Atypical low-copy number plasmid segregation systems, all in one?

Patricia Siguier, Manuel Campos, François Cornet, Jean-Yves Bouet, Catherine Guynet

Atypical low-copy number plasmid segregation systems, all in one?

Patricia Siguier[1], Manuel Campos[1], François Cornet[1], Jean-Yves Bouet[1] and Catherine Guynet[1*]

[1] Laboratoire de Microbiologie et de Génétique Moléculaires, Centre de Biologie Intégrative (CBI), Centre National de la Recherche Scientifique, Université de Toulouse, UPS, F-31000, Toulouse, France.

* Correspondence: Catherine.Guynet@univ-tlse3.fr, +33 5 61 33 59 15

KEY WORDS: segregation; partition; StbA; Par; plasmid R388; plasmid pSK1; hitchhiking; HTH

ABSTRACT

Plasmid families harbor different maintenances functions, depending on their size and copy number. Low copy number plasmids rely on active partition systems, organizing a partition complex at specific centromere sites that is actively positioned using NTPase proteins. Some low copy number plasmids lack an active partition system, but carry atypical intracellular positioning systems using a single protein that binds to the centromere site but without an associated NTPase. These systems have been studied in the case of the *Escherichia coli* R388 and of the *Staphylococcus aureus* pSK1 plasmids. Here we review these two systems, which appear to be unrelated but share common features, such as their distribution on plasmids of medium size and copy number, certain activities of their centromere-binding proteins, StbA and Par, respectively, as well as their mode of action, which may involve dynamic interactions with the nucleoid-packed chromosome of their hosts.

Contents

1. Introduction

As non-essential extra-chromosomal DNA molecules, plasmids require special strategies for efficient replication and stable propagation in growing bacterial populations. These include multimer resolution, addiction and partition systems, all of which have been identified in both Gram-negative and Gram-positive bacteria. Some high copy number plasmids rely on a stochastic distribution of their copies and do not require partition systems to ensure that the fraction of cells without plasmids is low enough for their maintenance. In contrast, large low copy number plasmids often require all three types of dedicated systems to be stably inherited (Sengupta and Austin, 2011).

Most well-studied low copy-number plasmids, which are maintained at less than 5-6 copies per chromosome, encode one or more partition system. Briefly, these systems assemble a dedicated nucleoprotein complex around a centromere-like site, then separate and position them at specific subcellular positions, allowing each daughter cell to receive at least one copy after cell division. Partition systems described so far share common features: they include a *cis*-acting centromere-like site and contain two genes organized in tandem in an autoregulated *par* operon that encode two *trans*-acting proteins. In all cases, one gene encodes a centromere binding protein (CBP) forming a nucleoprotein complex at the centromere (i.e. the partition/segregation complex) (for reviews: (Baxter and Funnell, 2014; Bouet and Funnell, 2019; Gerdes et al., 2000). The other gene codes for an NTPase interacting with the partition complex and essential for segregation.

The family of NTPase defines to which type belongs a given partition system. To date, three main types have been identified in bacteria. Type I systems (or ParABS), of which the well-studied models are the P1 and F plasmids are the paradigm, encode Walker-type ATPase

proteins (ParA). They are by far the most prevalent in low-copy-number plasmids and are the only type present on chromosomes. They are divided into two sub-types, Ia and Ib, which differ by the nature of the DNA-binding domain of their CBP (ParB), the size of the Par proteins and the mode of transcriptional autoregulation of the *par* operon. In type Ib, the CBPs include a ribbon-helix-helix (RHH) domain and regulate transcription of the *par* operon (Golovanov et al., 2003; Hayes, 2000). In type Ia, the CBPs contain a helix-turn-helix (HTH) domain and *par* transcription is rather regulated by the associated NTPase (Friedman and Austin, 1988; Hirano et al., 1998). Recently, type I ParB proteins have been shown to belong to a new family of cytosine triphosphate (CTP) dependent molecular switches, which is required for the partition complex assembly (Jalal et al., 2021; Osorio-Valeriano et al., 2021; Soh et al., 2019). Type II systems (or ParRMC), identified in the *Escherichia coli* plasmid R1, encode actin-like ATPases (ParM) and CBPs containing a RHH DNA-binding domain that contributes to the operon transcriptional autoregulation (Gerdes et al., 2000; Jensen et al., 1994). Type III systems (or TubZRC), exemplified by plasmid pXO1 of *Bacillus anthracis*, encode tubulin-like GTPases and HTH-carrying CBPs (Larsen et al., 2007; Ni et al., 2010).

In types I-III systems, plasmid segregation strictly relies on the activity of their NTPases, which provides energy to separate plasmid copies and to position them to specific subcellular locations. Type II, which are the best characterized, and III systems use filamentation-based mechanisms relying on the ATP-dependent polymerization of the NTPase to drive partition complexes toward opposite cell poles, either via a pushing or a pulling mechanism, respectively (Figure 1) (Aylett et al., 2011; Aylett and Löwe, 2012; Møller-Jensen et al., 2003; Vecchiarelli et al., 2012). Although a mechanism involving similar dynamic NTPase filaments has been initially proposed for type I systems (Ebersbach et al., 2006; Ringgaard et al., 2009), recent studies have shown that ParA ATPases rather promotes segregation by using a

Brownian-ratchet mechanism (Hu et al., 2017; Hwang et al., 2013; Le Gall et al., 2016; Lim et al., 2014; Vecchiarelli et al., 2014, 2013, 2012). This mechanism relies on ParA ATP-dependent nonspecific DNA binding activity to the bacterial nucleoid. Briefly, the partition complex binds to the nucleoid via ParA, which stimulates ParA release from the DNA. The ParA rebinding to the nucleoid is slow, such that a void of ParA is created on the nucleoid around the partition complexes. This asymmetric redistribution of unbound ParA induces the partition complexes to move away from each other, towards the ParA remaining bound to the nucleoid (Figure 1).

Partition systems are not mutually exclusive, since some naturally-occurring plasmids carry two different systems, generally one of each type, such as plasmids R27 and pB171 (Ebersbach and Gerdes, 2001; Lawley and Taylor, 2003; Planchenault et al., 2020). Although both systems contribute to plasmid stability for these plasmids, the type I system makes the larger contribution. Besides, an number of sequenced low-copy-number plasmids do not encode typical partition systems (types I-III), suggesting the existence of alternative systems (Planchenault et al., 2020). In that line, two additional and distinct segregation systems have been highlighted. Although unrelated, they share the particularity of involving a single plasmid-encoded DNA-binding protein: a CBP not associated with an NTPase (Guynet et al., 2011; Simpson et al., 2003). These proteins, which have no homology to each other or to other known partitioning proteins, are the Par and the StbA proteins of the staphylococcal plasmid pSK1 and the *Escherichia coli* conjugative plasmid R388, respectively. Plasmids R388 and pSK1 are low-copy-number plasmids. R388 is maintained at about 5 copies per chromosome (Guynet et al., 2011), and, to our knowledge, the copy number of pSK1 has not been determined experimentally, but pSK1 minireplicons are also about 5 copies per chromosome (Grkovic et al., 2003). This range of copy number, which is theoretically not high enough for

faithful vertical transmission to rely on a stochastic distribution (Nordström and Austin, 1989), requires an active segregation process.

Here, we present an update on the distribution and organization of the StbA system. We also look back at recent biochemical, structural and *in vivo* studies and discuss hypotheses on possible models of segregation of R388 and pSK1, as experimental plasmids for these unusual segregation pathways involving a single plasmid-encoded protein.

2. Genetic organizations and regulation of *par* and *stbA loci*

pSK1 Par is encoded by the *par* gene, which is divergently transcribed from the plasmid's replication initiation gene *rep* (Figure 2A; (Firth et al., 2000)). The intergenic region between *rep* and *par* genes contains several features involved in plasmid replication and stability. These include the promoter of *rep*, the promoter of an antisense RNA, complementary to the leader region of *rep* (P*rnaI*), mediating negative regulation of pSK1 copy number, and the centromere-like site (Chan et al., 2022; Kwong et al., 2008). The origin of replication of pSK1 is located within the *rep* gene (Kwong et al., 2008; Liu et al., 2012). This organization of the *rep-par* intergenic region is conserved among staphylococcal multiresistance plasmids (Kwong et al., 2008, 2004). This configuration, in which the partition locus and other maintenance features are located near the *rep* locus, is widespread in plasmids. The centromere site, recognized by Par, is composed of seven repeats of a 12-bp DNA sequence with the consensus sequence TTAGGYRSYWAR (Y=C/T, R=A/G, S=G/C, W=A/T) containing a palindrome TTAG$(X)_4$CTAA ((Chan et al., 2022; Simpson et al., 2003); Figure 2A). The promoter of pSK1 *par* has been identified among other putative promoters between the *par* and *rep* genes. The P*par* – 35 sequence is located between repeats 3 and 4 of the centromere site, the -10 sequence

encompasses repeats 2 and 3, and the transcriptional start point is in repeat 2. As observed in many partition systems, Par autoregulates the expression of its own gene by binding the centromere repeats in the promoter region ((Chan et al., 2022), Figure 2A).

R388 StbA is encoded in the *stb* operon, which is divergently transcribed from the *trwABC* mobility operon (MOB) involved in conjugative DNA processing ((Fernández-López et al., 2006; Guynet et al., 2011); Figure 2B). The intergenic region between the two operons contains the origin of conjugative transfer (*oriT*) and *stbS*, the centromere site. *stbS* is composed of two sets of five repeats of a 9-bp DNA sequence (the *stbDR*) with the consensus (T/C)TGCATCAT separated by 2 bp (Figure 2B). The P*stbA* – 35 and -10 sequences are predicted to be in the stbDRs of the centromere site. StbA was shown to repress its own promoter, as well as four other promoters of plasmid R388. These latter contain two to five *stbDR* repeats, and makes StbA an important transcriptional regulator of the plasmid (Fernandez-Lopez et al., 2014; Quèbre et al., 2022). Except *ardC*, which encodes an antirestriction protein, necessary for conjugation in certain conditions (González-Montes et al., 2020), the other genes regulated by StbA have unknown functions. In addition to *stbA*, the *stb* operon contains two other genes, *stbB* and *stbC*, which are not involved in plasmid segregation in *E. coli* (Guynet et al., 2011). StbB is involved in the control of conjugation (see below), while StbC functions remain unknown.

3. Distribution of single-protein partitioning systems

The *par* locus was essentially found in Staphylococcal plasmids. These usually range in size from small (usually less than 10 kb) plasmids replicating by a rolling circle mechanism, to larger low-copy-number theta-replicating plasmids (generally 15 to 60 kb). These latter usually carry

antibiotic resistance determinants and are referred to as staphylococcal multiresistance plasmids, which are divided into three groups based on resistance phenotypes and conjugative properties: the pSK41-like conjugative plasmids, the nonconjugative antimicrobial and heavy-metal resistance plasmids, and the pSK1 family (Firth et al., 2000; Shearer et al., 2011). The *par* locus distribution has been investigated only in staphylococcal genomes so far. In 2011, a survey of plasmids of a large number of natural isolates of staphylococci reported 93 fully sequenced plasmids ranging from 1290 bp to 64909 bp, of which about 60 were > 10 kb (Shearer et al., 2011). All the 29 small plasmids of less than 10 kb appeared to lack any known partition system. In contrast, the vast majority of large plasmids harbored segregation functions. Of these, genes homologous to the pSK1 *par* locus were the most frequent, present in about 80% of plasmids of more than 10 kb (Figure 3). Interestingly, *par* was mostly present in medium-sized plasmids (20 to 37 kb). Largest plasmids (above 38 kb) carried either a type II (all pSK41-like conjugative plasmids), or a type Ib partitioning system.

Homologs of the other single-protein mediated segregation system, StbA, were found in about 14 % of the plasmids from a database containing 971 secondary replicons from Enterobacteria selected for their representativity based on the diversity of their replication and transfer machineries and the genus they belong to (Planchenault et al., 2020). We further searched for StbA homologs in all Enterobacterial plasmids present in RefSeq (Materials and methods). In agreement with our previous results, StbA was found in 17 % of plasmids (957 of 5820), the large majority (92 %) ranging from 20-kb to 150-kb (Figure 3). Thus, as Par homologs in Staphylococci, StbA is preferentially found in medium-sized plasmids in enterobacteria.

We next built a phylogenetic tree based on amino acid sequences of homologs of StbA, from previously identified alleles ((Guynet et al., 2011), Figure 4, Materials and methods). StbA

family members fall into four major groups. Groups 1, 2 and 4 seem to be restricted to ϒ- and β-proteobacteria and group 3 also includes cyanobacteria (Table S1).

Overall, these data indicate that the Par and StbA proteins are widespread, at least in Staphylococcal and Proteobacterial species, respectively. They also suggest that their function is well adapted to medium size plasmids (i.e., ranging from 20-kb to 40-kb and to 150-kb, respectively).

4. What do proteins tell us about their functions?

4.1. Binding to the centromere-like site

Recent data, mostly from structural and *in vitro* approaches, have provided information on how the partition complex assembles, which is the first step in the segregation process. The crystal structures of the pSK1 Par and the R388 StbA N-terminal DNA-binding domains (residues 1-53, (PDB ID: 8CSH) and residues 1-75 (PDB ID: 7PC1), referred to as $StbA_{1-75}$ throughout the text, respectively) have been solved. They display a winged-helix-turn-helix (wHTH) for pSK1 and a typical HTH for StbA, but without structural homology to other characterized CBPs or to each other (Chan et al., 2022; Quèbre et al., 2022).

Par is a 245-residues protein predicted to contain three domains: a N-terminal HTH domain, a C-terminal disordered domain, which frame a short disordered region and a central coiled coil (CC) region that might be involved in oligomerization of the protein (Chan et al., 2022), Figure 5A).The structure of Par 53 N-terminal residues bound to a 18-mer DNA site consists in two Par subunits. These show no contacts with each other, and the α-helix 2 inserts into the major groove of the duplex DNA, which consist of a half-site of a centromere repeat

(Figure 5B, (Chan et al., 2022)). Further studies indicate that Par is able to bind to at least one centromere consensus half sequence for high affinity binding *in vitro*. In addition, the organization of the centromere in seven contiguous repeats but separated by 1 to 19-bp or even overlapped suggests that one Par DNA binding entity in bound to each repeat. Notably, most contacts between Par and the DNA are hydrophobic and water mediated, without highly specific hydrogen bond-base interactions, which might indicate that Par is able to make contacts with DNA sites that deviate from the consensus (Chan et al., 2022).

R388 StbA is a small protein of 110 residues that was shown to harbor two domains, a pretty conserved N-terminal half composed of the HTH, and a non-conserved C-terminal half with a predicted disordered region (residues 69 to 108)((Quèbre et al., 2022) Figure 5A). The HTH domain is a typical HTH with three α helices, of which α3 is supposed to be the recognition helix responsible for the specific binding to the major groove of DNA. StbA N-terminal HTH domain ($StbA_{1-75}$) contains the DNA binding activity required for specific binding to the *stbDR* sequences and for plasmid segregation (Quèbre et al., 2022). *In vitro* experiments (EMSA) further strongly suggest that FL StbA, as well as $StbA_{1-75}$, binds to the *stbDR*s with high cooperativity resulting in the binding of two StbA HTH domains to every two *stbDRs*. Since no structure of StbA in the presence of DNA is yet available, we attempted to generated a model of $StbA_{1-75}$ bound to DNA (Figure 5C). StbA N-terminal domain HTH domain is structurally related (DALI search) to the wHTH domain of several members of the PadR family transcription regulators, except that it lacks the C-terminal wing. The wHTH domain of the protein Rv3488 of Mycobacterium tuberculosis H37Rv (PDB ID: 5ZHC) showed the strongest structural similarity to StbA (Quèbre et al., 2022). We thus used Rv3488 as a query structure for a Vast+ search in order to find protein structures, in complex with a DNA substrate, which have the closest similar 3D shape (Madej et al., 2020). We found a replication terminator protein of

Bacillus subtilis, arranged in a dimer, in complex with part of one of its DNA termination sites (PDB ID: 1F4K; Figure S1 (Vivian et al., 2007)), and used it to build a model of DNA-bound StbA (Figure 5C). With all the caution in the different interpretations that working with models requires, some interesting observations can be made. In the model, StbA assembles as a dimer and, as expected, the recognition α3-helix of each monomer inserts into the major groove of the DNA. Noteworthy, both monomers contact the DNA on the same side. This is consistent with the organization of *stbS* in two arrays of five 9-bp *stbDRs* spaced by 2-bp, thus corresponding to a complete helix turn. This is reminiscent of the ParR proteins (type II CBPs) from plasmids pB171 and pSK41. These cooperatively assemble into a continuous structure on their cognate centromeres, which are organized very similarly to *stbS* (Møller-Jensen et al., 2007; Schumacher et al., 2007). The putative oligomerization of the N-terminal domain of StbA as suggested by the model will be discussed below.

The potential roles of the other domains of both proteins on binding to the DNA have been investigated. For both proteins, Par and StbA, the C-terminal disordered region does not appear to be involved in centromere binding (Chan et al., 2022; Quèbre et al., 2022). However, in the case of Par, deletion of the central CC domain or mutation at a specific position predicted to interfere with dimer formation (see below) led to reduced binding to the centromere *in vitro*. This suggests that, although the CC domain does not directly contact DNA, oligomerization of Par has an impact on DNA binding affinity (Chan et al., 2022).

4.2. Oligomerization

Protein oligomerization is proved to be an important property of proteins involved in segregation functions, whether for the assembly of large nucleoprotein partition complexes or the process generating the driving force that promotes adequate plasmid positioning.

Par protein structural data indicate that there are no interactions between the two Par wHTH (Figure 5B), which is confirmed by size exclusion chromatography (SEC) (Chan et al., 2022). *In silico* predictions combined with SEC, circular dichroism and microscopy assays also show strong evidence converging on the presence of a CC domain in the central domain of Par, mediating the formation of Par dimer-of-dimers (Chan et al., 2022). Besides, wild type Par expressing GFP at its C-terminus forms distinct fluorescent foci in *S. aureus* cells. Foci of GFP-tagged Par carrying a mutation in the wHTH domain that abolishes DNA binding are more diffuse but visible, suggesting that Par multimerizes in the absence of DNA binding, although potential aggregation artifacts cannot be ruled out. A model of the full-length (FL) Par protein bound to the DNA was proposed by combining the crystal structure of the wHTH, AlphaFold 2 modeling of the central CC and the disordered domains, and docking the ends of the domain from each dimer (Chan et al., 2022). It shows that the CC domain mediates the formation of Par A dimers, which assemble in dimer-of-dimers through interactions between the C-terminal ends of the CC domains. Altogether, these data strongly suggest that the oligomerization mediated by the CC domain is important in Par segregation function.

Regarding StbA, a body of evidence points to an oligomerization of the protein, but the involvement of each domain, the N-terminal HTH and the C-terminal disordered domain, as well as the oligomerization level, are not clear. StbA, as most bacterial transcription factors, including the PadR proteins to which it is structurally most related, is likely to dimerize. In that

line, SEC analysis and bacterial two-hybrid assays indicate that FL StbA forms dimers (Quèbre et al., 2022). Crystallographic data of StbA N-terminal HTH domain ($StbA_{1-75}$) showed that the asymmetric unit contained one monomer, which correlates with SEC assays showing that it is a monomer in solution. This suggests that the N-terminal domain of StbA is not involved in oligomerization of the protein, yet bacterial two-hybrid assays show that $StbA_{1-75}$ interacts with itself. Besides, the fact that it exhibits specific and cooperative binding to the *stbDRs* properties similar to those of FL StbA suggests that interactions between $StbA_{1-75}$ monomers might be promoted or stabilized by binding to the DNA (Quèbre et al., 2022). In agreement with this, EMSA experiments indicate that neither StbA nor $StbA_{1-75}$ are able to form stable complexes with a single *stbDR* site, leading to the hypothesis that two StbA HTH domains bind to every two *stbDRs*. StbA and $StbA_{1-75}$ form specific high molecular weight complexes in the presence of DNA carrying *stbDRs in vitro*, suggesting that the proteins are capable of higher order oligomerization.

Structural and modeling data suggest roughly three different models that could explain how $StbA_{1-75}$ may form dimers. The first one is predicted from the nature of the crystals. Indeed, although the asymmetric unit (a.s.u.) contained a single $StbA_{1-75}$ monomer, the crystals showed a threefold symmetry between three identical subunits (Quèbre et al., 2022). The N-terminus of helix α1 of one monomer packed into a hydrophobic pocket formed with residues from helices the α1 and α2 helices of the second monomer (Figure 6A). The second model follows AlphaFold 2 predictions using $StbA_{1-75}$ as a query (Figure 6B). The putative dimerization interface is similar, but in contrast to the first model, the dimer assembles in a head-to-tail fashion. A loop located within helix α1 allows the N-terminus of helix α1 of the second monomer to contact helices α1 and α2 of the first monomer. The third model arises from our model based on structural homologies with a replication terminator protein (see

299 above, Figure 5C), in which monomers interact through their α1 helices. All three models
300 suggest that the long α1 helix has an important role for StbA oligomerization.

301 The contribution of StbA C-terminal domain in the oligomerization of the protein is
302 unknown. One might expect that, as with the characterized PadR family proteins to which
303 StbA is structurally related, it would mediate dimerization through interactions with the HTH
304 domain of the other monomer, but there is no evidence for this yet. Its disordered nature
305 probably correlates with the failure to crystallize FL StbA despite numerous attempts, as well
306 as the very low confidence in all predictions with AlphaFold 2, which does not allow to propose
307 a 3D structure model of the FL protein. Although the C-terminal domain is not required for the
308 formation of specific StbA-*stbDRs* complexes *in vitro*, it is clearly necessary for StbA activities
309 *in vivo*. Indeed, $StbA_{1-75}$ exhibits only partial activities in segregation and subcellular
310 positioning, as well as reduced activity in repressing the *stbDR*-carrying promoters of plasmid
311 R388. The C-terminus of StbA might thus stabilize interactions between StbA and the *stbDR*
312 sites, and/or promote interactions with other partners, as discussed below (Quèbre et al.,
313 2022).

314

315 ## 5. Segregation mechanism models for single-protein systems

316 Although the pSK1 Par and R388 StbA segregation systems are unrelated, both involve a
317 single plasmid-encoded CBP, and thus share the absence of the NTPase encoded together with
318 the CBP in all partition systems described so far. This raises the question, how a single DNA
319 binding protein can ensure both the assembly of the segregation complex and the addressing
320 of the plasmid to both daughter cells.

321

5.1. Fate of the plasmids when cohabiting with chromosomal DNA

pSK1 minireplicons and R388 plasmids have been localized in live cells, using a TetR-GFP/*tetO* fluorescent repressor-operator system in *S. aureus* (Chan et al., 2022; Lau et al., 2003), and a $ParB_{P1}$-GFP/*parS* system in *E. coli* (Guynet et al., 2011; Li and Austin, 2002; Quèbre et al., 2022), respectively. The presence of Par correlates with a weak mobility of pSK1 minireplicons foci, mostly two in number per cell, and which are confined in restricted areas and separate into two or more foci in dividing cells (Figure 7A). For plasmid R388, most cells carrying R388 exhibit between 4 and 6 foci of fluorescent-tagged plasmid (Guynet et al., 2011). This roughly corresponds to the copy number of R388, suggesting that most foci contain a single copy of the plasmid. Foci appear to be evenly assorted within nucleoid area (Figure 7B).

In both cases, deletion of the CBP leads to a decrease in the number, as well as aberrant localization or variations in the dynamics of fluorescent-tagged plasmids, which is associated with plasmid instability. In the absence of Par, most cells do not contain any foci. In foci-containing cells, foci are highly mobile and do not separate, hence the absence of faithful inheritance in the daughter cells upon division (Figure 7C). The absence of StbA, which does not affect plasmid copy-number, correlates with a significant decrease in the number of foci (between 1 and 3), which are clustered in nucleoid-free spaces, mostly at one cell pole (Figure 7D, (Guynet et al., 2011).

These observations are reminiscent of the non-uniform plasmid distribution in the cell driven by entropic forces that tend to physically separate plasmids from the chromosome, resulting in plasmid exclusion from nucleoid space. Entropy had also been proposed previously as the driving force behind the spontaneous unmixing of daughter chromatids leading to chromosome segregation in dividing bacteria (Jun and Wright, 2010). High copy number

plasmids (>15 copies), for which no segregation systems have been identified, have a hybrid distribution, composed of both large clusters and single random molecules. They are located mainly in the nucleoid-free area, at the poles and around the nucleoid periphery with occasional movements between poles, which is conducive for efficient random assortment to the daughter cells at cell division (Hsu and Chang, 2019; Reyes-Lamothe et al., 2014; Wang, 2017). Besides, experimental studies in *E. coli* as well as *in silico* simulations demonstrated that DNA circles devoid of partition system, resulting either from excised chromosomal circles or natural plasmids, are unmixed from the chromosome but maintained at the nucleoid-cytoplasm transition (Planchenault et al., 2020). This unmixing strongly depends of the replicon size. While plasmids below 25 kb diffuse rapidly across the chromosome, which would ensure efficient random segregation, largest plasmids suffer the highest missegregation and require the presence of a partitioning system for stability (Planchenault et al., 2020). In the same vein, when mixed in cavities provided by artificial nanofluidic model systems, plasmid molecules are found excluded from large DNA molecules (Liu et al., 2022). Polymer physics further predicts that shape anisotropy of the cell influences the organization of DNA molecules in bacterial compartments, such that unmixing would be significantly increased in rod-shaped bacteria compared to round-shaped bacteria (Jun and Mulder, 2006; Jun and Wright, 2010; Liu et al., 2022).

These data can be accounted for when considered in light of the distributions of pSK1(Δ*par*) and R388(Δ*stbA*) minireplicons. Both make foci that appear to be released from the nucleoid. Eviction appears stronger in rod-shaped *E. coli* cells for R388(Δ*stbA*) plasmids, which are found primarily at the cell poles and, as far as can be deducted from 2-dimension images, at the nucleoid edges, than in round-shaped *S. aureus* cells for pSK1(Δ*par*) minireplicons, which appear to form clusters and foci that are more mobile than in the

presence of Par (Figures 7C and 7D). The Par and StbA proteins can therefore be considered as sub-cellular positioning systems that act to localize plasmids molecules to the nucleoid, which could allow plasmids to counter entropy. A potential strategy for efficient segregation and consistent with these observations would be the tethering of plasmids to the bacterial nucleoid by physical interactions, which would be ensured by the activities of Par and StbA.

5.2. "Chromosome hitchhiking" or "pilot fish" model

This idea of a physical association of plasmids with the host chromosome for their stable propagation, is reminiscent of the partitioning systems of eukaryotic extra-chromosomal elements, exemplified by the 2-micron yeast plasmid and episomes of certain mammalian viruses (reviewed in (Sau et al., 2019)). This is the so-called 'hitchhiking' model, positing that plasmids utilize the chromosomes as vehicles for segregation. Although the physical association between the partition proteins of these elements and chromosomes has been established, the molecular basis of the interactions, as well as direct evidence of the hitchhiking mechanism, are still missing.

The hitchhiking model has also been proposed for both pSK1 and R388 plasmids (Chan et al., 2022; Guynet et al., 2011). It was also called the 'pilot-fish' model, where plasmid copies of R388 resemble pilot-fishes, who take advantage of the bow wave created by the swimming of big marine predators, which would be the chromosome in the case of plasmids (Guynet and de la Cruz, 2011). In this model, Par and StbA CBPs would bind to their respective centromeres and the resulting segregation complex would associate with the nucleoid. This could be done through direct interactions with chromosomal DNA sequences, or through indirect interactions involving one or more unknown chromosome-associated host factors.

Figure 7E(i) represents direct binding of Par and StbA CBPs to a discreet DNA site on the chromosome. According to the model of the pSK1 Par dimer-of-dimers bound to DNA, one Par dimer could bind to the centromere site, while the other dimer would interact with the host chromosome. Searches of the *S. aureus* genomes showed that it carries many candidate sequences for Par binding. Indeed, more than 20 sites contain 9 of the 12-bp of the centromere consensus repeat of pSK1, and there are likely many more sites, given that experimental data show that Par is able to bind to one centromere consensus half site, and probably to DNA sites that deviate from the consensus, as suggested by the absence of specific hydrogen-bond base interactions (Chan et al., 2022). How CBPs binding to multiple sites scattered around the chromosome could ensure faithful segregation of the plasmids to new daughter cells remains however unclear. The case of StbA appears to be different. Even if more than 350 matches to the 9-bp *stbDR* repeats sequences are found on the *E. coli* chromosome, the combination of two *stbDRs* in direct repetition separated by 2 bp, required to detect binding *in vitro* (Quèbre et al., 2022), is not found.

Another way to mediate plasmid attachment to the chromosome would be through the intervention of another protein. In this view, Par and StbA C-terminal domains might be involved in interactions with a host-encoded protein that would make the link between the segregation complex and the nucleoid (Figure 7E(ii)). No potential host partners have been identified so far by genome-wide assays: in *S. aureus* using the yeast two-hybrid, and in *E. coli* using the bacterial two-hybrid screen. Attempted to specifically reveal interactions between StbA and E. coli nucleoid-associated proteins (NAPs) by bacterial two-hybrid assays were also unsuccessful. A clue could have been given by the structural similarities between the Par DNA-binding domain with the wHTH of the *Bacillus subtilis* protein RacA. RacA plays a critical role in DNA segregation during sporulation by tethering the chromosome at the poles, presumably

by binding to centromere-like DNA sites near the origin of replication and then interacting with the pole-located DivIVA protein via its C-terminal domain. (Ben-Yehuda et al., 2005, 2003; Chan et al., 2022; Schumacher et al., 2016). However, yeast-two hybrid assays did not reveal any interaction between Par and the *S. aureus* DivIVA (Chan et al., 2022).

The above hitchhiking models rely on interactions between partition complexes and the host nucleoid, requiring a certain degree of specificity hardly compatible with the broad host range of some plasmids as R388 (Fernández-López et al., 2006). Alternatively, plasmid copies may position relative to each other to optimize their distribution. Such a process could rely on repulsion forces between partition complexes by an unknown mechanism. Plasmid copies would thus occupy the largest possible volume, while remaining bound to the nucleoid by non-specific interactions, ensuring adequate positioning upon cell division (Figure 7E(iii)).

5.3. StbA and the interplay between vertical segregation and horizontal transfer

If the segregation process strives to keep plasmid copies in the centre of the cell, this calls into question the ability of conjugative plasmids such as R388 to undertake conjugation, which takes place at the membrane. Conjugation is the process by which plasmids are transferred from a donor bacterium to a recipient bacterium through a conjugative pore (T4SS, type IV secretion system), which is established at the bacterial membrane between the two cells in physical contact (Cabezón et al., 2015; Llosa et al., 2002; Virolle et al., 2020). In this context, the *stb* operon the first evidence for a mechanistic interplay between segregation (vertical transfer) and conjugation (horizontal transfer). Indeed, StbA is not only the sole plasmid-encoded protein involved in R388 segregation, but is also an inhibitor of conjugation. StbA

439 inactivation makes R388 a super-spreader plasmid with up to 50-fold higher transfer
440 frequency than the wild type, whereas StbA presence makes the second protein encoded in
441 the opreron *stb*, StbB, strictly required for conjugative transfers (Guynet et al., 2011). Note
442 that as mentioned above, StbB, which contains a variant Walker A nucleotide triphosphate-
443 binding motif related to that found in the Soj/MinD superfamily of ATPases, including type I
444 par motor proteins, is not involved in R388 segregation. Strikingly, the activities of StbA and
445 StbB in controlling vertical and horizontal transfer of R388 correlate with subcellular
446 localization of plasmid copies: StbA-mediated segregation of R388 and inhibition of
447 conjugation correlates with the confinement of plasmids to nucleoid areas, whereas StbB
448 stimulation of conjugation correlates with the presence of plasmid copies in nucleoid-free
449 zones at the membrane T4SS. Plasmid R388 thus appears to encode in the same operon two
450 proteins that act to position plasmids. These counteract each other, and go against entropic
451 forces, which tend to position plasmids at the edge of the nucleoid, i.e. neither at the nucleoid
452 nor near the membrane, where StbA and StbB position plasmids, respectively. The StbAB
453 system can illustrate evolutionary parsimony applied to plasmid physiology, using only two
454 actors to ensure faithful plasmid propagation.

455

456 ## 6. Conclusion

457 Despite recent progress in characterizing the proteins involved in the two unrelated one-
458 protein plasmid segregation systems described so far, the mechanisms underlying the
459 assembly of their segregation complex, as well as how the plasmid copies separate and
460 strategically position themselves to ensure their stable maintenance in progeny are not yet
461 elucidated. No host protein that might play a role analogous to that of the NTPases of typical
462 partition systems, by an as yet unknown mechanism, has been identified. Polymer physics

studies could also potentially help in the understanding of the behavior of selfish elements and chromosomal DNA in the very crowded yet dynamic bacterial compartment. Although a hitchhiking mode of segregation is appealing for the pSK1 and R388 plasmid, no direct or indirect evidence for an association with the chromosome or any other host factor exists. The next step forward would be to identify potential plasmid tethering sites on the nucleoid and /or host factors if any. This mode of segregation would involve DNA-protein and protein-protein interactions that also have to be characterized. Whatever the molecular basis governing the one-protein segregation systems, it seems to be adapted to medium-sized plasmids. These plasmids may be small enough not to interfere with chromosome dynamics, either physically due to their size, or metabolically due to the number of genes they carry. They may thus keep a medium copy number (about 5 copies per chromosome in both cases), allowing stable inheritance to rely on an even subcellular assortment using the nucleoid as a scaffold, with no need of active transportation of plasmid copies before cell division.

7. Materials and methods

7.1. StbA reference dataset

The sequence of StbA alleles described in (Guynet et al., 2011) was used as a query in an iterative BLASTP search at NCBI (Genbank non-redundant database, May 2019). These alleles were those found in plasmids R388, R46, NAH7, pTF-FC2, and RP4. Since the full sequence of the latter was not available, the StbA allele of plasmid pBS228, which was identical, was used, and the StbA allele of plasmid pET49 was added following the iterative search. For each potential non-redundant StbA sequence, the corresponding plasmid sequence was downloaded and annotated using SnapGene software (www.snapgene.com), and oriTfinder

(https://tool-mml.sjtu.edu.cn/oriTfinder/oriTfinder.html) was used to determine possible *oriT* site, relaxases, and type IV secretion system genes. Hypothetical StbA sequences were retained if at least, either *stbB* or the MOB gene encoding the relaxase was present nearby.

7.2. Enterobacterial StbA dataset

An HMM profile was constructed using a MAFFT alignment of the conserved N-terminal domains of the StbA reference set. This HMM was then used as a query to search for StbA homologs in the full set of Enterobacteria plasmids (5820 plasmids) available in RefSeq (May 2019). We obtained 977 results for StbA in 957 plasmids (16 plasmids carried 2 StbA copies and 2 plasmids had 3 copies).

7.3. Phylogenetic tree

StbA sequences were analyzed via NGPhylogeny (https://ngphylogeny.fr/). They were aligned using MAFFT (Katoh and Standley, 2013) and fitted with TrimAl (Capella-Gutiérrez et al., 2009). The tree was built from the StbA reference dataset, to which we added 18 sequences (highlited in blue in Table S1) selected from previous results (Planchenault et al., 2020) for a better representation of all StbA sequences. A phylogenetic tree was constructed with PhyML (maximum-likelihood) (Guindon et al., 2010) (evolutionary model LG ; Tree topology search : SPR (Subtree Pruning and Regraphing)) and visualized with iTOL ((Letunic and Bork, 2021), https://itol.embl.de).

8. Acknowlegments

We thank all members of the Gedy team and to numerous colleagues for fruitful discussions and contributions to our understanding of the Stb system. This work was supported by French National Research Agency, grant number ANR-18- CE35-0008.

9. References

Aylett, C.H.S., Löwe, J., 2012. Superstructure of the centromeric complex of TubZRC plasmid partitioning systems. Proc Natl Acad Sci U S A 109, 16522–16527. https://doi.org/10.1073/pnas.1210899109

Aylett, C.H.S., Löwe, J., Amos, L.A., 2011. New insights into the mechanisms of cytomotive actin and tubulin filaments. Int Rev Cell Mol Biol 292, 1–71. https://doi.org/10.1016/B978-0-12-386033-0.00001-3

Baxter, J.C., Funnell, B.E., 2014. Plasmid Partition Mechanisms. Microbiol Spectr 2. https://doi.org/10.1128/microbiolspec.PLAS-0023-2014

Ben-Yehuda, S., Fujita, M., Liu, X.S., Gorbatyuk, B., Skoko, D., Yan, J., Marko, J.F., Liu, J.S., Eichenberger, P., Rudner, D.Z., Losick, R., 2005. Defining a centromere-like element in Bacillus subtilis by Identifying the binding sites for the chromosome-anchoring protein RacA. Mol Cell 17, 773–782. https://doi.org/10.1016/j.molcel.2005.02.023

Ben-Yehuda, S., Rudner, D.Z., Losick, R., 2003. RacA, a bacterial protein that anchors chromosomes to the cell poles. Science 299, 532–536. https://doi.org/10.1126/science.1079914

Bouet, J.-Y., Funnell, B.E., 2019. Plasmid Localization and Partition in Enterobacteriaceae. EcoSal Plus 8. https://doi.org/10.1128/ecosalplus.ESP-0003-2019

Cabezón, E., Ripoll-Rozada, J., Peña, A., de la Cruz, F., Arechaga, I., 2015. Towards an integrated model of bacterial conjugation. FEMS Microbiol Rev 39, 81–95. https://doi.org/10.1111/1574-6976.12085

Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chan, H.Y., Jensen, S.O., LeBard, R.J., Figgett, W.A., Lai, E., Simpson, A.E., Brzoska, A.J., Davies, D.S., Connolly, A.M., Cordwell, S.J., Travis, B.A., Salinas, R., Skurray, R.A., Firth, N., Schumacher, M.A., 2022. Molecular Analysis of pSK1 par: A Novel Plasmid Partitioning System Encoded by Staphylococcal Multiresistance Plasmids. J Mol Biol 434, 167770. https://doi.org/10.1016/j.jmb.2022.167770

Coppens, L., Lavigne, R., 2020. SAPPHIRE: a neural network based classifier for σ70 promoter prediction in Pseudomonas. BMC Bioinformatics 21, 415. https://doi.org/10.1186/s12859-020-03730-z

Ebersbach, G., Gerdes, K., 2001. The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. Proc Natl Acad Sci U S A 98, 15078–15083. https://doi.org/10.1073/pnas.261569598

Ebersbach, G., Ringgaard, S., Møller-Jensen, J., Wang, Q., Sherratt, D.J., Gerdes, K., 2006. Regular cellular distribution of plasmids by oscillating and filament-forming ParA ATPase of plasmid pB171. Mol Microbiol 61, 1428–1442. https://doi.org/10.1111/j.1365-2958.2006.05322.x

549 Fernandez-Lopez, R., Del Campo, I., Revilla, C., Cuevas, A., de la Cruz, F., 2014. Negative feedback and
550 transcriptional overshooting in a regulatory network for horizontal gene transfer. PLoS Genet
551 10, e1004171. https://doi.org/10.1371/journal.pgen.1004171
552 Fernández-López, R., Garcillán-Barcia, M.P., Revilla, C., Lázaro, M., Vielva, L., de la Cruz, F., 2006.
553 Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid
554 evolution. FEMS Microbiol Rev 30, 942–966. https://doi.org/10.1111/j.1574-
555 6976.2006.00042.x
556 Firth, N., Apisiridej, S., Berg, T., O'Rourke, B.A., Curnock, S., Dyke, K.G.H., Skurray, R.A., 2000.
557 Replication of Staphylococcal Multiresistance Plasmids. J Bacteriol 182, 2170–2178.
558 Friedman, S.A., Austin, S.J., 1988. The P1 plasmid-partition system synthesizes two essential proteins
559 from an autoregulated operon. Plasmid 19, 103–112. https://doi.org/10.1016/0147-
560 619x(88)90049-2
561 Gerdes, K., Møller-Jensen, J., Bugge Jensen, R., 2000. Plasmid and chromosome partitioning:
562 surprises from phylogeny. Mol Microbiol 37, 455–466. https://doi.org/10.1046/j.1365-
563 2958.2000.01975.x
564 Golovanov, A.P., Barillà, D., Golovanova, M., Hayes, F., Lian, L.-Y., 2003. ParG, a protein required for
565 active partition of bacterial plasmids, has a dimeric ribbon-helix-helix structure. Mol
566 Microbiol 50, 1141–1153. https://doi.org/10.1046/j.1365-2958.2003.03750.x
567 Grkovic, S., Brown, M.H., Hardie, K.M., Firth, N., Skurray, R.A., 2003. Stable low-copy-number
568 Staphylococcus aureus shuttle vectors. Microbiology (Reading) 149, 785–794.
569 https://doi.org/10.1099/mic.0.25951-0
570 Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms
571 and methods to estimate maximum-likelihood phylogenies: assessing the performance of
572 PhyML 3.0. Syst Biol 59, 307–321. https://doi.org/10.1093/sysbio/syq010
573 Guynet, C., Cuevas, A., Moncalián, G., de la Cruz, F., 2011. The stb operon balances the requirements
574 for vegetative stability and conjugative transfer of plasmid R388. PLoS Genet 7, e1002073.
575 https://doi.org/10.1371/journal.pgen.1002073
576 Guynet, C., de la Cruz, F., 2011. Plasmid segregation without partition. Mob Genet Elements 1, 236–
577 241. https://doi.org/10.4161/mge.1.3.18229
578 Hayes, F., 2000. The partition system of multidrug resistance plasmid TP228 includes a novel protein
579 that epitomizes an evolutionarily distinct subgroup of the ParA superfamily. Mol Microbiol
580 37, 528–541. https://doi.org/10.1046/j.1365-2958.2000.02030.x
581 Hirano, M., Mori, H., Onogi, T., Yamazoe, M., Niki, H., Ogura, T., Hiraga, S., 1998. Autoregulation of
582 the partition genes of the mini-F plasmid and the intracellular localization of their products in
583 Escherichia coli. Mol Gen Genet 257, 392–403. https://doi.org/10.1007/s004380050663
584 Hsu, T.-M., Chang, Y.-R., 2019. High-Copy-Number Plasmid Segregation-Single-Molecule Dynamics in
585 Single Cells. Biophys J 116, 772–780. https://doi.org/10.1016/j.bpj.2019.01.019
586 Hu, L., Vecchiarelli, A.G., Mizuuchi, K., Neuman, K.C., Liu, J., 2017. Brownian ratchet mechanisms of
587 ParA-mediated partitioning. Plasmid 92, 12–16.
588 https://doi.org/10.1016/j.plasmid.2017.05.002
589 Hwang, L.C., Vecchiarelli, A.G., Han, Y.-W., Mizuuchi, M., Harada, Y., Funnell, B.E., Mizuuchi, K., 2013.
590 ParA-mediated plasmid partition driven by protein pattern self-organization. EMBO J 32,
591 1238–1249. https://doi.org/10.1038/emboj.2013.34
592 Jalal, A.S., Tran, N.T., Stevenson, C.E., Chimthanawala, A., Badrinarayanan, A., Lawson, D.M., Le, T.B.,
593 2021. A CTP-dependent gating mechanism enables ParB spreading on DNA. Elife 10, e69676.
594 https://doi.org/10.7554/eLife.69676
595 Jensen, R.B., Dam, M., Gerdes, K., 1994. Partitioning of plasmid R1. The parA operon is autoregulated
596 by ParR and its transcription is highly stimulated by a downstream activating element. J Mol
597 Biol 236, 1299–1309. https://doi.org/10.1016/0022-2836(94)90059-0
598 Jun, S., Mulder, B., 2006. Entropy-driven spatial organization of highly confined polymers: lessons for
599 the bacterial chromosome. Proc Natl Acad Sci U S A 103, 12388–12393.
600 https://doi.org/10.1073/pnas.0605305103

601 Jun, S., Wright, A., 2010. Entropy as the driver of chromosome segregation. Nat Rev Microbiol 8,
602 600–607. https://doi.org/10.1038/nrmicro2391
603 Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7:
604 Improvements in Performance and Usability. Mol Biol Evol. 30, 772–780.
605 Kwong, S.M., Lim, R., LeBard, R.J., Skurray, R.A., Firth, N., 2008. Analysis of the pSK1 replicon, a
606 prototype from the staphylococcal multiresistance plasmid family. Microbiology (Reading)
607 154, 3084–3094. https://doi.org/10.1099/mic.0.2008/017418-0
608 Kwong, S.M., Skurray, R.A., Firth, N., 2004. Staphylococcus aureus multiresistance plasmid pSK41:
609 analysis of the replication region, initiator protein binding and antisense RNA regulation. Mol
610 Microbiol 51, 497–509. https://doi.org/10.1046/j.1365-2958.2003.03843.x
611 Larsen, R.A., Cusumano, C., Fujioka, A., Lim-Fong, G., Patterson, P., Pogliano, J., 2007. Treadmilling of
612 a prokaryotic tubulin-like protein, TubZ, required for plasmid stability in Bacillus
613 thuringiensis. Genes Dev 21, 1340–1352. https://doi.org/10.1101/gad.1546107
614 Lau, I.F., Filipe, S.R., Søballe, B., Økstad, O.-A., Barre, F.-X., Sherratt, D.J., 2003. Spatial and temporal
615 organization of replicating Escherichia coli chromosomes. Mol Microbiol 49, 731–743.
616 https://doi.org/10.1046/j.1365-2958.2003.03640.x
617 Lawley, T.D., Taylor, D.E., 2003. Characterization of the double-partitioning modules of R27:
618 correlating plasmid stability with plasmid localization. J Bacteriol 185, 3060–3067.
619 https://doi.org/10.1128/JB.185.10.3060-3067.2003
620 Le Gall, A., Cattoni, D.I., Guilhas, B., Mathieu-Demazière, C., Oudjedi, L., Fiche, J.-B., Rech, J.,
621 Abrahamsson, S., Murray, H., Bouet, J.-Y., Nollmann, M., 2016. Bacterial partition complexes
622 segregate within the volume of the nucleoid. Nat Commun 7, 12107.
623 https://doi.org/10.1038/ncomms12107
624 Letunic, I., Bork, P., 2021. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
625 display and annotation. Nucleic Acids Res 49, W293–W296.
626 https://doi.org/10.1093/nar/gkab301
627 Li, Y., Austin, S., 2002. The P1 plasmid in action: time-lapse photomicroscopy reveals some
628 unexpected aspects of plasmid partition. Plasmid 48, 174–178.
629 https://doi.org/10.1016/s0147-619x(02)00104-x
630 Lim, H.C., Surovtsev, I.V., Beltran, B.G., Huang, F., Bewersdorf, J., Jacobs-Wagner, C., 2014. Evidence
631 for a DNA-relay mechanism in ParABS-mediated chromosome segregation. Elife 3, e02758.
632 https://doi.org/10.7554/eLife.02758
633 Liu, M.A., Kwong, S.M., Pon, C.K., Skurray, R.A., Firth, N., 2012. Genetic requirements for replication
634 initiation of the staphylococcal multiresistance plasmid pSK41. Microbiology (Reading) 158,
635 1456–1467. https://doi.org/10.1099/mic.0.057620-0
636 Liu, Z., Capaldi, X., Zeng, L., Zhang, Y., Reyes-Lamothe, R., Reisner, W., 2022. Confinement anisotropy
637 drives polar organization of two DNA molecules interacting in a nanoscale cavity. Nat
638 Commun 13, 4358. https://doi.org/10.1038/s41467-022-31398-x
639 Llosa, M., Gomis-Rüth, F.X., Coll, M., de la Cruz Fd, F., 2002. Bacterial conjugation: a two-step
640 mechanism for DNA transport. Mol Microbiol 45, 1–8. https://doi.org/10.1046/j.1365-
641 2958.2002.03014.x
642 Madej, T., Marchler-Bauer, A., Lanczycki, C., Zhang, D., Bryant, S.H., 2020. Biological Assembly
643 Comparison with VAST. Methods Mol Biol 2112, 175–186. https://doi.org/10.1007/978-1-
644 0716-0270-6_13
645 Møller-Jensen, J., Borch, J., Dam, M., Jensen, R.B., Roepstorff, P., Gerdes, K., 2003. Bacterial mitosis:
646 ParM of plasmid R1 moves plasmid DNA by an actin-like insertional polymerization
647 mechanism. Mol Cell 12, 1477–1487. https://doi.org/10.1016/s1097-2765(03)00451-9
648 Møller-Jensen, J., Ringgaard, S., Mercogliano, C.P., Gerdes, K., Löwe, J., 2007. Structural analysis of
649 the ParR/parC plasmid partition complex. EMBO J 26, 4413–4422.
650 https://doi.org/10.1038/sj.emboj.7601864
651 Ni, L., Xu, W., Kumaraswami, M., Schumacher, M.A., 2010. Plasmid protein TubR uses a distinct mode
652 of HTH-DNA binding and recruits the prokaryotic tubulin homolog TubZ to effect DNA

653 partition. Proc Natl Acad Sci U S A 107, 11763–11768.
654 https://doi.org/10.1073/pnas.1003817107
655 Nordström, K., Austin, S.J., 1989. Mechanisms that contribute to the stable segregation of plasmids.
656 Annu Rev Genet 23, 37–69. https://doi.org/10.1146/annurev.ge.23.120189.000345
657 Osorio-Valeriano, M., Altegoer, F., Das, C.K., Steinchen, W., Panis, G., Connolley, L., Giacomelli, G.,
658 Feddersen, H., Corrales-Guerrero, L., Giammarinaro, P.I., Hanßmann, J., Bramkamp, M.,
659 Viollier, P.H., Murray, S., Schäfer, L.V., Bange, G., Thanbichler, M., 2021. The CTPase activity
660 of ParB determines the size and dynamics of prokaryotic DNA partition complexes. Mol Cell
661 81, 3992-4007.e10. https://doi.org/10.1016/j.molcel.2021.09.004
662 Planchenault, C., Pons, M.C., Schiavon, C., Siguier, P., Rech, J., Guynet, C., Dauverd-Girault, J., Cury, J.,
663 Rocha, E.P.C., Junier, I., Cornet, F., Espéli, O., 2020. Intracellular Positioning Systems Limit the
664 Entropic Eviction of Secondary Replicons Toward the Nucleoid Edges in Bacterial Cells. J Mol
665 Biol 432, 745–761. https://doi.org/10.1016/j.jmb.2019.11.027
666 Quèbre, V., Del Campo, I., Cuevas, A., Siguier, P., Rech, J., Le, P.T.N., Ton-Hoang, B., Cornet, F., Bouet,
667 J.-Y., Moncalian, G., de la Cruz, F., Guynet, C., 2022. Characterization of the DNA Binding
668 Domain of StbA, A Key Protein of A New Type of DNA Segregation System. J Mol Biol 434,
669 167752. https://doi.org/10.1016/j.jmb.2022.167752
670 Reyes-Lamothe, R., Tran, T., Meas, D., Lee, L., Li, A.M., Sherratt, D.J., Tolmasky, M.E., 2014. High-copy
671 bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are
672 randomly partitioned at cell division. Nucleic Acids Res 42, 1042–1051.
673 https://doi.org/10.1093/nar/gkt918
674 Ringgaard, S., van Zon, J., Howard, M., Gerdes, K., 2009. Movement and equipositioning of plasmids
675 by ParA filament disassembly. Proc Natl Acad Sci U S A 106, 19369–19374.
676 https://doi.org/10.1073/pnas.0908347106
677 Sau, S., Ghosh, S.K., Liu, Y.-T., Ma, C.-H., Jayaram, M., 2019. Hitchhiking on chromosomes: A
678 persistence strategy shared by diverse selfish DNA elements. Plasmid 102, 19–28.
679 https://doi.org/10.1016/j.plasmid.2019.01.004
680 Schumacher, M.A., Glover, T.C., Brzoska, A.J., Jensen, S.O., Dunham, T.D., Skurray, R.A., Firth, N.,
681 2007. Segrosome structure revealed by a complex of ParR with centromere DNA. Nature 450,
682 1268–1271. https://doi.org/10.1038/nature06392
683 Schumacher, M.A., Lee, J., Zeng, W., 2016. Molecular insights into DNA binding and anchoring by the
684 Bacillus subtilis sporulation kinetochore-like RacA protein. Nucleic Acids Res 44, 5438–5449.
685 https://doi.org/10.1093/nar/gkw248
686 Sengupta, M., Austin, S., 2011. Prevalence and significance of plasmid maintenance functions in the
687 virulence plasmids of pathogenic bacteria. Infect Immun 79, 2502–2509.
688 https://doi.org/10.1128/IAI.00127-11
689 Shearer, J.E.S., Wireman, J., Hostetler, J., Forberger, H., Borman, J., Gill, J., Sanchez, S., Mankin, A.,
690 Lamarre, J., Lindsay, J.A., Bayles, K., Nicholson, A., O'Brien, F., Jensen, S.O., Firth, N., Skurray,
691 R.A., Summers, A.O., 2011. Major families of multiresistant plasmids from geographically and
692 epidemiologically diverse staphylococci. G3 (Bethesda) 1, 581–591.
693 https://doi.org/10.1534/g3.111.000760
694 Simpson, A.E., Skurray, R.A., Firth, N., 2003. A single gene on the staphylococcal multiresistance
695 plasmid pSK1 encodes a novel partitioning system. J Bacteriol 185, 2143–2152.
696 https://doi.org/10.1128/jb.185.7.2143-2152.2003
697 Soh, Y.-M., Davidson, I.F., Zamuner, S., Basquin, J., Bock, F.P., Taschner, M., Veening, J.-W., De Los
698 Rios, P., Peters, J.-M., Gruber, S., 2019. Self-organization of parS centromeres by the ParB
699 CTP hydrolase. Science 366, 1129–1133. https://doi.org/10.1126/science.aay3965
700 Vecchiarelli, A.G., Hwang, L.C., Mizuuchi, K., 2013. Cell-free study of F plasmid partition provides
701 evidence for cargo transport by a diffusion-ratchet mechanism. Proc Natl Acad Sci U S A 110,
702 E1390-1397. https://doi.org/10.1073/pnas.1302745110

Vecchiarelli, A.G., Mizuuchi, K., Funnell, B.E., 2012. Surfing biological surfaces: exploiting the nucleoid for partition and transport in bacteria. Mol Microbiol 86, 513–523. https://doi.org/10.1111/mmi.12017
Vecchiarelli, A.G., Neuman, K.C., Mizuuchi, K., 2014. A propagating ATPase gradient drives transport of surface-confined cellular cargo. Proc Natl Acad Sci U S A 111, 4880–4885. https://doi.org/10.1073/pnas.1401025111
Virolle, C., Goldlust, K., Djermoun, S., Bigot, S., Lesterlin, C., 2020. Plasmid Transfer by Conjugation in Gram-Negative Bacteria: From the Cellular to the Community Level. Genes (Basel) 11, E1239. https://doi.org/10.3390/genes11111239
Vivian, J.P., Porter, C.J., Wilce, J.A., Wilce, M.C.J., 2007. An asymmetric structure of the Bacillus subtilis replication terminator protein in complex with DNA. J Mol Biol 370, 481–491. https://doi.org/10.1016/j.jmb.2007.02.067
Wang, Y., 2017. Spatial distribution of high copy number plasmids in bacteria. Plasmid, SI: ISPB Plasmid 2016 91, 2–8. https://doi.org/10.1016/j.plasmid.2017.02.005

Figure legends

Figure 1. Mechanisms of typical partition systems. As indicated, centromere binding proteins are represented as yellow circles, and NTPases as ovals (ATP-bound, light blue; ADP-bound, dark blue). **(A)** Type I partition systems. The Brownian ratchet mechanism relies on the ParA ATPase, which binds to the nucleoid in a nonspecific and ATP-dependent manner to DNA. ParB/*parS* partition complexes associate with the nucleoid via ParA-ATP, and then stimulate the release of ParA from the DNA by ATP hydrolysis or conformational change. Due to the slow re-binding of ParA to the nucleoid, a void of ParA is created and serves as a barrier so that the partition complexes move toward opposite directions by following the ParA remaining bound to the nucleoid. **(B)** Type II partition systems. The R1 paradigm plasmid uses ATP-dependent polymerization of the actin-like ParM ATPase to push plasmids poleward. ParR/*parC* partition complexes bind to the terminal ParM-ATP subunits at the growing end of the filament. Hydrolysis of ATP to ADP leads to destabilization of the filaments, allowing entry of the ParM-ATP subunits. The filaments are polar and associate antiparallel, so that the plasmids are pushed in a bidirectional manner. **(C)** Type III partition systems. The pulling mechanism, exemplified by pXO1, involves polymerization of the tubulin-like TubZ GTPase, which forms

polar and dynamic filaments in a treadmilling-like pattern. These grow at the plus end by addition of TubZ-GTP and disassemble at the minus end, from which the TubR/*tubC* partition complexes are pulled to the pole.

Figure 2. Genetic organization of *par* (A) and *stb* (B) regions. Promoters are indicated by black bent arrows, dotted black the putative promoter P*stb* (predicted with SAPPHIRE, (Coppens and Lavigne, 2020)). Gray boxes represent −10 and −35 sequences of P*par* and the putative P*stb*. As indicated, Par and StbA repress their own promoter. Direct repeats of the centromere-like sites are represented by orange and yellow (*stbDRs*) solid arrows, respectively. The origin of conjugative transfer of plasmid R388 is shown as a vertical arrow. The scale is not respected.

Figure 3. Distribution of Par (A) and StbA (B) segregation systems. The event plots represent the presence of Par or StbA (blue vertical lines) in the 92 staphylococcal or in the 5820 enterobacterial replicons (red vertical lines), respectively, and which were ranked according to their size (x-axis). The data for staphylococcal plasmids are from (Shearer et al., 2011).

Figure 4. Phylogenetic tree of the StbA representative sequences. Accession numbers of plasmids are indicated and the reference plasmids are in large bold type. Based on the tree topology, the sequences could be separated in four groups. As indicated: Group 1 (red) includes plasmids R46 (AY046276), pNAH7 (AB237655), R721 (NC_002525) and R388 (NC_028464) – Group 2 (green) includes pBS228 (NC_008357) – Group 3 (blue) include

plasmid pTF-FC2 (TFETFFC2) – Group 4 (orange) include plasmid pET49 (CU468131). The tree display was obtained with online iTOL ((Letunic and Bork, 2021), https://itol.embl.de).

Figure 5. Par and StbA DNA binding domains. **(A)** Schematic representing Par and StbA structural domains. **(B)** Structure of Par DNA-binding domains bound to DNA. The two subunits do not make contact, and each binds to a half site of the centromere. α-helices and β-strands of the winged-HTH domain are labelled (PDB ID: 8CSH, (Chan et al., 2022)). **(C)** Model of StbA DNA-binding domain bound to DNA. The StbA N-terminal domain (PDB ID: 7PC1) and the replication terminator protein of *Bacillus subtilis* bound to DNA (PDB ID: 1F4K) structures were superimposed. In the resulting model, StbA assembles as a head-to-tail dimer, and the recognition helix α3 of each monomer inserts into the major groove of the DNA. α-helices of the HTH domain are indicated. The display of structures and analyses were performed with UCSF Chimera (developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311).

Figure 6. Models of StbA DNA-binding domain dimerization. **(A)** The structure of the DNA-binding domain ($StbA_{1-75}$) is from our crystal structure. The crystals showed a threefold symmetry between three identical units, revealing possible interactions between monomers. One monomer interacts with helix α3 and the C-terminal part of α1 of another monomer via the N-terminal part of its helix α1. The third monomer is not show for clarity. **(B)** The structure of the DNA-binding domain is from modeling with AlphaFold 2. The arrangement of the two monomers is head-to-tail, so that the N-terminal part of helix α1 of one monomer is packed between helix α1 and α2 of the other monomer and vice versa.

Figure 7. Models for plasmid segregation mediated by single-protein systems. Symbols are indicated in the legend of the figure (left panel). **(A)-(D)** Schematics showing the sub-cellular positioning of pSK1 minireplicons and R388 plasmids in *S. aureus* and *E. coli* cells, respectively. pSK1 minireplicons are confined in restricted areas in the presence of Par and separate into two or more foci in dividing cells **(A)**, whereas in the absence of Par they are highly mobile (shown as a grey dotted line) and do not separate **(C)**. Plasmids R388 are evenly distributed in the nucleoid area in the presence of StbA **(B)**, whereas they are clustered and excluded from the nucleoid in the absence of StbA **(D)**. **(E)** Proposed models for positioning and segregation mediated with a single-protein system. **(i)** and **(ii)** represent a hitchhiking mechanism, in which plasmid molecules are attached on the bacterial nucleoid either through direct interactions between the segregation protein and the chromosomal DNA **(i)**, or through interactions with one or more host proteins that bind to the nucleoid **(ii)**. The plasmids thus take advantage of the segregation of the chromosomes to distribute themselves between the two daughter cells during bacterial division. In **(iii)**, the partition complexes would interact with each other rather than with the nucleoid to partition plasmids into the nucleoid space and ensure daughter cells to receive at least one copy of the plasmid.

Supplementary material

Figure S1. (A) Structure of the replication terminator protein of *Bacillus subtilis* bound to DNA (PDB ID: 1F4K). **(B)** Structural superposition of the replication terminator protein of *B. subtilis* (green) and $StbA_{1-75}$ (crimson and pink) dimers. Part of the α-helix located C-terminal to the winge is not shown for clarity.

802 **Table S1.** List of plasmids used to build the phylogenetic tree. For each, the accession number,
803 the name, the size, the host and the group of StbA proteins (group 1 in red, group 2 in green,
804 group 3 in blue and group 4 in orange) are indicated. The reference plasmids are in bold type,
805 and the 18 additional sequences are highlighted in blue.

806

807

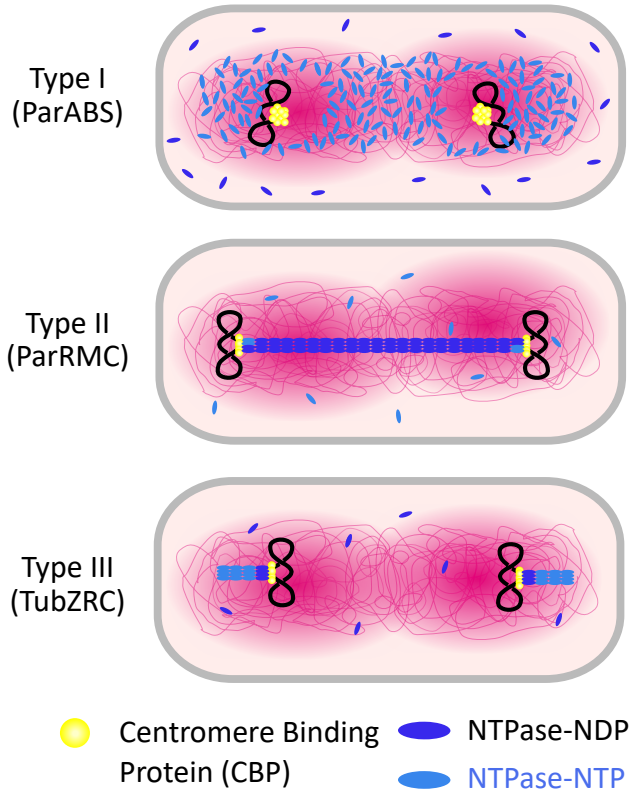
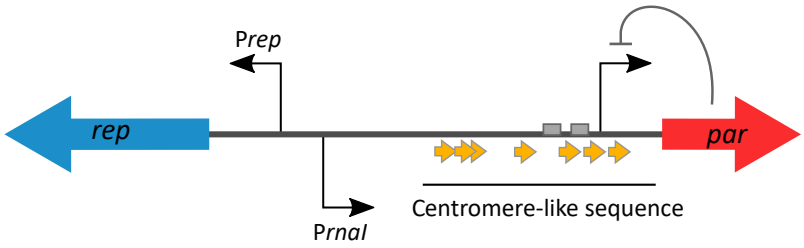

Type I
(ParABS)
Type II
(ParRMC)
Type III
(TubZRC)
Centromere Binding
Protein (CBP)
NTPase-NDP
NTPase-NTP

Figure 1

A


Prep
rep
par
PrnaI
Centromere-like sequence

B

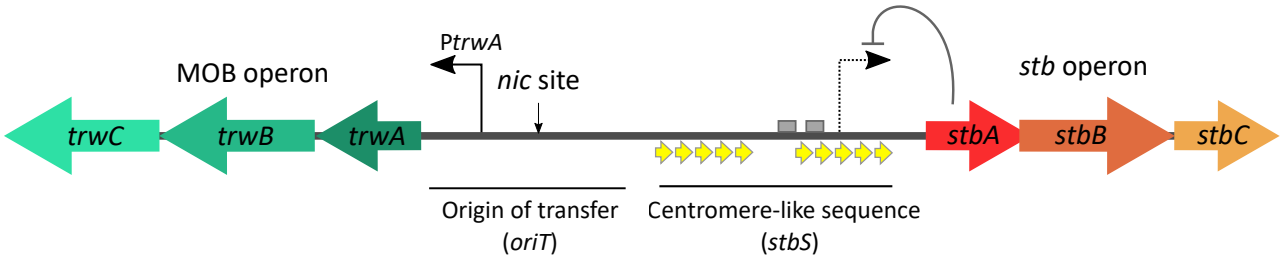
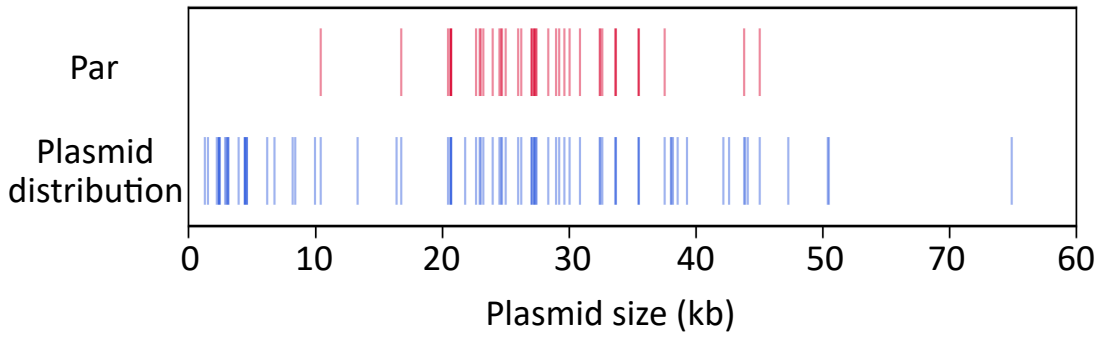

PtrwA
MOB operon
nic site
stb operon
trwC
trwB
trwA
stbA
stbB
stbC
Origin of transfer
(oriT)
Centromere-like sequence
(stbS)

Figure 2


Par
Plasmid distribution
0
10
20
30
40
50
70
60
Plasmid size (kb)

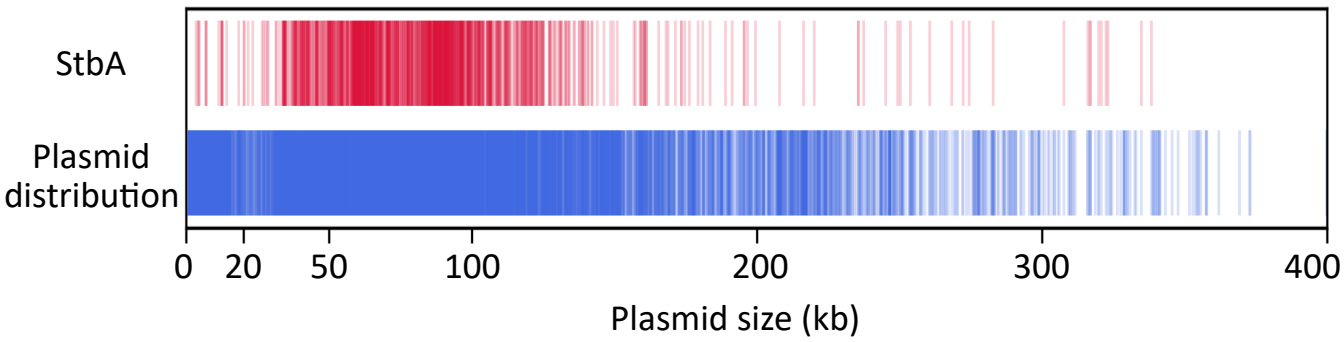

StbA
Plasmid distribution
0
20
50
100
200
300
400
Plasmid size (kb)

Figure 3

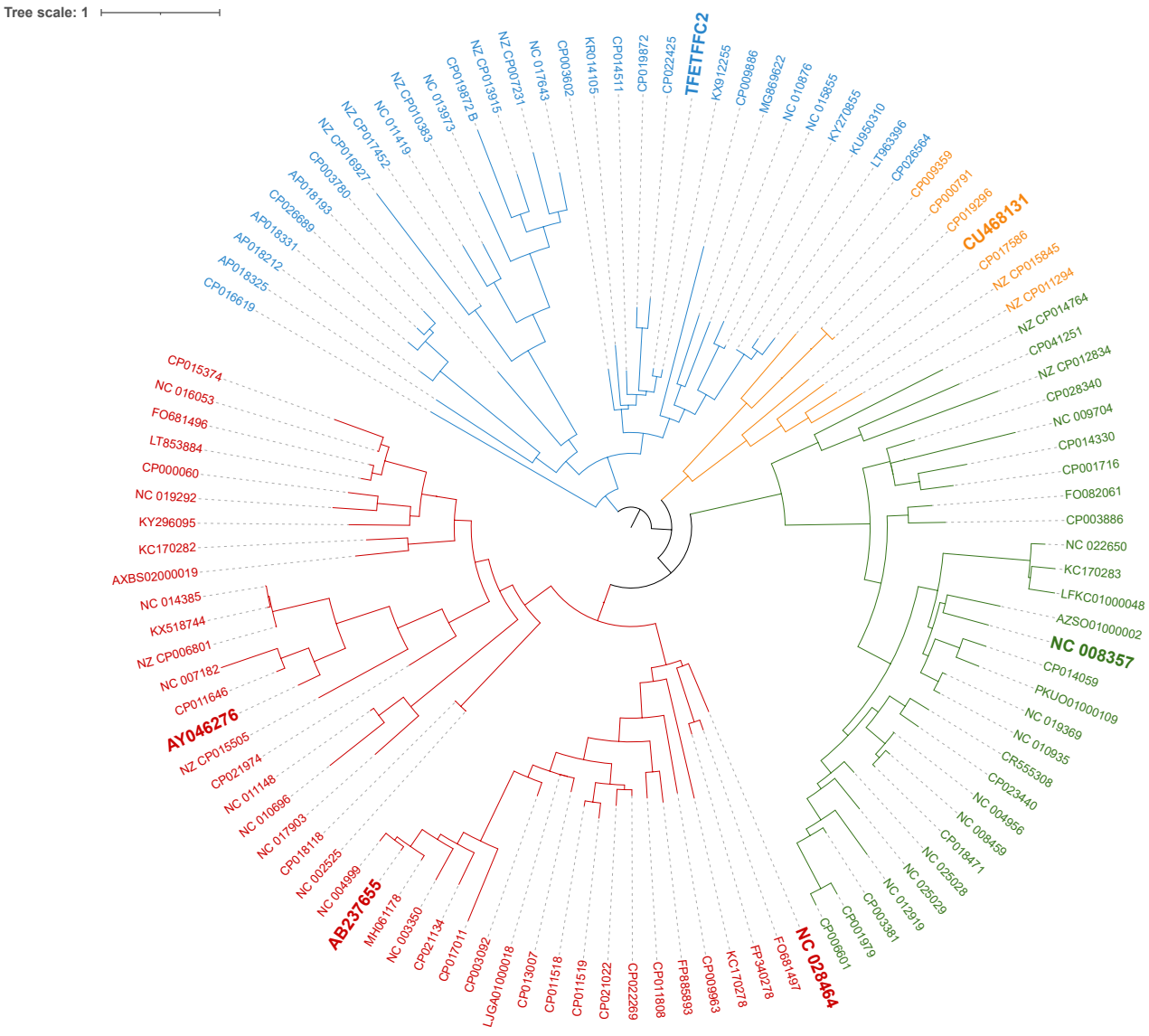
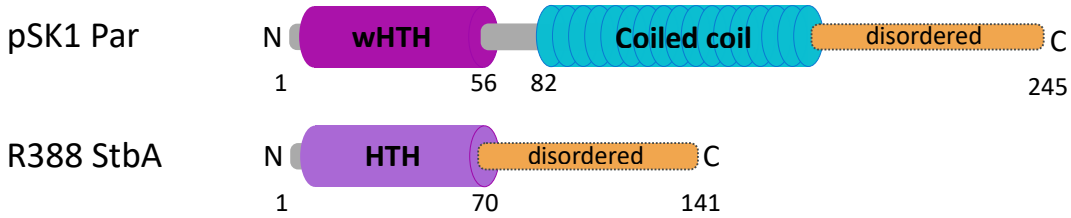
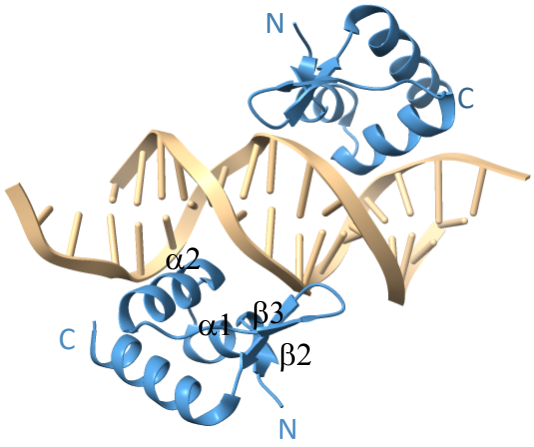

Tree scale: 1
TFETFFC2
CU468131
NC 008357
NC 028464
AB237655
AY046276
Group 1 (R388, NAH7, R46, R721)
Group 2 (pBS228)
Group 3 (pTF-FC2)
Group 4 (pET49)

Figure 4

A


pSK1 Par
N
wHTH
Coiled coil
disordered
C
1
56
82
245
R388 StbA
N
HTH
disordered
C
1
70
141

B


N
C
α2
α1
β3
β2
C
N

C

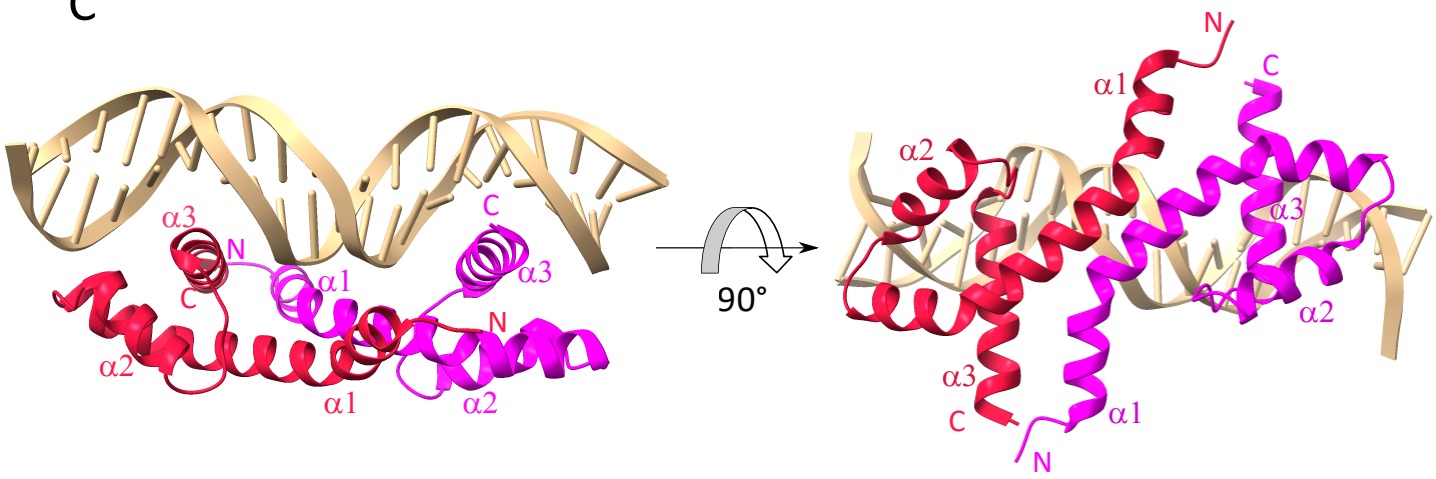

α3
N
C
α1
C
α3
N
α2
α1
α2
90°
N
α1
C
α2
α3
α2
α3
C
α1
N

Figure 5

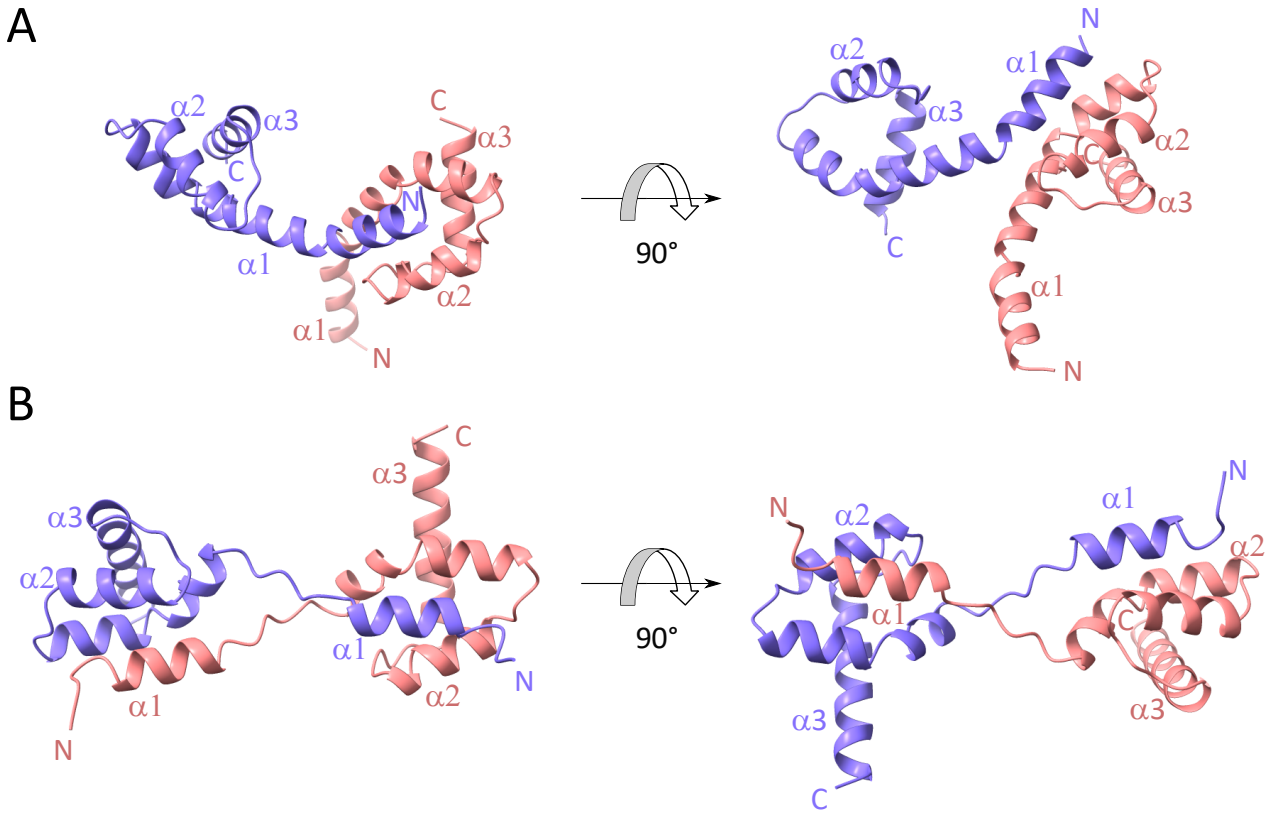

A
α2
α3
C
α3
C
N
α1
α2
α1
N
90°
N
α2
α1
α3
α2
C
α3
α1
N
B
C
α3
α3
α2
α1
N
α1
α2
N
90°
N
α2
α1
N
α2
α1
C
α3
α3
C

Figure 6

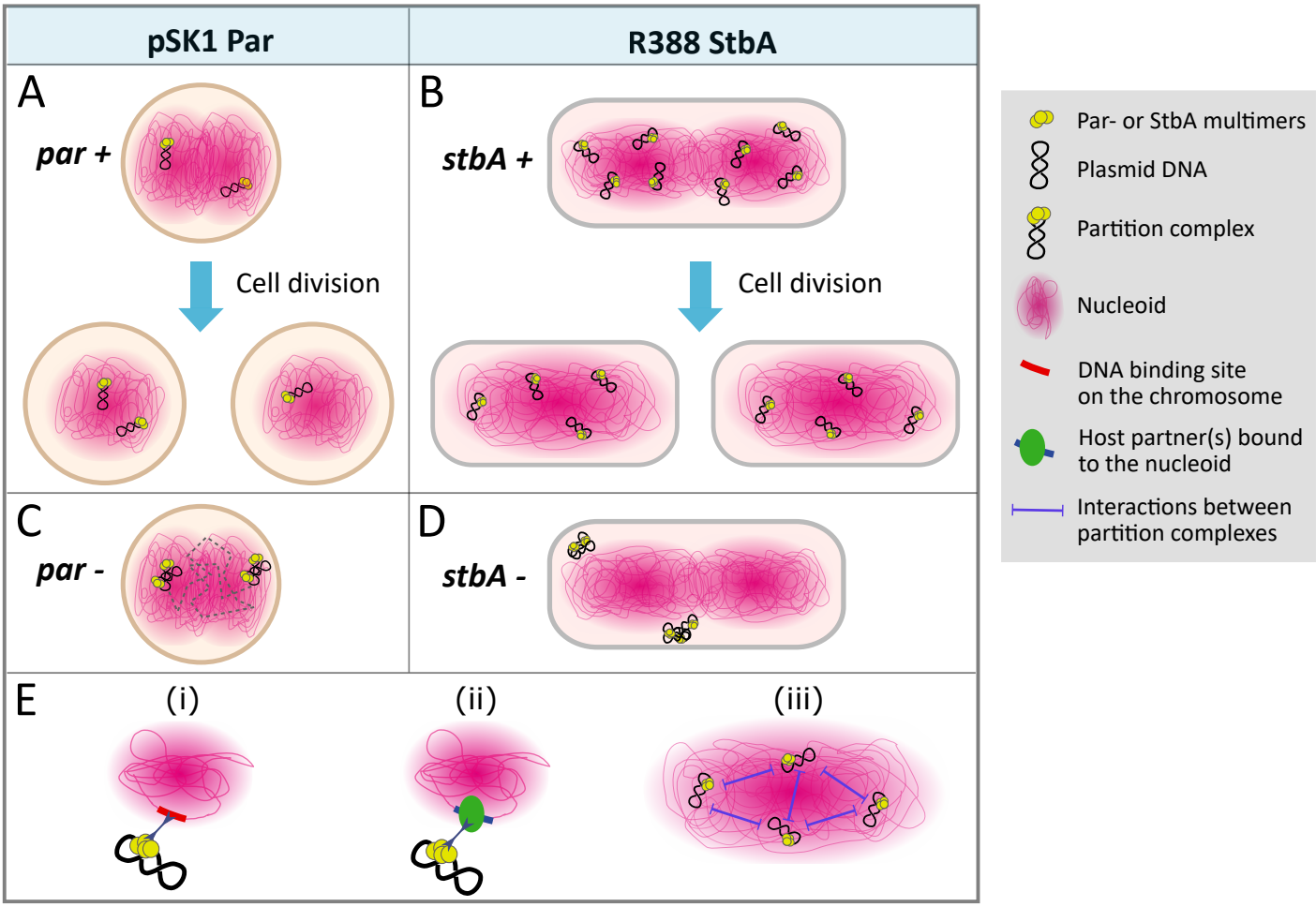
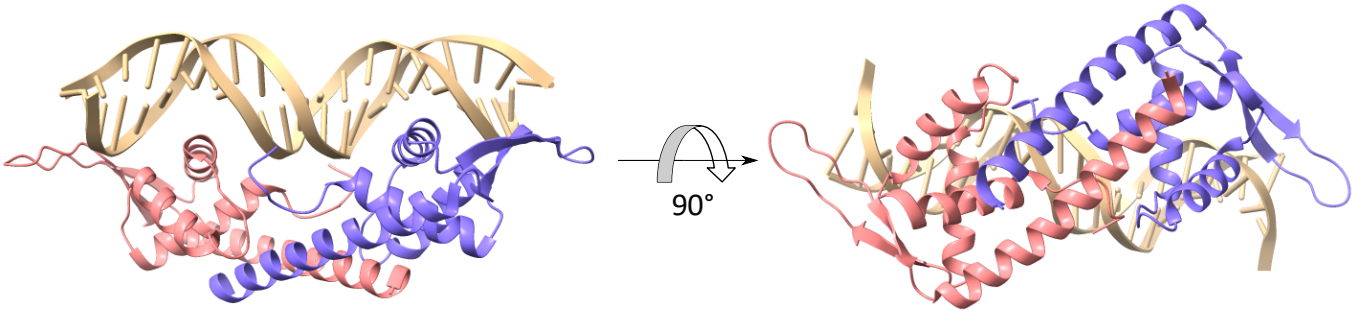

pSK1 Par
R388 StbA
A
par +
Cell division
B
stbA +
Cell division
C
par -
D
stbA -
E
(i)
(ii)
(iii)
Par- or StbA multimers
Plasmid DNA
Partition complex
Nucleoid
DNA binding site on the chromosome
Host partner(s) bound to the nucleoid
Interactions between partition complexes

Figure 7

A



B

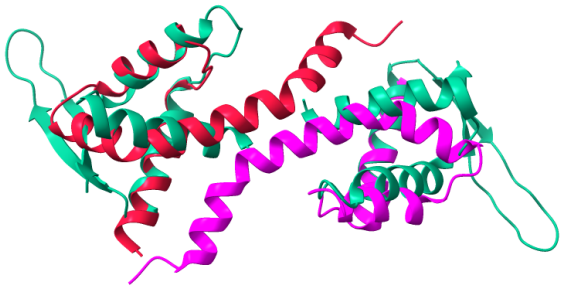


Figure S1

group	Plasmid ID	Host	Plasmid	Plasmid size (kb)	Host taxonomy
1	CP015374	Pandoraea pnomenusa strain MCB032	plasmid unnamed 3	29016	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Pandoraea
1	NC_016053	Xanthomonas arboricola pv. pruni str. CFBP 5530	pXap41	41102	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	FO681496	Xanthomonas fuscans subsp. fuscans str. 4834-R	p1b	19514	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	LT853884	Xanthomonas fragariae strain PD885	pPD885-27	27106	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	CP000060	Pseudomonas savastanoi pv. phaseolicola 1448A	sma11	51711	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	NC_019292	Pseudomonas savastanoi NCPPB 3335	pPsv48C	42103	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	KY296095	Pseudomonas aeruginosa strain 14057	p14057-KPC	516636	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	KC170282	Uncultured bacterium	pMBUI6	47999	Bacteria; environmental samples
1	AXBS02000019	Xylella fastidiosa 6c (whole genome shotgun sequence)	pXF6c	39572	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xylella
1	NC_014385	Escherichia coli	pEC_L46	144871	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia
1	KX518744	Escherichia coli strain HYEC7	pHYEC7-110	110226	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia
1	NZ_CP006801	Klebsiella pneumoniae subsp. pneumoniae PittNDM01	p3	70814	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Klebsiella/Raoultella group; Klebsiella.
1	NC_007182	Sodalis glossinidius	pSG1	81553	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Bruguierivoracaceae; Sodalis
1	CP011646	Klebsiella pneumoniae strain CAV1596	pKPC_CAV1596-97	96702	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Klebsiella/Raoultella group; Klebsiella
1	**AY046276**	Salmonella enterica subsp. enterica serovar Typhimurium	R46	50969	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Salmonella
1	NZ_CP015505	Klebsiella pneumoniae strain SKGH01	unnamed 5	34611	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Klebsiella/Raoultella group; Klebsiella.
1	CP021974	Methylophaga nitratireducenticrescens strain GP59	pGP59-34	33560	Bacteria; Pseudomonadota; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Methylophaga
1	NC_021148	Salmonella enterica subsp. enterica serovar Agona str. SL483	unnamed	37978	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Salmonella.
1	NC_010696	Erwinia tasmaniensis Et1/99	pET35	35494	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Erwiniaceae; Erwinia.
1	NC_017903	Escherichia coli Xuzhou21	p0157_Sa1	37785	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia.
1	CP018118	Escherichia coli strain MRSN346638	pMRSN346638_64.5	64467	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia
1	NC_002525	Escherichia coli K-12	R721	75582	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia
1	NC_004999	Pseudomonas putida NCIB 9816-4	pDTG1	83042	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	**AB237655**	Pseudomonas putida	pNAH7	82232	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	MH061178	Pseudomonas thivervalensis strain P101	pPHE101	60958	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	NC_003350	Pseudomonas putida	pWW0	116580	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	CP021134	Pseudomonas fragi strain NMC25	unnamed2	54359	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
1	CP017011	Pseudomonas syringae pv. actinidiae strain NZ-47	pPsa22180b	55944	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas syringae
1	CP003092	Burkholderia sp. YI23	byi_3p	115232	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Burkholderia
1	LJGA01000018	Xanthomonas citri pv. citri strain NIGEB-88 (whole genome shotgun sequence)	pXCC_55	55677	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	CP013007	Xanthomonas citri pv. malvacearum strain XcmH1005	pXcmN	59644	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	CP011518	Pandoraea oxalativorans strain DSM 23570	pPO70_1	633357	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Pandoraea
1	CP011519	Pandoraea oxalativorans strain DSM 23570	pPO70-2	126976	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Pandoraea
1	CP021022	Xanthomonas citri pv. phaseoli var. fuscans strain CFBP6167	pG	42380	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	CP022269	Xanthomonas citri pv. vignicola strain CFBP7112	p1B	42275	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	CP011808	Pandoraea faecigallinarum strain DSM 23572	pPF72-1	386625	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Pandoraea
1	FP885893	Ralstonia solanacearum CMR15	pRSC35	35008	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Ralstonia
1	CP009963	Collimonas arenae strain Ca135	Collimonas_plasmid	41440	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Oxalobacteraceae; Collimonas
1	KC170278	Uncultured bacterium	pMBUI4	37247	Bacteria; environmental samples
1	FP340278	Xanthomonas albilineans str. GPE PC73	plasmI1	31555	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	FO681497	Xanthomonas fuscans subsp. fuscans str. 4834-R	p1c	41950	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
1	**NC_028464**	Escherichia coli	R388	33913	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia
2	CP006601	Cycloclasticus zancles 78-ME	p7ME01	42347	Bacteria; Pseudomonadota; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Cycloclasticus
2	CP001979	Marinobacter adhaerens HP15	pHP-42	42349	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Marinobacteraceae; Marinobacter
2	CP003381	Methylophaga frappieri strain JAM7	plasmid	47825	Bacteria; Pseudomonadota; Gammaproteobacteria; Thiotrichales; Piscirickettsiaceae; Methylophaga
2	NC_012919	Photobacterium damselae subsp. piscicida	pP9014	55851	Bacteria; Pseudomonadota; Gammaproteobacteria; Vibrionales; Vibrionaceae; Photobacterium
2	NC_025029	Uncultured bacterium pAKD4	pAKD4	56803	Bacteria; environmental samples
2	NC_025028	Uncultured bacterium pMCBF6	pMCBF6	66615	Bacteria; environmental samples
2	CP018471	Xanthomonas vesicatoria strain LM159	pLM159.2	62784	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
2	NC_008459	Bordetella pertussis	pBP136	41268	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Alcaligenaceae; Bordetella
2	NC_004956	Pseudomonas sp.	pADP-1	108845	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
2	CP023440	Thauera sp. K11	pTX1	140963	Bacteria; Pseudomonadota; Betaproteobacteria; Rhodocyclales; Zoogloeaceae; Thauera
2	CR555308	Azoarcus sp. EbN1	plasmid 2	223670	Bacteria; Pseudomonadota; Betaproteobacteria; Rhodocyclales; Rhodocyclaceae; Aromatoleum
2	NC_010935	Comamonas testosteroni CNB-1	pCNB	91181	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Comamonadaceae; Comamonas
2	NC_019369	Burkholderia cepacia	pYS1	82988	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Burkholderia; Burkholderia cepacia complex
2	PKUO01000109	Sedimenticola sp. isolate BM503 sc_pri(whole genome shotgun sequence)		70620	Bacteria; Pseudomonadota; Gammaproteobacteria; Sedimenticola.
2	CP014059	Achromobacter xylosoxidans strain FDAARGOS_147	plasmid	21970	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Alcaligenaceae; Achromobacter
2	**NC_008357**	Pseudomonas aeruginosa	pBS228	89147	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
2	AZSO01000002	Hydrogenophaga sp. T4 (whole genome shotgun sequence)		42735	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Comamonadaceae; Hydrogenophaga.
2	LFKC01000048	Escherichia coli strain HMLN-1 (whole genome shotgun sequence)	unnamed	55844	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia
2	KC170283	Uncultured bacterium	pDS1	40596	Bacteria; environmental samples
2	NC_022650	Escherichia coli JJ1886	plasmid pJJ1886_4	55956	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia
2	CP003886	Legionella pneumophila subsp. pneumophila LPE509	plasmid	73490	Bacteria; Pseudomonadota; Gammaproteobacteria; Legionellales; Legionellaceae; Legionella
2	FO082061	Methylomicrobium alcaliphilum str. 20Z	MEALZ_p	128415	Bacteria; Pseudomonadota; Gammaproteobacteria; Methylococcales; Methylococcaceae; Methylotuvimicrobium
2	CP001716	Candidatus Accumulibacter phosphatis clade IIA str. UW-1	pAph01	167595	Bacteria; Proteobacteria; Betaproteobacteria; Candidatus Accumulibacter
2	CP014330	Xylella fastidiosa strain FB7	plasmid	39408	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xylella
2	NC_009704	Yersinia pseudotuberculosis IP 31758	p_59kb	58679	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Yersiniaceae; Yersinia.
2	CP028340	Thauera aromatica K172	pKJK172	53761	Bacteria; Pseudomonadota; Betaproteobacteria; Rhodocyclales; Zoogloeaceae; Thauera
2	NZ_CP012834	Salmonella enterica subsp. enterica serovar Cerro str. CFSAN001588	pCFSAN001588_001	53952	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Salmonella.
2	CP041251	Raoultella electrica strain DSM 102253	unnamed4	35253	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Klebsiella/Raoultella group; Raoultella
2	NZ_CP014764	Klebsiella pneumoniae strain KPNIH39	pKPN-704	36707	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Klebsiella/Raoultella group; Klebsiella.
4	CP019296	Vibrio campbellii strain LMB29	pLMB143	143114	Bacteria; Pseudomonadota; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio
4	CP000791	Vibrio campbellii ATCC BAA-1116	pVIBHAR	89008	Bacteria; Pseudomonadota; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio
4	CP009359	Vibrio tubiashii ATCC 19109	p48	47973	Bacteria; Pseudomonadota; Gammaproteobacteria; Vibrionales; Vibrionaceae; Vibrio; Vibrio oreintalis group
4	NZ_CP011294	Salmonella enterica subsp. diarizonae strain 11-01854	unnamed2	57955	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Salmonella.
4	NZ_CP015845	Escherichia coli O157:H7 strain FRIK2455	p35K	34689	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia.
4	CP017586	Pantoea stewartii subsp. stewartii DC283	pDSJ05	34447	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Erwiniaceae; Pantoea
4	**CU468131**	Erwinia tasmaniensis strain ET1/99	pET49	48751	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Erwiniaceae; Erwinia
3	CP026564	Pseudomonas avellanae strain R2leaf	p2_tig5	102862	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
3	LT963396	Pseudomonas cerasi isolate PL963 genome assembly	Pp1	127474	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
3	KU950310	Pseudomonas syringae pv. actinidiae strain SR198	pMG2_SR198	111158	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas syringae
3	KY270855	Pseudomonas putida strain 12969	p12969-2	109708	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
3	NC_015855	Pseudomonas putida	pGRT1	133451	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas
3	NC_010876	Xanthomonas axonopodis pv. glycines strain 8ra	pXAG81	26721	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
3	MG869622	Polaromonas sp. H8N	pH8NP2	38325	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Comamonadaceae; Polaromonas
3	CP009886	Xylella fastidiosa strain Hib4	pXF64-HB	64251	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xylella
3	KX912255	Enterobacter cloacae strain H140960786	pJF-786	25354	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Enterobacter; Enterobacter cloacae complex
3	**TFETFFC2**	Thiobacillus ferrooxidans (incomplete)	pTF-FC2		Bacteria; Pseudomonadota; Acidithiobacillia; Acidithiobacillales; Acidithiobacillaceae; Acidithiobacillus
3	CP022425	Vitreoscilla filiformis strain ATCC 15551	pVF2	40016	Bacteria; Pseudomonadota; Betaproteobacteria; Neisseriales; Neisseriaceae; Vitreoscilla
3	CP019872	Pseudomonas syringae pv. tomato strain B13-200	pB13-200A	125801	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas syringae
3	CP014511	Burkholderia sp. PAMC 28687 strain PAMC28687	plasmid 4	20973	Bacteria; Pseudomonadota; Betaproteobacteria; Burkholderiales; Burkholderiaceae; Burkholderia
3	KR014105	Aeromonas hydrophila strain WCHAH01	pGES5	32664	Bacteria; Pseudomonadota; Gammaproteobacteria; Aeromonadales; Aeromonadaceae; Aeromonas
3	CP003602	Chamaesiphon minutus PCC 6605	pCHA6605.02	31019	Bacteria; Cyanobacteria; Synechococcales; Chamaesiphonaceae; Chamaesiphon
3	NC_017643	Escherichia coli UMNK88	pUMNK88_H1y	65549	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia.
3	NZ_CP007231	Yersinia similis strain 228	plasmid	60687	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Yersiniaceae; Yersinia.
3	NZ_CP013915	Serratia fonticola strain GS2	pSF002	93737	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Yersiniaceae; Serratia.
3	CP019872	Pseudomonas syringae pv. tomato strain B13-200	pB13-200A	125801	Bacteria; Pseudomonadota; Gammaproteobacteria; Pseudomonadales; Pseudomonadaceae; Pseudomonas; Pseudomonas syringae
3	NC_013973	Erwinia amylovora ATCC 49946	plasmid 2	71487	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Erwiniaceae; Erwinia.
3	NZ_CP010383	Enterobacter hormaechei subsp. steigerwaltii strain 34998	p34998-106.409kb	106410	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Enterobacter; Enterobacter cloacae complex.
3	NC_011419	Escherichia coli SE11	pSE11-1	100021	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Escherichia.
3	NZ_CP017452	Klebsiella sp. LTGPAF-6F	unnamed2	89552	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Klebsiella/Raoultella group; Klebsiella.
3	NZ_CP016927	Klebsiella pneumoniae isolate 23	pIncL_M_DHQP1400954	72093	Bacteria; Pseudomonadota; Gammaproteobacteria; Enterobacterales; Enterobacteriaceae; Klebsiella/Raoultella group; Klebsiella.
3	CP003780	Xanthomonas citri subsp. citri Aw12879	pXcaw58	58317	Bacteria; Pseudomonadota; Gammaproteobacteria; Xanthomonadales; Xanthomonadaceae; Xanthomonas
3	AP018193	Nostoc sp. NIES-2111	plasmid9	38211	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc
3	CP026689	Nostoc sp. 'Peltigera membranacea cyanobiont N6' strain N6	pNPM7	29551	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc
3	AP018331	Nostoc commune HK-02	plasmid5	34879	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc
3	AP018212	Calothrix brevissima NIES-22	plasmid5	57802	Bacteria; Cyanobacteria; Nostocales; Calotrichaceae; Calothrix
3	AP018325	Nostoc sp. HK-01	plasmid7	29094	Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc
3	CP016619	Microvirga sp. V5/3M	unnamed2	977332	Bacteria; Pseudomonadota; Alphaproteobacteria; Hyphomicrobiales; Methylobacteriaceae; Microvirga

Table S1