterobacteriaceae. Anthrax lethal toxin and the protective antigen are encoded on plasmids, while several plant pathogens, such as *Pantoea agglomerans* and *Agrobacterium tumefaciens*, harbour plasmids which are essential for virulence [110-112]. Furthermore, in some cases virulence plasmids contribute not only to virulence but also to resistance to antibiotics and other environmental stresses [108, 113, 114]. Much is still to be learned about how plasmids are transferred within bacterial populations by transformation and/or conjugation. *Shigella* and *Yersinia* virulence plasmids harbour inactive mobilisation systems, suggesting that they probably acquired the plasmid by conjugation. However, both these plasmids can be mobilised with the help of a co-resident conjugative plasmid [27, 115]. Understanding the mechanisms of plasmid replication, partitioning, and post-seregational killing should offer insights that can be employed to eliminate virulence plasmids from bacterial populations [116].

BOX 1: Enterobacteriaceae, a large family of commensal and pathogenic bacteria

Enterobacteriaceae are a large and heterogeneous family of Gram-negative bacteria comprising non-spore-forming bacilli that are aerobic or facultative anaerobes [117]. Also termed "enterics", their name mainly refers to numerous genera that live in the animal and human gastro-intestinal tract, either as harmless components of the flora (*e.g.* commensal *Escherichia coli*), representing a small portion of the human microbiota [118], or as pathogens (e.g. *Salmonella*, *Shigella*, *Yersinia*, pathogenic *E. coli*). Many species can be found in soil, either during faeco-oral transmission, or as free-living organisms (*e.g. Buttiauxella* and *Rahnella*) and some are associated with plants (*e.g. Erwinia* genus). Some members are also opportunistic pathogens, such as *Klebsiella, Proteus* and *Enterobacter* spp. [119, 120], which can cause disease in immunocompromised individuals [121, 122].

The human pathogenic Enterobacteriaceae cause similar diseases, characterised by multi-step pathogenesis, generally consisting of colonisation +/- invasion of the intestinal mucosa, where they trigger inflammation, leading to symptoms such as diarrhoea, dysentery and fever. *Escherichia coli* is the most abundant and well-studied member of the Enterobacteriaceae, and there are five recognised enteric pathotypes of this species: enteroinvasive (EIEC), enterohaemorrhagic (EHEC), enteropathogenic (EPEC), enterotoxigenic (ETEC), and enteroaggregative (EAEC) *E. coli*. These bacteria constitute an important cause of disease [123,124,125,126], with EHEC and EIEC responsible for outbreaks in wealthy countries and all pathotypes are endemic in less wealthy countries [3, 127]. *Shigella* has been often considered as another pathotype of *E. coli* given their close evolutionary relationship [128, 129] and is the leading cause of bacillary dysentery in humans worldwide. *Salmonella* and *Yersinia* are two further important pathogenic members of the Enterobacteriaceae.

DEFINITIONS:

Enteropathogen: a pathogenic bacterium that infects the intestinal tract of humans and animals causing diarrhoea, gastroenteritis and localised lymphadenitis.

***Escherichia coli* pathotypes**: classes of pathogenic *E. coli* which are characterised by their capacity to cause specific diseases through a defined set of virulence factors.

Haemolytic uraemic syndrome: disease characterised by acute renal failure, haemolytic anaemia (*i.e.* inappropriate destruction of erythrocytes) and thrombocytopaenia (low levels of circulating platelets).

Yersiniosis: acute gastrointestinal infection caused by *Y. enterocolitica* or *Y. pseudotuberculosis*, characterised by enteritis, diarrhoea and fever. Rarely associated with more severe complications such as ileitis, septicaemia and acute arthritis.

Pathogenicity Island: a region on a chromosome or plasmid containing clusters of virulence genes often flanked by mobile genetic elements or direct repeats that could mediate mobility of the entire region.

Pyroptosis: mechanism of inflammatory cell death, characterised by a rapid disruption of the plasmalemma driven by stimulation of the pore-forming activity of Gasdermin D, and accompanied by the concomitant release of pro-inflammatory cytokines, such as IL-1β and IL-18, and chromatin fragmentation.

Insertion sequence (ISs): short transposable DNA elements that can move on the same DNA molecule or between different DNA molecules. They are only composed of genes encoding proteins implicated in mobility such as transposases and regulatory elements. Distinct from transposons, ISs do not carry any accessory genes (*e.g.* encoding antibiotic resistance).

Plasmid incompatibility: a phenomenon whereby two plasmids cannot co-exist in the same bacterial cell. It occurs when plasmids share one or more elements that control their replication, partitioning or copy number. Based on sequence homology, plasmids are classified into different incompatibility groups, so plasmids belonging to the same group are incompatible with each other, but compatible with plasmids in different incompatibility groups.

Replicon: DNA region that includes genes that are sufficient for plasmid replication and copy number control, and where replication is initiated. Depending on the sequence of the replicon, plasmids are classified into different replicon groups.

Prototypic plasmid: some of the first discovered plasmids that have been employed as models to study plasmid biology.

Theta replication: a mechanism of replication whereby the synthesis of the leading and lagging DNA strands are coupled, leading to the formation of theta-shaped intermediates, from which the mechanism derives its name.

DnaA boxes: short stretches of DNA which are bound by DnaA. The interaction between DnaA and DnaA boxes localised at the origin is essential to unwind DNA before the start of DNA replication.

Co-integration: phenomenon that leads two circular plasmids to combine, maintaining the sequence of each plasmid intact, thus, producing a single plasmid from two.

NTPase: Nucleoside triphosphatase; a family of enzymes that catalyses the hydrolysis of a nucleoside triphosphate (NTP) to a nucleoside diphosphate (NDP). The reaction releases

574 energy, often inducing a conformational change in protein structure that allows the protein
575 to drive other chemical reactions.

576 **Centromere-binding proteins (CBP)**: A family of proteins that specifically binds to
577 centromere-like DNA sites, which can contain multiple CBP-bound sequences and therefore
578 be recognised by multiple CBP, leading to the formation of nucleoprotein complexes.

579 **DNA *par* sites:** centromere-like DNA sequences often containing repeated sequences and
580 specifically bound by CBPs; it is required in cis for plasmid partitioning and forms the
581 partitioning complex when associated with CBP.

582 **Walker-type NTPase**: NTPases characterised by Walker motifs, amino acid sequences that
583 play an important role in NTP binding and hydrolysis.

584 **Actin-like ATPase**: a family of ATPases that contain ATP-binding domains that are homologous
585 to those in actin. For ParM, the structure of the ATPase resembles that of actin, implying that,
586 like actin, ParM can form filaments.

587 **DNA gyrase:** group of essential enzymes defined as topoisomerases and responsible for ATP-
588 dependent conversion of relaxed DNA into a negatively supercoiled form.

589 **Poly-cistronic operon:** an operon containing multiple genes, which are transcribed as a single
590 mRNA from which the proteins are translated.

591 **Resolvase:** large family of site-specific recombinases. Resolvases play an essential role in
592 resolving plasmid multimers into monomers.

593 **mRNA interferase:** class of endoribonucleases, which cleave mRNA at a specific site, blocking
594 protein synthesis.

FIGURE LEGENDS

Figure 1. General features of laboratory and virulence plasmids. (A) Laboratory plasmids are generally smaller than 5 kb in size, and contain: i) replicon which can confer either high (≥50), medium (~20-40) or low (~1-10) copies number of the plasmid per cell; ii) a selectable marker for selection of bacteria containing the plasmid; iii) multiple cloning site (MCS), which contains several enzyme restriction sites to allow insertion of genes of interest; iv) a constitutive or inducible promoter, with its relative regulator ORF, that enables the expression of downstream genes. **(B)** A typical enteropathogenic virulence plasmid is larger than 40 kb and possesses: i) a low copy number replicon; ii) one or multiple TA systems, generally belonging to the type I or type II family; iii) different virulence genes that encode for different pathogenesis-related factors; iv) one or two partitioning systems.

Figure 2. The virulence plasmids of *Shigella flexneri* (pWR501) and *Salmonella enterica* Typhimurium (pSLT). The features on the plasmids follow the colour code in Figure 1. Plasmid maps were constructed on the sequencing data from Venkatesan *et al*. and McClendall *et al.* [28, 130].

Figure 3. The genetic structure of replicons in virulence plasmids. Plasmid replicons are mainly divided into **(A)** antisense RNA-regulated replicons and **(B**) iteron-regulated replicons. In both, binding of Rep to the origin (*ori*) is required to initiate DNA replication. In antisense RNA-regulated replicons **(A)**, plasmid copy number is controlled by an unstable antisense RNA, which prevents RepA translation at high concentrations by RNA interference. In RepFIIA, the most common antisense RNA-regulated system in virulence plasmids, the antisense RNA copA binds to *tap*, the untranslated leader of RepA mRNA which is essential for RepA translation [54]. CopB is a second element that also controls plasmid copy number by repressing

620 *repA* transcription, at its promoter PrepA. In iteron-regulated replicons **(B)**, the origin contains
621 iterons, which are directly repeated sequences that are specifically bound by Rep monomers.
622 A second set of iterons downstream *rep* controls plasmid copy number either by titrating Rep
623 monomers (the Titration model), or/and by forming Rep-DNA multimers that can interact
624 with Rep-*ori* complexes, blocking the initiation of replication (the Handcuffing Model) [18, 51, 57,
625 58]. RepFIA and RepFIB are two examples of iteron-regulated replicons in virulence plasmids,
626 which have iterons upstream and downstream of *rep* [59, 60].

627

628 **Figure 4. Partitioning systems in virulence plasmids: a schematic of type I and II par loci**
629 **and their different models of mechanism. (A)** In both type I and II par loci, the gene encoding
630 the NTPase component (blue) upstream of the centromere-binding protein (CBP, yellow),
631 while the centromere-like element (*par*, black) is either located downstream of the CBP (in
632 type Ia), upstream the NTPase (in type II), or at both positions (in type Ib). **(B)** Schematic
633 representation of the "pulling" and "diffusion ratchet" models for type I partitioning, and the
634 "pushing" model for type II partitioning. In the pulling model, NTPases-NTPs bind
635 cooperatively to nucleoid DNA, forming filaments. Contact between the partitioning complex
636 and the filaments induces the NTP hydrolysis, resulting in disassembly of polymers. Each
637 plasmid is pulled in the direction of NTPase depolymerisation [73]. In the diffusion ratchet
638 model, the NTPase-NTP binds non-specifically to nucleoid DNA and interacts with the
639 partitioning complex (NTPase-NTP link). This stimulates NTP hydrolysis and the NTPase-NDP
640 is then released from DNA, generating an area of NTPase-free nucleoid DNA around the site
641 of interaction with the plasmid; movement of the plasmid is driven by the slow re-association
642 of the NTPase and its fast de-association from nucleoid DNA [75]. In the "pushing" model, the

643 partitioning complex increases the stability of NTPase-NTP, stimulating the formation of
644 NTPase filaments that push the plasmid towards opposite poles [73]. Eventually, NTP hydrolysis
645 destabilises the filaments, and the polymers are disassembled; in contrast to the pulling
646 model, polymerisation and depolymerisation do not occur simultaneously [77].

647

648 **Figure 5. Cross-talk between par1 and par2 in pB171.** The two partitioning loci of pB171 (*par1*
649 and *par2*) share a central *par* region (*parC1*), which contains the promoters of both *parMR*
650 and *parAB* (P1 and P2, respectively), two R repeats (R1 and R2) and seventeen B repeats. ParR
651 and ParB homodimers bind to the R and B repeats, respectively. As both proteins bind co-
652 operatively to the repeated sequences, they form higher order nucleoprotein complexes that
653 repress the activity the P1 promoter only, in the case of ParR, and both the P1 and P2
654 promoters, in the case of ParB [79, 80].

655

656 **Figure 6. Summary of the principal elements that influence the biology of virulence**
657 **plasmids.** The contribution of each element to the biology of the plasmid is shown.

658

659 # TABLES

660 **Table 1. Comparison of prototypic plasmids with virulence plasmids from enteric bacteria**

661

662 **Table 2. Characteristics of TA systems in virulence plasmids.** [1] "Toxin target" refers to the

663 specific molecule with which the toxin of the TA system of interest interacts.

664

References

1. Stone, A.B. R factors: plasmids conferring resistance to antibacterial agents. *Sci Prog* **62**, 89-101 (1975).
2. Yang, F. et al. Genome dynamics and diversity of Shigella species, the etiologic agents of bacillary dysentery. *Nucleic Acids Res* **33**, 6445-58 (2005).
3. Kaper, J.B., Nataro, J.P. & Mobley, H.L. Pathogenic Escherichia coli. *Nat Rev Microbiol* **2**, 123-40 (2004).
4. Sansonetti, P.J., d'Hauteville, H., Ecobichon, C. & Pourcel, C. Molecular comparison of virulence plasmids in Shigella and enteroinvasive Escherichia coli. *Ann Microbiol (Paris)* **134A**, 295-318 (1983).
5. Lan, R., Lumb, B., Ryan, D. & Reeves, P.R. Molecular evolution of large virulence plasmid in Shigella clones and enteroinvasive Escherichia coli. *Infect Immun* **69**, 6303-9 (2001).
6. Clements, A., Young, J.C., Constantinou, N. & Frankel, G. Infection strategies of enteric pathogenic Escherichia coli. *Gut Microbes* **3**, 71-87 (2012).
7. Johnson, T.J. & Nolan, L.K. Pathogenomics of the virulence plasmids of Escherichia coli. *Microbiol Mol Biol Rev* **73**, 750-74 (2009).
8. Burland, V. et al. The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli O157:H7 *Nucleic Acids Research* **26** (1998).
9. Nguyen, Y. & Sperandio, V. Enterohemorrhagic E. coli (EHEC) pathogenesis. *Front Cell Infect Microbiol* **2**, 90 (2012).
10. Tobe, T. et al. Complete DNA sequence and structural analysis of the enteropathogenic Escherichia coli adherence factor plasmid. *Infect Immun* **67**, 5455-62 (1999).
11. Froehlich, B., Parkhill, J., Sanders, M., Quail, M.A. & Scott, J.R. The pCoo plasmid of enterotoxigenic Escherichia coli is a mosaic cointegrate. *J Bacteriol* **187**, 6509-16 (2005).
12. Jonsson, R. et al. A Novel pAA Virulence Plasmid Encoding Toxins and Two Distinct Variants of the Fimbriae of Enteroaggregative Escherichia coli. *Front Microbiol* **8**, 263 (2017).
13. Benjelloun-Touimi, Z., Si Tahar, M., Montecucco, C., Sansonetti, P.J. & Parsot, C. SepA, the 110 kDa protein secreted by Shigella flexneri: two-domain structure and proteolytic activity. *Microbiology* **144 (Pt 7)**, 1815-22 (1998).
14. Eslava, C. et al. Pet, an autotransporter enterotoxin from enteroaggregative Escherichia coli. *Infect Immun* **66**, 3155-63 (1998).
15. Eng, S. et al. Salmonella: A review on pathogenesis, epidemiology and antibiotic resistance. *Frontiers in Life Sciences* **8**, 284-293 (2015).
16. Rotger, R. & Casadesús, J. The virulence plasmids of Salmonella. *Int Microbiol* **2**, 177-84 (1999).
17. Mehigh, R.J., Sample, A.K. & Brubaker, R.R. Expression of the low calcium response in Yersinia pestis. *Microb Pathog* **6**, 203-17 (1989).
18. Funnell, E.B. & Phillips, J.G. Plasmid Biology (American Society for Microbiology Press, 2004).
19. Hu, P. et al. Structural organization of virulence-associated plasmids of Yersinia pestis. *J Bacteriol* **180**, 5192-202 (1998).
20. Nesbakken, T., Kapperud, G., Sorum, H. & Dommarsnes, K. Structural variability of 40-50 Mdal virulence plasmids from Yersinia enterocolitica. Geographical and ecological distribution of plasmid variants. *Acta Pathol Microbiol Immunol Scand B* **95**, 167-73 (1987).
21. Venkatesan, M.M. & Ranallo, R.T. Live-attenuated Shigella vaccines. *Expert Rev Vaccines* **5**, 669-86 (2006).
22. Visser, L.G., Annema, A. & van Furth, R. Role of Yops in inhibition of phagocytosis and killing of opsonized Yersinia enterocolitica by human granulocytes. *Infect Immun* **63**, 2570-5 (1995).
23. Pfeiffer, M.L., DuPont, H.L. & Ochoa, T.J. The patient presenting with acute dysentery--a systematic review. *J Infect* **64**, 374-86 (2012).

24. Mortality, G.B.D. & Causes of Death, C. Global, regional, and national life expectancy, all-cause mortality, and cause-specific mortality for 249 causes of death, 1980-2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet* **388**, 1459-1544 (2016).
25. Lima, I.F., Havt, A. & Lima, A.A. Update on molecular epidemiology of Shigella infection. *Curr Opin Gastroenterol* **31**, 30-7 (2015).
26. Kotloff, K.L. et al. Global burden of Shigella infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ* **77**, 651-66 (1999).
27. Sansonetti, P.J., Kopecko, D.J. & Formal, S.B. Involvement of a plasmid in the invasive ability of Shigella flexneri. *Infect Immun* **35**, 852-60 (1982).
28. Venkatesan, M.M. et al. Complete DNA sequence and analysis of the large virulence plasmid of Shigella flexneri. *Infect Immun* **69**, 3271-85 (2001).
29. 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**, 760-71. (2000).
30. Zychlinsky, A., Prevost, M.C. & Sansonetti, P.J. Shigella flexneri induces apoptosis in infected macrophages. *Nature* **358**, 167-9 (1992).
31. Ogawa, M., Handa, Y., Ashida, H., Suzuki, M. & Sasakawa, C. The versatility of Shigella effectors. *Nat Rev Microbiol* **6**, 11-6 (2008).
32. Suzuki, T. et al. Differential regulation of caspase-1 activation, pyroptosis, and autophagy via Ipaf and ASC in Shigella-infected macrophages. *PLoS Pathog* **3**, e111 (2007).
33. Schuch, R. & Maurelli, A.T. Virulence plasmid instability in Shigella flexneri 2a is induced by virulence gene expression. *Infect Immun* **65**, 3686-92 (1997).
34. Pilla, G., McVicker, G. & Tang, C.M. Genetic plasticity of the Shigella virulence plasmid is mediated by intra- and inter-molecular events between insertion sequences. *PLoS Genet* **13**, e1007014 (2017).
35. Brotcke Zumsteg, A., Goosmann, C., Brinkmann, V., Morona, R. & Zychlinsky, A. IcsA is a Shigella flexneri adhesin regulated by the type III secretion system and required for pathogenesis. *Cell Host Microbe* **15**, 435-45 (2014).
36. Schroeder, G.N. & Hilbi, H. Molecular pathogenesis of Shigella spp.: controlling host cell signaling, invasion, and death by type III secretion. *Clin Microbiol Rev* **21**, 134-56 (2008).
37. Gulig, P.A. & Doyle, T.J. The Salmonella typhimurium virulence plasmid increases the growth rate of salmonellae in mice. *Infect Immun* **61**, 504-11 (1993).
38. Fields, P.I., Swanson, R.V., Haidaris, C.G. & Heffron, F. Mutants of Salmonella typhimurium that cannot survive within the macrophage are avirulent. *Proc Natl Acad Sci U S A* **83**, 5189-93 (1986).
39. Gulig, P.A., Doyle, T.J., Clare-Salzler, M.J., Maiese, R.L. & Matsui, H. Systemic infection of mice by wild-type but not Spv- Salmonella typhimurium is enhanced by neutralization of gamma interferon and tumor necrosis factor alpha. *Infect Immun* **65**, 5191-7 (1997).
40. Zhu, Y. et al. Structural insights into the enzymatic mechanism of the pathogenic MAPK phosphothreonine lyase. *Mol Cell* **28**, 899-913 (2007).
41. Haneda, T. et al. Salmonella type III effector SpvC, a phosphothreonine lyase, contributes to reduction in inflammatory response during intestinal phase of infection. *Cell Microbiol* **14**, 485-99 (2012).
42. Boyd, E.F. & Hartl, D.L. Salmonella virulence plasmid. Modular acquisition of the spv virulence region by an F-plasmid in Salmonella enterica subspecies I and insertion into the chromosome of subspecies II, IIIa, IV and VII isolates. *Genetics* **149**, 1183-90 (1998).
43. Baumler, A.J. et al. The pef fimbrial operon of Salmonella typhimurium mediates adhesion to murine small intestine and is necessary for fluid accumulation in the infant mouse. *Infect Immun* **64**, 61-8 (1996).
44. Baumler, A.J., Tsolis, R.M. & Heffron, F. Contribution of fimbrial operons to attachment to and invasion of epithelial cell lines by Salmonella typhimurium. *Infect Immun* **64**, 1862-5 (1996).

45. Heffernan, E.J., Harwood, J., Fierer, J. & Guiney, D. The Salmonella typhimurium virulence plasmid complement resistance gene rck is homologous to a family of virulence-related outer membrane protein genes, including pagC and ail. *J Bacteriol* **174**, 84-91 (1992).
46. Heffernan, E.J. et al. Mechanism of resistance to complement-mediated killing of bacteria encoded by the Salmonella typhimurium virulence plasmid gene rck. *J Clin Invest* **90**, 953-64 (1992).
47. Rosselin, M. et al. Rck of Salmonella enterica, subspecies enterica serovar enteritidis, mediates zipper-like internalization. *Cell Res* **20**, 647-64 (2010).
48. Vandenbosch, J.L., Rabert, D.K., Kurlandsky, D.R. & Jones, G.W. 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* **57**, 850-7 (1989).
49. Sengupta, M. & Austin, S. Prevalence and significance of plasmid maintenance functions in the virulence plasmids of pathogenic bacteria. *Infect Immun* **79**, 2502-9 (2011).
50. Chattoraj, D.K. Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol Microbiol* **37**, 467-76 (2000).
51. Helinski, D.R., Toukdarian, A.E. & Novick, R. in Escherichia coli and Salmonella 2295-2324 (1996).
52. Pinto, U.M., Pappas, K.M. & Winans, S.C. The ABCs of plasmid replication and segregation. *Nat Rev Microbiol* **10**, 755-65 (2012).
53. del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M. & Diaz-Orejas, R. Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**, 434-64 (1998).
54. Nordstrom, K. Plasmid R1--replication and its control. *Plasmid* **55**, 1-26 (2006).
55. Cox, K.E.L. & Schildbach, J.F. Sequence of the R1 plasmid and comparison to F and R100. *Plasmid* **91**, 53-60 (2017).
56. Masai, H., Kaziro, Y. & Arai, K. Definition of oriR, the minimum DNA segment essential for initiation of R1 plasmid replication in vitro. *Proc Natl Acad Sci U S A* **80**, 6814-8 (1983).
57. Trawick, J.D. & Kline, B.C. A two-stage molecular model for control of mini-F replication. *Plasmid* **13**, 59-69 (1985).
58. Uga, H., Matsunaga, F. & Wada, C. Regulation of DNA replication by iterons: an interaction between the ori2 and incC regions mediated by RepE-bound iterons inhibits DNA replication of mini-F plasmid in Escherichia coli. *EMBO J* **18**, 3856-67 (1999).
59. Willetts, N. & Skurray, R. in Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology 1110-1133 (American Society for Microbiology, washington DC, 1987).
60. Saadi, S., Maas, W.K., Hill, D.F. & Bergquist, P.L. Nucleotide sequence analysis of RepFIC, a basic replicon present in IncFI plasmids P307 and F, and its relation to the RepA replicon of IncFII plasmids. *J Bacteriol* **169**, 1836-46 (1987).
61. Kawasaki, Y., Matsunaga, F., Kano, Y., Yura, T. & Wada, C. The localized melting of mini-F origin by the combined action of the mini-F initiator protein (RepE) and HU and DnaA of Escherichia coli. *Mol Gen Genet* **253**, 42-9 (1996).
62. Muraiso, K., Tokino, T., Murotsu, T. & Matsubara, K. Replication of mini-F plasmid in vitro promoted by purified E protein. *Mol Gen Genet* **206**, 519-21 (1987).
63. Ishiai, M., Wada, C., Kawasaki, Y. & Yura, T. Replication initiator protein RepE of mini-F plasmid: functional differentiation between monomers (initiator) and dimers (autogenous repressor). *Proc Natl Acad Sci U S A* **91**, 3839-43 (1994).
64. Manwaring, N.P., Skurray, R.A. & Firth, N. Nucleotide sequence of the F plasmid leading region. *Plasmid* **41**, 219-25 (1999).
65. Lilly, J. & Camps, M. Mechanisms of Theta Plasmid Replication. *Microbiol Spectr* **3**, PLAS-0029-2014 (2015).
66. Lan, R., Alles, M.C., Donohoe, K., Martinez, M.B. & Reeves, P.R. Molecular evolutionary relationships of enteroinvasive Escherichia coli and Shigella spp. *Infect Immun* **72**, 5080-8 (2004).

67. Tinge, S.A. & Curtiss, R. Conservation of Salmonella typhimurium virulence plasmid maintenance regions among Salmonella serovars as a basis for plasmid curing. *Infect Immun* **58**, 3084-92 (1990).
68. Rodríguez-Peña, J.M., Buisan, M., Ibáñez, M. & Rotger, R. Genetic map of the virulence plasmid of Salmonella enteritidis and nucleotide sequence of its replicons. *Gene* **188**, 53-61 (1997).
69. McNally, A., Thomson, N.R., Reuter, S. & Wren, B.W. 'Add, stir and reduce': Yersinia spp. as model bacteria for pathogen evolution. *Nat Rev Microbiol* **14**, 177-90 (2016).
70. Reuter, S. et al. Parallel independent evolution of pathogenicity within the genus Yersinia. *Proc Natl Acad Sci U S A* **111**, 6768-73 (2014).
71. Wang, H. et al. Increased plasmid copy number is essential for Yersinia T3SS function and virulence. *Science* **353**, 492-5 (2016).
72. Ebersbach Gitte & Gerdes, K. Plasmid Segregation Mechanisms. *Annual Review of Genetics* (2005).
73. Gerdes, K., Howard, M. & Szardenings, F. Pushing and pulling in prokaryotic DNA segregation. *Cell* **141**, 927-42 (2010).
74. Brooks, A.C. & Hwang, L.C. Reconstitutions of plasmid partition systems and their mechanisms. *Plasmid* **91**, 37-41 (2017).
75. Vecchiarelli, A.G., Mizuuchi, K. & Funnell, B.E. Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. *Mol Microbiol* **86**, 513-23 (2012).
76. Vecchiarelli, A.G., Seol, Y., Neuman, K.C. & Mizuuchi, K. A moving ParA gradient on the nucleoid directs subcellular cargo transport via a chemophoresis force. *Bioarchitecture* **4**, 154-9 (2014).
77. Møller-Jensen, J., Jensen, R.B., Löwe, J. & Gerdes, K. Prokaryotic DNA segregation by an actin-like filament. *EMBO J* **21**, 3119-27 (2002).
78. Jensen, R.B. & Gerdes, K. Mechanism of DNA segregation in prokaryotes: ParM partitioning protein of plasmid R1 co-localizes with its replicon during the cell cycle. *EMBO J* **18**, 4076-84 (1999).
79. Ebersbach, G. & Gerdes, K. The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. *Proc Natl Acad Sci U S A* **98**, 15078-83 (2001).
80. Ringgaard, S., Ebersbach, G., Borch, J. & Gerdes, K. Regulatory cross-talk in the double par locus of plasmid pB171. *J Biol Chem* **282**, 3134-45 (2007).
81. Cerin, H. & Hackett, J. The parVP region of the Salmonella typhimurium virulence plasmid pSLT contains four loci required for incompatibility and partition. *Plasmid* **30**, 30-8 (1993).
82. Chaudhuri, R.R. et al. Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative Escherichia coli strain 042. *PLoS One* **5**, e8801 (2010).
83. Brzozowska, I. & Zielenkiewicz, U. Regulation of toxin-antitoxin systems by proteolysis. *Plasmid* **70**, 33-41 (2013).
84. Gerdes, K. & Maisonneuve, E. Bacterial persistence and toxin-antitoxin loci. *Annu Rev Microbiol* **66**, 103-23 (2012).
85. Harms, A., Brodersen, D.E., Mitarai, N. & Gerdes, K. Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Mol Cell* (2018).
86. Dienemann, C., Bøggild, A., Winther, K.S., Gerdes, K. & Brodersen, D.E. Crystal structure of the VapBC toxin-antitoxin complex from Shigella flexneri reveals a hetero-octameric DNA-binding assembly. *J Mol Biol* **414**, 713-22 (2011).
87. Winther, K.S. & Gerdes, K. Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proc Natl Acad Sci U S A* **108**, 7403-7 (2011).
88. McVicker, G. & Tang, C.M. Deletion of toxin-antitoxin systems in the evolution of Shigella sonnei as a host-adapted pathogen. *Nat Microbiol* **2**, 16204 (2016).
89. Lobato-Márquez, D., Molina-García, L., Moreno-Córdoba, I., García-Del Portillo, F. & Díaz-Orejas, R. Stabilization of the Virulence Plasmid pSLT of Salmonella Typhimurium by Three

Maintenance Systems and Its Evaluation by Using a New Stability Test. *Front Mol Biosci* **3**, 66 (2016).
90. Lobato-Márquez, D., Moreno-Córdoba, I., Figueroa, V., Díaz-Orejas, R. & García-del Portillo, F. Distinct type I and type II toxin-antitoxin modules control Salmonella lifestyle inside eukaryotic cells. *Sci Rep* **5**, 9374 (2015).
91. Bernard, P. & Couturier, M. Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol* **226**, 735-45 (1992).
92. 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. *Mol Microbiol* **15**, 1031-7 (1995).
93. Di Cesare, A. et al. Diverse distribution of Toxin-Antitoxin II systems in Salmonella enterica serovars. *Sci Rep* **6**, 28759 (2016).
94. Christensen, S.K. & Gerdes, K. RelE toxins from bacteria and Archaea cleave mRNAs on translating ribosomes, which are rescued by tmRNA. *Mol Microbiol* **48**, 1389-400 (2003).
95. Pedersen, K. et al. The bacterial toxin RelE displays codon-specific cleavage of mRNAs in the ribosomal A site. *Cell* **112**, 131-40 (2003).
96. Gerdes, K. et al. Mechanism of postsegregational killing by the hok gene product of the parB system of plasmid R1 and its homology with the relF gene product of the E. coli relB operon. *EMBO J* **5**, 2023-9 (1986).
97. Thisted, T., Sorensen, N.S. & Gerdes, K. Mechanism of post-segregational killing: secondary structure analysis of the entire Hok mRNA from plasmid R1 suggests a fold-back structure that prevents translation and antisense RNA binding. *J Mol Biol* **247**, 859-73 (1995).
98. Thisted, T. & Gerdes, K. Mechanism of post-segregational killing by the hok/sok system of plasmid R1. Sok antisense RNA regulates hok gene expression indirectly through the overlapping mok gene. *J Mol Biol* **223**, 41-54 (1992).
99. Loh, S.M., Cram, D.S. & Skurray, R.A. Nucleotide sequence and transcriptional analysis of a third function (Flm) involved in F-plasmid maintenance. *Gene* **66**, 259-68 (1988).
100. Jurenas, D. et al. AtaT blocks translation initiation by N-acetylation of the initiator tRNAfMet. *Nat Chem Biol* **13**, 640-646 (2017).
101. Cheverton, A.M. et al. A Salmonella Toxin Promotes Persister Formation through Acetylation of tRNA. *Mol Cell* **63**, 86-96 (2016).
102. Jiang, Y., Pogliano, J., Helinski, D.R. & Konieczny, I. ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of Escherichia coli gyrase. *Mol Microbiol* **44**, 971-9 (2002).
103. Gerdes, K. Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. *J Bacteriol* **182**, 561-72 (2000).
104. Kasari, V., Mets, T., Tenson, T. & Kaldalu, N. Transcriptional cross-activation between toxin-antitoxin systems of Escherichia coli. *BMC Microbiol* **13**, 45 (2013).
105. Wessner, F. et al. Regulatory crosstalk between type I and type II toxin-antitoxin systems in the human pathogen Enterococcus faecalis. *RNA Biol* **12**, 1099-108 (2015).
106. Hyland, E.M., Wallace, E.W. & Murray, A.W. A model for the evolution of biological specificity: a cross-reacting DNA-binding protein causes plasmid incompatibility. *J Bacteriol* **196**, 3002-11 (2014).
107. Helaine, S. & Kugelberg, E. Bacterial persisters: formation, eradication, and experimental systems. *Trends Microbiol* **22**, 417-24 (2014).
108. Aviv, G. et al. A unique megaplasmid contributes to stress tolerance and pathogenicity of an emergent Salmonella enterica serovar Infantis strain. *Environ Microbiol* **16**, 977-94 (2014).
109. Chen, C.L., Su, L.H., Janapatla, R.P., Lin, C.Y. & Chiu, C.H. Genetic analysis of virulence and antimicrobial-resistant plasmid pOU7519 in Salmonella enterica serovar Choleraesuis. *J Microbiol Immunol Infect* (2017).
110. Barash, I. & Manulis-Sasson, S. Virulence mechanisms and host specificity of gall-forming Pantoea agglomerans. *Trends Microbiol* **15**, 538-45 (2007).

111. Koehler, T.M. Bacillus anthracis genetics and virulence gene regulation. *Curr Top Microbiol Immunol* **271**, 143-64 (2002).
112. Watson, B., Currier, T.C., Gordon, M.P., Chilton, M.D. & Nester, E.W. Plasmid required for virulence of Agrobacterium tumefaciens. *J Bacteriol* **123**, 255-64 (1975).
113. Franco, A. et al. Emergence of a Clonal Lineage of Multidrug-Resistant ESBL-Producing Salmonella Infantis Transmitted from Broilers and Broiler Meat to Humans in Italy between 2011 and 2014. *PLoS One* **10**, e0144802 (2015).
114. Aviv, G., Rahav, G. & Gal-Mor, O. Horizontal Transfer of the Salmonella enterica Serovar Infantis Resistance and Virulence Plasmid pESI to the Gut Microbiota of Warm-Blooded Hosts. *MBio* **7** (2016).
115. Hammerl, J.A., Freytag, B., Lanka, E., Appel, B. & Hertwig, S. The pYV virulence plasmids of Yersinia pseudotuberculosis and Y. pestis contain a conserved DNA region responsible for the mobilization by the self-transmissible plasmid pYE854. *Environ Microbiol Rep* **4**, 433-8 (2012).
116. Bikard, D. & Barrangou, R. Using CRISPR-Cas systems as antimicrobials. *Curr Opin Microbiol* **37**, 155-160 (2017).
117. Brenner, D.J. in The Prokaryotes (eds. Balows, A., Truper, H.G., Dworkin, M., Harder, W. & Shleifer, K.H.) 2673-2695 (Springer-Verlag, New York, 1992).
118. Eckburg, P.B. et al. Diversity of the human intestinal microbial flora. *Science* **308**, 1635-8 (2005).
119. Chow, J., Tang, H. & Mazmanian, S.K. Pathobionts of the gastrointestinal microbiota and inflammatory disease. *Curr Opin Immunol* **23**, 473-80 (2011).
120. Davin-Regli, A. & Pages, J.M. Enterobacter aerogenes and Enterobacter cloacae; versatile bacterial pathogens confronting antibiotic treatment. *Front Microbiol* **6**, 392 (2015).
121. Zechner, E.L. Inflammatory disease caused by intestinal pathobionts. *Curr Opin Microbiol* **35**, 64-69 (2017).
122. Round, J.L. & Mazmanian, S.K. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* **9**, 313-23 (2009).
123. Radosavljevic, V., Finke, E.J. & Belojevic, G. Escherichia coli O104:H4 outbreak in Germany--clarification of the origin of the epidemic. *Eur J Public Health* **25**, 125-9 (2015).
124. Pennington, T.H. E. coli O157 outbreaks in the United Kingdom: past, present, and future. *Infect Drug Resist* **7**, 211-22 (2014).
125. Heiman, K.E., Mody, R.K., Johnson, S.D., Griffin, P.M. & Gould, L.H. Escherichia coli O157 Outbreaks in the United States, 2003-2012. *Emerg Infect Dis* **21**, 1293-1301 (2015).
126. Yoder, J.S. et al. Outbreak of enterotoxigenic Escherichia coli infection with an unusually long duration of illness. *Clin Infect Dis* **42**, 1513-7 (2006).
127. Croxen, M.A. et al. Recent advances in understanding enteric pathogenic Escherichia coli. *Clin Microbiol Rev* **26**, 822-80 (2013).
128. 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**, 10567-72 (2000).
129. Lan, R. & Reeves, P.R. Escherichia coli in disguise: molecular origins of Shigella. *Microbes Infect* **4**, 1125-32 (2002).
130. McClelland, M. et al. Complete genome sequence of Salmonella enterica serovar Typhimurium LT2. *Nature* **413**, 852-6 (2001).