Quantitative measurements and modelling of bacterial DNA segregation

Dissertation

kumulativ

Zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

Dem Fachbereich Biologie
der Philipps-Universität Marburg
vorgelegt von

Robin Clemens Köhler
aus Münster, Deutschland

Marburg an der Lahn, Deutschland, Mai 2023

Die vorliegende Dissertation wurde von Mai 2020 bis Mai 2023 am Max-Planck-Institut für terrestrische Mikrobiologie unter Leitung von Dr. Seán Murray angefertigt.

Vom Fachbereich Biologie der Philipps-Universität Marburg (Hochschulkennziffer 1180) als Dissertation angenommen am _____

Erstgutachter: Dr. Seán Murray

Zweitgutachter: Prof. Dr. Martin Thanbichler

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Peter Graumann

Prof. Dr. Lennart Randau

Tag der Disputation:_____

Erklärung

Ich versichere, dass ich meine Dissertation mit dem Titel „Quantitative measurements and modelling of bacterial DNA segregation“ selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfsmittel bedient habe.

Diese Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den:_____

_____

Robin Köhler

Erklärung zum Eigenanteil der Doktorandin/des Doktoranden an den vorgelegten Publikationen/Manuskripten

Titel der Publikation:

High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS

Erschienen in:

Köhler, R., Kaganovitch, E., & Murray, S. M. (2022). High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS. ELife, 11, e78743. https://doi.org/10.7554/eLife.78743

Die oben aufgeführte Publikation soll im Rahmen dieser kumulativen Dissertation verwendet werden. Meine wissenschaftlichen Anteile an der Veröffentlichung sind die Folgenden: Data Generation, Data curation, Software, Formal analysis, Investigation, Visualization, Methodology, Writing

Titel der Publikation:

High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS

Erschienen in:

Köhler, R., Sadhir, I., & Murray, S. M. (2023). ★Track: Inferred counting and tracking of replicating DNA loci. loci. Biophysical Journal, 122(9), 1577–1585. https://doi.org/10.1016/j.bpj.2023.03.033

Die oben aufgeführte Publikation soll im Rahmen dieser kumulativen Dissertation verwendet werden. Meine wissenschaftlichen Anteile an der Veröffentlichung gestalten sind die Folgenden: Experiments, Data curation, Software, Investigation, Visualization, Methodology, Writing

Die oben genannten Publikationen sind kein Bestandteil einer weiteren wissenschaftlichen Qualifikationsarbeit

Ich, Seán Murray bin mit der Abfassung der Dissertation als kumulative Dissertation einverstanden und bestätige die vorstehende Angabe:

_____ _____

Datum, Erstbetreuer(in) Datum, Doktorand(in)

Table of Contents

1. Abstract/Summary

1.1. Deutsch

Eine genaue Segregation von genetischem Material ist für das Überleben und die Vermehrung aller Organismen, ob eukaryotisch oder prokaryotisch, unerlässlich. Neben dem Chromosom sind Bakterien dafür bekannt, extrachromosomale DNA-Moleküle zu tragen, die als Plasmide bezeichnet werden und in bestimmten Situationen Wachstumsvorteile bieten und oft Antibiotikaresistenzen kodieren. Die Kopienzahlen von Plasmiden können je nach spezifischem Plasmid stark variieren, von mehreren hundert Kopien pro Zelle bis zu weniger als einer Handvoll. Plasmide mit hoher Kopienzahl können sich auf ihre Anzahl verlassen, um sicherzustellen, dass mindestens ein Plasmid während der Zellteilung an jede Tochterzelle verteilt wird. Plasmide mit geringer Kopienzahl hingegen benötigen spezielle Partitionierungssysteme, um eine korrekte Segregation sicherzustellen und ihren Verlust während der Zellteilung zu verhindern. Das ParABS-System ist das häufigste dieser Partitionierungssysteme und besteht aus drei Komponenten: parS, einer zentromerähnlichen Sequenz in der Nähe des Replikationsursprungs auf dem Plasmid; ParA, eine Walker-Typ-ATPase, die unspezifisch an das Nukleoid in ihrem dimeren ATP-gebundenen Zustand bindet; ParB ist ebenfalls ein Protein, das Dimere bildet und spezifisch an eine parS-Stelle bindet. Sobald es an parS gebunden ist, breitet sich ParB mehrere Kilobasen in beide Richtungen aus, und zusammen bilden parS und ParB einen Nukleoproteinkomplex. Dieser Komplex interagiert mit nukleoidgebundenem ParA, was zur gleichmäßigen Positionierung von Plasmiden entlang der Längsachse der Zelle führt. Diese Art der Positionierung stellt sicher, dass beide Tochterzellen bei der Zellteilung eine gleiche Anzahl von Plasmiden erhalten. Der Hauptfokus dieser Arbeit liegt auf dem F-Plasmid, einem Plasmid mit geringer Kopienzahl, das ein ParABS-System besitzt und durchschnittlich drei Kopien pro Zelle aufweist. Die Platzierung dieses Plasmids und anderer ParABS-tragender Plasmide innerhalb der Zelle ist Gegenstand von Diskussionen. Einige vermuten, dass Plasmide innerhalb der Zelle durch oszillatorische Bewegung positioniert werden, während andere gerichtete Bewegung vorschlagen, die die Plasmide direkt an ihren Zielpositionen platzieren. Durch den Einsatz von hochdurchsatzdatenerfassung und -analyse konnten wir Tausende von Zellzyklen mit F-Plasmiden erfassen, um die wahren Dynamiken der Plasmidpositionierung dieses Systems aufzudecken. Um die Bewegung des F-Plasmids zu verfolgen, verwenden wir ParB, das an ein fluoreszierendes Protein gekoppelt ist, das mit dem Plasmid kolokalisiert und helle Foci bildet, die leicht verfolgbar sind. Darüber hinaus entwickelten wir basierend auf einem zuvor entwickelten Modell ein vereinheitlichendes Modell, das die *in vivo* Beobachtungen nicht nur des F-Plasmids, sondern auch eines anderen entfernt verwandten ParABS-Systems (pB171) genau reproduziert. Unsere Ergebnisse, basierend auf experimentellen Daten und Simulationen, zeigen, dass das F-Plasmid eine zielgerichtete

und regelmäßige Positionierung aufweist: Die Plasmide bewegen sich präzise zu ihren Zielpositionen und stellen somit sicher, dass sie gleichmäßig in der ganzen Zelle verteilt sind. Unser Modell zeigt, dass das F-Plasmid knapp über der Schwelle einer oszillatorischen Instabilität arbeitet. Wie von unserem Modell vorhergesagt, überschreitten große Zellen mit einem einzigen Plasmid diese Schwelle. Im Gegensatz dazu zeigten Zellen mit einem anderen Plasmid, den pB171-Plasmiden, bei niedrigen Plasmidkonzentrationen deutliche Pol-zu-Pol-Oszillationen, aber wenn die Konzentration zunahm, wurden die Plasmide regelmäßig positioniert. Trotz dieser signifikanten Unterschiede in der Plasmidbewegung konnte unser Modell diese unterschiedlichen Plasmiddynamiken mit einem einzigen dimensionslosen Parameter namens λ erklären, der die durchschnittliche Diffusionsstrecke von ParA auf dem Nukleoid beschreibt. Weiterhin ergaben unsere Simulationen ein interessantes Ergebnis: Wenn sich das System dem oszillatorischen Regime näherte, verringerte sich sein Energieverbrauch. Dieser Fund liefert eine mögliche Erklärung dafür, warum diese Systeme so nahe zur oszillatorischen Instabilität arbeiten.

Eine Herausforderung, auf die wir während der Untersuchung des ParABS-Systems stießen, war die Schwierigkeit, Plasmide bei hohen Zahlen genau zu verfolgen. Um dieser Herausforderung zu begegnen, entwickelten wir einen Algorithmus namens ⋆Track, der beständige Objekte wie Plasmide mit hoher Genauigkeit verfolgt. Wir testeten diesen Algorithmus sowohl für F-Plasmide als auch für chromosomale Loci und machten interessante Beobachtungen und Erkenntnisse in beiden Systemen.

Aufbauend auf unseren bisherigen Ergebnissen untersuchten wir auch das Zusammenspiel zwischen Plasmidbewegung und ParA-Lokalisierung. Es ist bekannt, dass ParA von einer Zellhälfte zur anderen oszilliert, aber die Funktion und biologische Relevanz dieser Oszillationen sind unbekannt. Durch den Einsatz von Hochdurchsatz-Datenerfassung und -analyse identifizierten wir die Merkmale dieser Oszillationen und fanden heraus, dass sie zum Vorschein kommen, sobald das Nukleoid eine Verengung aufweist. Wir schlagen vor, dass dieser Effekt eine Folge der "ParA-ParA-Rekrutierung" ist, bei der nukleoidgebundenes ParA zytosolisches ParA rekrutiert, um in enger Nähe den Nukleoid zu binden. Um diese Behauptung zu unterstützen, haben wir die ParA-ParA-Rekrutierung in unser Computermodell integriert, was es uns ermöglichte, die in vivo beobachteten Oszillationen von ParA genau nachzubilden. Unsere Hypothese für die biologische Relevanz dieses Mechanismus ist, dass ParA-Oszillationen dem System ermöglichen, Plasmide zwischen zwei getrennten Schwester-Nukleoiden vor der Zellteilung aufzuteilen, wodurch eine symmetrische Vererbung der Plasmide an den Nachwuchs ermöglicht wird.

Zusammenfassend präsentiert diese Dissertation neue Erkenntnisse über das ParABS-System und seine Rolle bei der präzisen Segregation von Plasmiden. Unsere Ergebnisse ebnen den Weg für zukünftige Forschungen, um die Feinheiten dieses komplexen, aber minimalistischen Systems weiter zu entschlüsseln.

1.2. English

Accurate segregation of genetic material is essential for the survival and proliferation of all organisms, whether eukaryotic or prokaryotic. In addition to the chromosome, bacteria are known to carry extrachromosomal DNA molecules called plasmids, which can provide growth advantages in certain situations and often encode antibiotic resistances. The copy numbers of plasmids can differ greatly based on the specific plasmid, ranging from hundreds of copies per cell for some to less than a handful for others. Plasmids with high copy numbers can rely on their abundance to ensure that at least one plasmid is distributed to each daughter cell during cell division. Low-copy number plasmids, on the other hand, require specialized partitioning systems to ensure their faithful segregation and prevent their loss during cell division. The ParABS system is the most common of these partitioning systems, consisting of three key components: *parS*, a centromere-like sequence located near the origin of replication on the plasmid; ParA, a walker-type ATPase that binds nonspecifically to the nucleoid in its dimeric ATP-bound state; ParB, is also a protein that forms dimers and binds specifically to a parS site. Once bound to parS, ParB spreads several kilobases in both directions and together *parS* and ParB form a nucleoprotein complex. This complex interacts with nucleoid-bound ParA, resulting in the equidistant positioning of plasmids along the long axis of the cell. This kind of positioning ensures that both daughter cells receive an equal number of plasmids upon cell division.

The primary focus of this thesis is the F-plasmid, a low-copy number plasmid that harbors a ParABS system and maintains an average of three copies per cell. The placement of this plasmid and other ParABS carrying plasmids within the cell has been a topic of debate. Some suggest that plasmids within the cell are positioned through oscillatory motion, while others propose directed motion, placing the plasmids directly at their target positions. By utilizing high-throughput data acquisition and analysis, we captured thousands of cell cycles containing F-plasmids to reveal the true dynamics of plasmid positioning of this system. To track the motion of the F-plasmid, we used ParB fused to a fluorophore, which colocalizes with the plasmid and forms bright foci that are easily trackable. Furthermore, based on a previously developed molecular-scale model, we developed a unifying model that accurately reproduces the *in vivo* observations not only of the F-plasmid but also of another distantly related ParABS system (pB171). Our findings, based on both experimental data and simulations, prove that the F-plasmid exhibits true regular positioning: Plasmids move precisely to their target positions, ensuring they are equidistantly spaced throughout the cell. However, our model indicates that the F-plasmid operates just above the threshold of an oscillatory instability. As predicted by our model, large cells with a single plasmid crossed this threshold and exhibited low amplitude oscillations. In contrast, cells containing a different plasmid, the pB171 plasmids, exhibited clear pole-to-pole oscillations at low plasmid concentrations, but as the concentration increased, the plasmids became regularly positioned. Despite these significant differences in plasmid locomotion,

our model was able to explain these differing plasmid dynamics using a single dimensionless parameter named λ, which describes the average distance that ParA diffuses on the nucleoid. Further, our simulations yielded an interesting result: as the system approached the oscillatory regime, its energy consumption decreased. This finding provides a possible explanation for why these systems operate in such close proximity to the oscillatory instability.

One challenge we encountered during the investigation of the ParABS system was the difficulty of tracking plasmids accurately at high numbers. To address this challenge, we developed an algorithm called ⋆Track, which tracks persistent objects like plasmids with high accuracy. We tested this algorithm on both F-plasmid and chromosomal loci, revealing interesting observations and findings in both systems.

Building upon our previous results, we also examined the interplay between plasmid movement and ParA localization. ParA is known to oscillate from one cell half to the other, but the function and biological relevance of these oscillations are unknown. By using high-throughput data acquisition and analysis, we identified the characteristics of these oscillations and found that they start to emerge once the nucleoid shows a constriction. We propose that this effect is a consequence of "ParA-ParA recruitment" in which nucleoid-bound ParA recruits cytosolic ParA to bind in close proximity to the nucleoid. To support this claim, we incorporated ParA-ParA recruitment into our computational model, which enabled us to accurately replicate the *in vivo* observed oscillations of ParA. Our hypothesis for the biological relevance of this mechanism is that ParA oscillations enable the system to partition plasmids between two separated sister nucleoids before division, allowing symmetric inheritance of plasmids to the offspring.

In conclusion, this thesis presents novel insights into the ParABS system and its role in the accurate segregation of plasmids. Our findings pave the way for future research to further unravel the intricacies of this complex yet minimalistic system.

2. Introduction

Before diving into the introduction, it should be noted that at the end of this section, a concise overview of the most crucial and relevant information is provided (section 2.3. Important features of the ParABS system). This streamlined overview is intended to offer sufficient context for comprehending the content of the three papers presented in this cumulative thesis.

The segregation of genetic material is a fundamental process that allows cells to divide and propagate their genetic information to their progeny. This process is essential for the survival of all living organisms, from simple prokaryotic bacteria to complex eukaryotic organisms such as plants and

animals. While the basic mechanisms of DNA segregation are conserved across different organisms, there are important differences in the way prokaryotic and eukaryotic cells accomplish this task.

In eukaryotes, this process is called mitosis, in which the cell undergoes a series of highly coordinated processes. First, the cell's chromosomes condense, the nucleolus dissolves, and the chromosomes attach to the spindle fibers, which then pull them apart and toward opposite poles of the cell. Once the chromosomes have been fully separated, the cell undergoes cytokinesis, during which the cytoplasm and organelles are divided between the two daughter cells.

In bacteria, the segregation of chromosomes is not dependent on a single strategy. Instead, it involves a variety of mechanisms and physical principles, given that the chromosome is free-floating in the cytosol rather than being enveloped in a membrane. Typically, multiple mechanisms are simultaneously active within a given bacterial organism, working together to ensure that replicated genetic material is symmetrically distributed to the daughter cells. A few examples are:

- Entropic forces: Spontaneous segregation of two intermingled DNA polymers. Studies have demonstrated that it is entropically favorable for two DNA polymers confined within a cell to be separated (Jun & Mulder, 2006).

- Condensation through SMC (**s**tructural **m**aintenance of **c**hromosomes): SMCs densely compact the chromosome by locally looping DNA. They are recruited to the origin of replication and move down both DNA arms simultaneously. As they do so, they extrude the DNA as a loop and bring the DNA arms together, akin to a zipper (Tran et al., 2017; X. Wang et al., 2017, 2018). This process leads to the separation and segregation of sister chromosomes (Lindow et al., 2002).

- ParABS: The ParABS system, which plays a vital role in chromosome segregation for the majority of bacteria, is composed of two proteins, ParA and ParB, along with a centromere-like binding site. Typically located near the origin of replication and present in multiple copies, the parS forms a nucleoprotein complex together with ParB. ParA plays a crucial role in positioning the ParB-DNA complex, thereby ensuring the accurate distribution of replicated chromosomes to daughter cells (Livny et al., 2007; Mohl et al., 2008; Viollier et al., 2004).

Bacteria are known to harbor extrachromosomal DNA in the form of circular molecules called plasmids, which are typically smaller in size than the chromosome but usually present in higher numbers. Certain plasmids have a copy number of more than 50 (Lin-Chao et al., 1992) and can segregate into daughter cells through Brownian diffusion. However, low copy number plasmids, some

of which have fewer than 5 copies per cell (Köhler et al., 2022), cannot rely on random diffusion and require an active system similar to the chromosome.

There are two types of systems that ensure faithful segregation of low copy number plasmids. Let us begin with the ParMRC system (Type II). It forms growing ParM-filaments, which connect to two plasmids, pushing them apart, each towards one pole (Campbell & Mullins, 2007). Each end of a growing ParM-filament eventually binds, via the DNA-adaptor protein ParR, to a *parC* sequence located on a plasmid. The ParM filament continues to grow and exerts a force on the plasmids, pushing them apart and to the poles of the cell. Once the filament reaches the pole, it rapidly and completely depolymerizes. The probability of a growing filament connecting to two plasmids is higher at the pole with more plasmids, resulting in a greater number of plasmids being moved away from that pole than towards it. This results in an even distribution of plasmids at both cell poles.

Type I systems for plasmid partitioning share the same name and components as the ParABS system, which is responsible for chromosomal segregation (mentioned above). However, plasmid partitioning systems of that type operate differently. They position the plasmids equidistantly along the long axis of the cell, such that after division, both daughter cells inherit half of the plasmids. The ParABS system involves three main components: (i) ParA, which is a Walker-type ATPase, (ii) parS, a centromere-like binding site, and (iii) ParB, a protein that binds to parS. ParB dimerizes, binds to parS, and slides along DNA for several kilobases (Breier & Grossman, 2007), enabling the buildup of a substantial amount of ParB on the plasmid, forming a condensed nucleoprotein complex. Additionally, ParB stimulates the ATPase activity of ParA. ParA, also forming dimers, binds to DNA non-specifically (G. E. Lim et al., 2005) in its dimeric-ATP-bound state, coating the nucleoid of the cell. Upon ATP hydrolysis, ParA detaches from the DNA and undergoes a prolonged cytosolic conformational transition before re-binding to DNA (Vecchiarelli et al., 2010).

This thesis will mainly focus on the ParABS system found on low copy number plasmids. As there are analogous ParABS systems found on various plasmids, we will employ the notation outlined in Figure 1 for clarity and consistency in the next section.

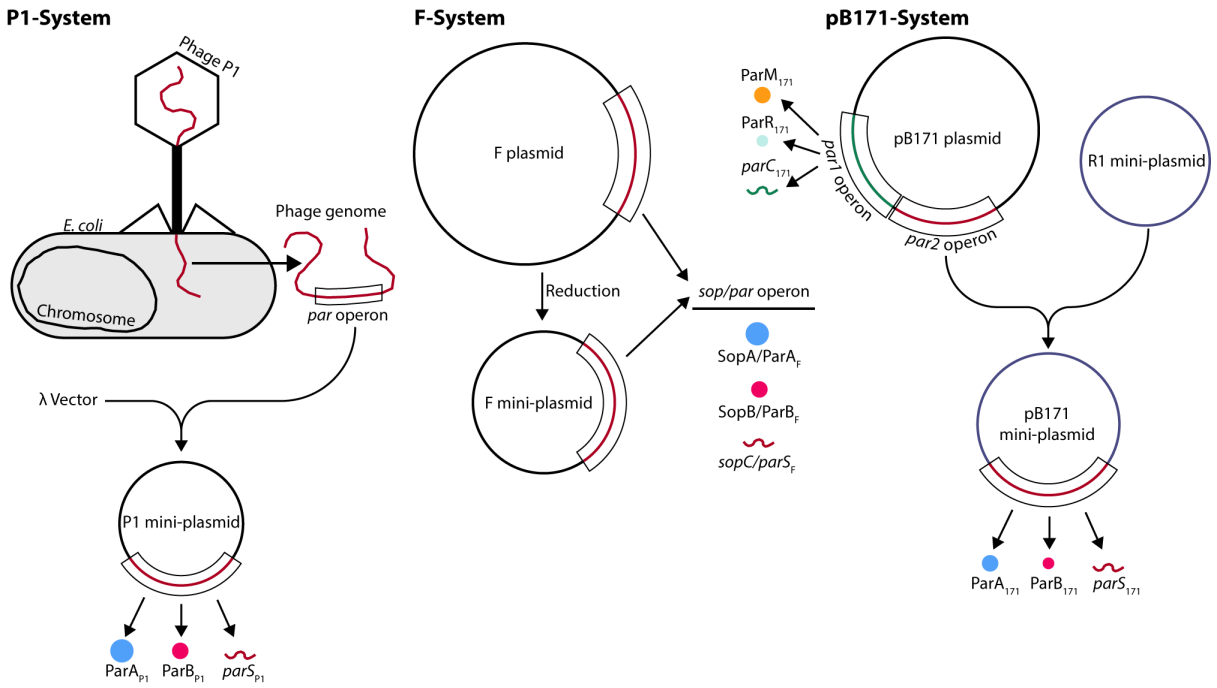

Figure 1. Overview of three different ParABS systems originating from different backgrounds. The **P1-System** originates from the P1 phage. During the lysogenic cycle, the phage genome exists in the form of a plasmid inside the host. It needs an active partitioning system, due to its low copy number. Its partitioning system was cloned into a λ vector to form a stable plasmid. The **F-Plasmid**, the F stands for fertility factor, provides bacteria with the ability to transfer genes from a bacterium carrying the factor to another bacterium that lacks it through a process called conjugation. The F mini-plasmid is a reduced version of the F plasmid that only carries essential genes necessary for plasmid maintenance. The **pB171-system** is an *E. coli* virulence factor. It contains two types of partitioning systems: *par1*, a Type II system (ParMRC) & *par2*, a Type I system (ParABS). A fragment containing *par2* was cloned into the R1 mini-plasmid.

2.1. History of the ParABS system

In the following sections, a comprehensive review will be provided, detailing the discovery and historical background of the ParABS system, as well as the scientific efforts aimed at understanding it up to the present day. A graphic overview of the history can be found at the end of this section in Figure 2. The main focus of this historical review will be on three different ParABS systems depicted in Figure 1 and for all studies, the mini-version of the plasmid was used, unless it was created as part of that study.

The initial partitioning system we will explore originates from an unlikely source: The bacteriophage P1. During the lysogenic cycle, the phage genome is distinct from the majority of phages, as it exists in the form of a plasmid within the host bacterium rather than integrating its genetic material into the host's chromosome. Given its low copy nature, this plasmid needs an active partition system to maintain stability. By inserting a fragment of this plasmid into a λ vector, researchers successfully generated a stable mini-plasmid (Austin et al., 1982). Through multiple deletions within this

mini-plasmid, a genetic map was constructed, and various regions were identified. One of these regions was found to be necessary for the plasmid to remain stable and to avoid getting lost during cell division. This development of a simple and easily-studiable plasmid marked a significant milestone, as it enabled researchers to investigate in depth the mechanisms involved in plasmid maintenance.

One year later, in 1983, the partition system of another plasmid, the F-plasmid, was identified (Ogura & Hiraga, 1983), and three distinct regions were found to be necessary for the plasmid's stability. Two of them, named sopA and sopB, are genes and the last one is named sopC (in this section referred to as $parA_F$, $parB_F$ and $parS_F$). The paper also mentions a personal communication between Hayakawa and Matsubara, which suggests that they have proof that $ParB_F$ can bind to the $parS_F$ region. Remarkably, it took 5 years for evidence of this to be published (Davis & Austin, 1988; Funnell, 1988).

The sequencing of the region required for stable plasmid maintenance of the P1 unit-copy plasmid led to the identification of two crucial genes, parA and parB (in this section referred to as $ParA_{P1}$ and $ParB_{P1}$) (Abeles et al., 1985). In this study, the term 'par' was coined as an abbreviation for plasmid partition. Additionally, the study found that both genes, $parA_{P1}$ and $parB_{P1}$, are located in close proximity to each other and share the same reading direction, with parA situated upstream of parB.

Shortly thereafter, the sequences of $parA_F$ and $parB_F$ were determined (Mori et al., 1986), and it was discovered that there was homology between $parA_F$/$parB_F$ from the F-plasmid and $parA_{P1}$/$parB_{P1}$ from the P1-plasmid. This was the first instance in which it was realized that many low copy number plasmids might utilize a similar mechanism for plasmid positioning.

The P1-plasmid has a defined minimal partition site ($parS_{P1}$) consisting of 34 base pairs (bp), which is called $parS_{P1}$ (Martin et al., 1987). This site contains a 13 bp inverted repeat and is sufficient to accurately direct the segregation of low copy number plasmids. An astonishing experiment demonstrated that this site could confer stability. The experiment involved a small target plasmid that held a 49-bp fragment containing a $parS_{P1}$ site, which on its own was intrinsically unstable. However, the presence of both ParA and $ParB_{P1}$ stabilized the plasmid. The fact that these three components alone were sufficient to ensure faithful plasmid segregation served as proof that the system is self-sufficient and does not require any additional components.

Around the same time, it was confirmed that the claims made by Hayakawa and Matsubara were valid, as $ParB_{P1}$ was shown to specifically bind to the $parS_{P1}$ site (Davis & Austin, 1988; Funnell, 1988).

Another substantial advance was the discovery that $ParA_{F\&P1}$ can bind ATP since it contained a type I nucleotide-binding motif (Motallebi-Veshareh et al., 1990). Further, *in vitro* experiments demonstrated that both the purified $ParA_F$ and $ParA_{P1}$ exhibited a weak basal hydrolytic activity. When both non-specific double-stranded DNA and ParB were supplemented in combination, the ATPase-activity of $ParA_{P1}$ was significantly enhanced, whereas using either ParB or non-specific double-stranded DNA alone only moderately stimulated the activity (Castaing et al., 2008; Davis et al., 1992; Watanabe et al., 1992).

As the field progressed, it was discovered that purified $ParB_{P1}$ forms dimers, and it became clear that this dimerization plays a crucial role in the accurate partitioning of plasmids (Funnell, 1991). In the same study, the authors determined the number of $ParB_{P1}$ dimers to be between 2,000 to 4,000 dimers per cell.

At the time, little was understood regarding the regulation of $ParA_{P1}$ and $ParB_{P1}$ expression. Upon closer examination, the genes $parA_{P1}$ and $parB_{P1}$ were found to be under the control of the same promoter and are subject to repression by $ParA_{P1}$, which binds to inverted repeats near the promoter (Davis et al., 1992; Hayes et al., 1994). This complicated future studies of ParA knockouts since this would lead to an overexpression of ParB.

Visual evidence for the partitioning of plasmids by the ParABS system was not obtained until more than 10 years after its discovery (Gordon et al., 1997). In this study, the authors introduced the lac-operator-target-array into both the F-plasmid and P1-plasmid and used GFP–LacI fluorescence to tag it. For the first time, the dynamics of plasmid motion within living cells were visualized. The authors showed that when there was only a single copy of the F-plasmid or P1-plasmid present within the cell, it was located at the center. After this plasmid replicated, both sister plasmids rapidly moved apart to the quarter positions, thereby ensuring their stability.

The discovery of more and more Par and Par-like systems led to the classification of the system into different types (Gerdes et al., 2000). Type I, to which all ParABS systems for plasmid partitioning belong, was subdivided into two subtypes, type Ia to which both P1 and F-plasmid belong and type Ib. The primary difference between these two subtypes is the size of their constituent proteins. ParA of type Ia are roughly 400 amino acids (aa) long, while ParA of type Ib are only half as big (~200aa). The relative size difference is even greater for ParB (type Ia: ~300aa, type Ib ~100). Type II systems are described in the previous section.

The next major finding came from a type Ib system, pB171. Using a fluorescent $ParA_{171}$ fusion, the authors showed that ParA is confined to the nucleoid and exhibits oscillatory movement within its boundaries (Ebersbach & Gerdes, 2001). This localization of $ParA_{171}$ suggests that $ParA_{171}$ has a high

affinity for DNA and that the majority of $ParA_{171}$ is in a nucleoid-bound state. The oscillatory behaviour could not be explained.

In vitro experiments have demonstrated that Soj, a member of the conserved ParA protein family involved in chromosome segregation in *Bacillus subtilis*, binds to DNA non-specifically in an ATP-dependent manner (Leonard et al., 2005). In the same study, the authors showed that Soj forms filaments in an ATP- and DNA-dependent manner. Subsequently, it was shown that both $ParA_F$ and $ParA_{171}$ also form filaments *in vivo*, this time only requiring ATP to do so (Ebersbach et al., 2006; G. E. Lim et al., 2005). These findings indicate that ParA has some cooperative capabilities beyond merely forming dimers, under the right conditions.

These discoveries gave rise to various hypotheses on how the mechanism of the ParABS system works, with some suggesting a potential similarity with the eukaryotic spindle apparatus (Ebersbach et al., 2006; Hatano et al., 2007). This mechanism proposed helical ParA filaments which oscillate in the cell from pole to pole while positioning plasmids. This mechanism was supported by *in vivo* observations of both oscillating and heterogeneous localizations of ParA-GFP, which resembled filamentous helices (G. E. Lim et al., 2005). However, the veracity of *in vivo* filaments remained unverified, and the observations of heterogeneity of ParA could potentially be attributed to the use of non-monomeric fluorophores.

Another study examined the crystal structures of $ParA_{P1}$ and found that it is made up of three domains: An elongated alpha-helix at the N-terminal end, which facilitates dimerization; a winged-HTH; and a C-domain containing a Walker-box (Dunham et al., 2009). This study showed that similarly to $ParB_{P1}$, $ParA_{P1}$ also forms dimers.

In this study (Vecchiarelli et al., 2010), the authors investigate $ParA_{P1}$'s DNA-binding process. When $ParA_{P1}$ binds ATP, it undergoes a slow, multi-step conformational transition, allowing it to non-specifically bind DNA. This mechanism serves as a time-delay switch, preventing instantaneous rebinding of $ParA_{P1}$ to DNA after detaching due to hydrolysis. The researchers propose that this time delay, combined with the stimulation of $ParA_{P1}$'s ATPase activity by $ParB_{P1}$ bound to plasmid DNA, results in an uneven distribution of nucleoid-bound $ParA_{P1}$. This uneven distribution generates the driving force for plasmid movement in the direction of the greatest $ParA_{P1}$ gradient. The authors refer to this proposed mechanism as the "diffusion-ratchet”.

In 2013 the ParABS system was reconstituted *in vitro* inside a DNA-carpeted flow cell, which acted as an artificial nucleoid (Hwang et al., 2013). The DNA-carpet inside the flow cell was pre-incubated with $ParA_F$-ATP until it saturated and no more ParA were able to bind to the DNA-carpet. In the subsequent step, $ParB_F$-coated plasmids were introduced into the flow cell to investigate their

interactions with the carpet using total internal reflection fluorescence microscopy. Notably, no proof was found for the development of filaments, and instead, ParA and ParB briefly anchored plasmids to the carpet (Hwang et al., 2013). Moreover, DNA-bound $ParA_F$ was found to undergo transient unbinding, which allowed it to diffuse on the DNA carpet. This was demonstrated by placing two DNA-carpets (one preincubated with $ParA_F$) in close proximity to each other. Remarkably, when researchers placed the plasmids with magnetic beads, coated with *parS*$_F$-DNA/$ParB_F$, on the carpet and confined those beads to the DNA-carpet using a magnetic field, the beads exhibited a period of immobilization followed by directed motion, leaving zones depleted of $ParA_F$ in their wake (Vecchiarelli et al., 2014). These observations further support their previously proposed diffusion-ratchet mechanism.

A study examining the role of $ParA_F$ in the partition of the F-plasmid found that the frequency of $ParA_F$ oscillations was affected by the interaction between ParB and ParA (Ah-Seng et al., 2013). Surprisingly, the authors found that $ParA_F$ oscillated faster when its ATPase activity was less stimulated by $ParB_F$-mutants (compared to wild type), which is counterintuitive because less energy consumption/hydrolysis should “normally” result in slower oscillations. Interestingly, the loss rate of plasmids was not substantially affected by changes in frequency/$ParB_F$ mutants, which again raised questions about the nature and purpose of $ParA_F$ oscillations.

Around the same time, a model was developed that not only set the stage for all subsequent models but also revolutionized our understanding of positioning in general (Sugawara & Kaneko, 2011). The model proposes a mechanism in which a gradient of a substrate generates a "chemophoretic" pulling force on a cargo. This pulling force provides the energy and information to position the cargo. Although the model does not specify a concrete molecular mechanism, it still offered an explanation of regular positioning to the observations made both *in vivo* and *in vitro*: The movement and positioning of plasmids are based on the flux of incoming ParA.

Building upon this concept, several models emerged. Among the earliest was a model proposed by Ietswaart et al. (Ietswaart et al., 2014). This stochastic model proposes that nucleoid-bound ParA, in their ATP-bound state, form filaments that subsequently attach to ParB-coated plasmids. Once attached, ParB on the plasmid starts hydrolyzing the ParA it is connected to and converting it to its ADP-bound state, causing its detachment from the nucleoid and release into the cytosol. This process results in the shortening of the filament, causing the plasmid to be pulled along. ParAs which are released into the cytosol are now in their ADP-bound state and have to remain in this state for a certain duration of time until they are converted back to their ATP-bound state, allowing them once again to bind to the nucleoid. This model is capable of producing regular positioning. The authors attributed this behaviour to the observation that only plasmids at regular positions maintain a constant flux of incoming ParA filaments from each side.

In the same year the DNA-relay model was published. It proposed that the force to move the plasmid is generated by the elastic fluctuations of the nucleoid (Lim et al., 2014; Surovtsev et al., 2016). Specifically, this molecular-level model assumes that ParA binds to the nucleoid in its ATP-bound state and neither diffuses on the nucleoid nor forms any polymers. After binding to the DNA, ParAs undergo elastic fluctuations around an equilibrium position similar to what was observed for DNA-loci (Lim et al., 2014). This mechanism allows ParA to explore the space surrounding its equilibrium position, and when in close proximity to a ParB-coated plasmid, ParA and ParB instantly form a tether, anchoring the plasmid to the nucleoid. The tether exerts force on the plasmid, trying to pull it towards its equilibrium position. While in a tether, ParB stimulates the hydrolysis activity of ParA and once the ATP of the ParA hydrolyzes, the tether disassembles and the ParA is released into the cytosol. Similar to the model of Ietwaart et al. it can rebind the nucleoid in its ATP-bound state after a waiting period in the cytosol. Another notable addition was the incorporation of a basal hydrolysis rate of ParA, allowing it to detach from the nucleoid without interacting with ParB. In this model, plasmids move in the direction of the highest gradient of ParA, leaving behind a depleted zone of ParA. This causes directed movement of the plasmid until it reaches the edge of the nucleoid, at which point it stops and moves in the opposite direction. Consequently, a single plasmid moves from pole to pole. Remarkably, plasmids repel each other by recruiting/tethering ParA away from neighboring plasmids. In simulations, this repulsion prevented plasmids from crossing each other along the long axis of the cell, confining each plasmid to a narrow region on the nucleoid. As a time average, this appeared similar to regular positioning in cells that contain more than one plasmid.

In the following years, the cell-centric version of the Brownian-ratchet model was published (Hu et al., 2017b, 2017a). This model was an iteration of a previous model which was based on the observations from the DNA-carpeted flow cell experiment and it did not contain any cell boundaries (Hu et al., 2015). While the newest iteration of the Brownian-ratchet model shares the same underlying concept of force generation as the DNA-relay model, there are some notable differences. Specifically, it is more intricate, defining most of the molecular mechanisms in more detail, while still assuming that the force needed to move the plasmid is generated by chromosomal fluctuations, the Brownian-ratchet model included several additional details: Tethers do not form instantly but at a very high rate; tethers break if they are too far extended from their equilibrium position; the lifetime of a tether changes based on its extension; tethers have an equilibrium length equal to the postulated distance between the plasmid and nucleoid; ParA can diffuse on the nucleoid; after the ATP of nucleoid-bound ParA is hydrolyzed and before it is released into the cytosol, there is a transient nucleoid-bound ADP state; there is a finite number of ParB on the plasmid. The inclusion of these details did not provide any significant new insights into the ParABS system, it only made it more difficult to identify the impactful parameters, like the inclusion of ParA diffusion on the nucleoid. This allowed the model to produce a plethora of different behaviours, as shown in this thesis, but the

authors did not identify this parameter as one of the most important parameters of their model. Another downside to the Brownian-ratchet model is that the number of ParA was not explicitly implemented. There is an infinite cytosolic pool and, based on a refilling rate (how fast cytosolic ParA binds to the nucleoid), different concentrations of ParA on the nucleoid were attained. The refilling rate they chose led to an unreasonably high number of nucleoid-associated ParA dimer (67000), while studies estimate the actual amount of ParA at least one magnitude less. Irrespective of these downsides, the authors showed, using their model, that plasmids can move in different modes: diffusive, local excursion/regular positioning (the latter only for higher numbers of plasmids), oscillatory and static (plasmids do not move). They achieved these various modes by changing the lifetime of the ParA-ParB tethers and the refilling rate.

Both the DNA-relay and Brownian-ratchet models have been used to study plasmid dynamics, with little emphasis on ParA. Both models produced oscillating ParA, similar to what was observed *in vivo*, only when a single oscillating plasmid was present. However, the literature has reported multiple instances of oscillating ParA even at higher plasmid numbers. The observed discrepancy regarding the oscillations of ParA raises the question of their purpose in the absence of any functional requirement.

During the same period, another study cast doubt on the existence of ParA filaments *in vivo*. Using super-resolution microscopy techniques to localize all components in 3D, researchers found that the ParABS system's components reside within the nucleoid interior (Le Gall et al., 2016). No evidence of filaments was found. Instead, they found that ParA colocalized with regions of high DNA-density within the nucleoid. Additionally, they visualized the effect of $ParA_F$ mutants, which either cannot bind to DNA or are unable to hydrolyze ATP, on the system. The researchers proposed an alternative mechanism for plasmid positioning and segregation, wherein plasmids are attracted to these high-density regions within the nucleoid through interactions facilitated by the $ParA_F$. In this manner, plasmids effectively hitchhike on replicated nucleoids, ensuring faithful segregation to the daughter cells.

In 2019, another long-standing question of how *parS* sites are able to load hundreds of ParB dimers onto the plasmid was answered. It was discovered that ParB is capable of binding and hydrolyzing CTP, which enables it to slide along the DNA (Osorio-Valeriano et al., 2019; Taylor et al., 2021). This process occurs in a specific chain of events: A ParB dimer initially binds to a *parS* site; the dimer then functions as a clamp in a CTP-dependent manner, closing and forming a ring around the DNA; ParB rings exhibit a decreased affinity for *parS* sites in this closed state, ergo they are able to move along the DNA strand, thereby freeing up the *parS* site (Soh et al., 2019). Once CTP is hydrolyzed, ParB dissociates from the DNA and returns to the cytosol. This mechanism allows one *parS* site to load multiple ParB, and it is hypothesized that CTP acts as a timing mechanism to regulate how far ParB spreads along the DNA.

Returning to $ParA_F$, evidence was found that three $ParA_F$ dimers cooperatively bind to the promoter region (Boudsocq et al., 2021). There is a single perfect inverted repeat (IR) motif, consisting of two hexamer sequences spaced by 28-bp, that serves as the primary $ParA_F$ DNA binding site. A similar but degenerated motif overlaps the primary motif. A winged-HTH domain of $ParA_F$ may facilitate cooperative $ParA_F$ binding to these three regions and allow $ParA_F$ to control the expression of both $ParB_F$ and itself. The idea that ParA exhibits cooperative interactions with itself is not unlikely, given that earlier *in vitro* investigations demonstrated that $ParA_F$ is capable of forming filaments at high concentrations, thus providing support for $ParA_F$'s cooperative nature.

Further, the investigation of the C-terminal of $ParA_F$ showed that several residues in the final helix of $ParA_F$ play a critical role in the maintenance of plasmids (Mishra et al., 2022). Although $ParA_F$ can still bind to the nucleoid and interact with $ParB_F$ when the last five residues are removed, it has a significant impact on the partitioning of plasmids. Moreover, specific mutations in residues within the final C-terminal helix only slightly affect the localization of $ParA_F$ to the nucleoid, but others completely abolish nsDNA binding and disrupt plasmid maintenance functions.

In the same year, our "Hopping and Relay" model was published (Köhler et al., 2022), building upon the foundation of the DNA-relay model, integrating basal hydrolysis and ParA diffusion on the nucleoid. This publication is included in a later chapter of this thesis.

A recent publication tackled the question of how ParB can bypass roadblocks such as actively expressed genes (Tišma et al., 2022). The study found that DNA-bound ParB is capable of recruiting cytosolic ParB to DNA strands in close proximity. This allows ParB to spread to regions separated by roadblocks from the *parS* site

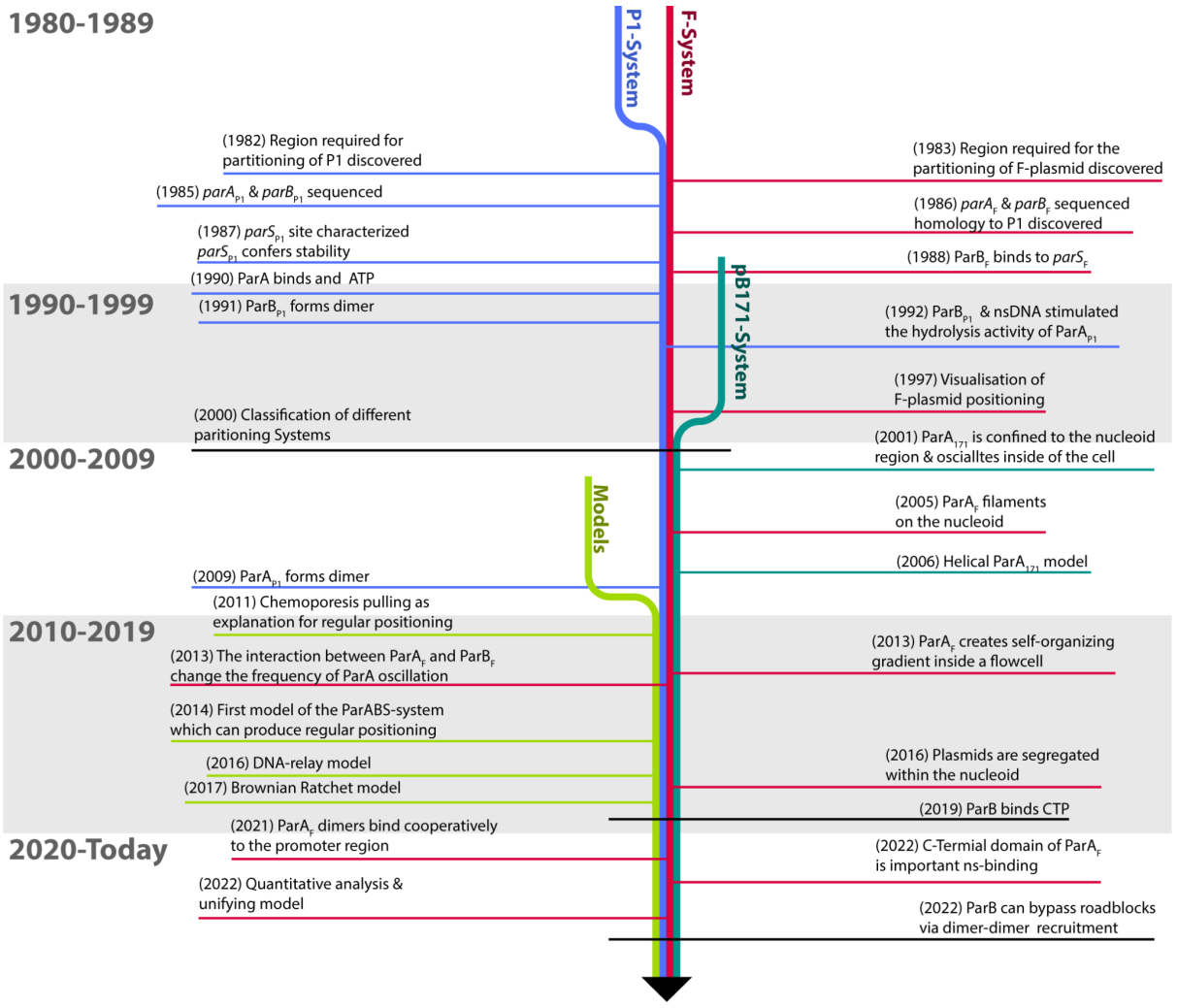

Figure 2. Timeline of the ParABS system. The P1-system findings are denoted in blue, F-system in pink, pB171-system in dark green, and other systems in black. Models denoted in light green.

2.2. Related Systems

Exploring systems related to ParABS may offer valuable insights and a deeper understanding of its underlying principles.

2.2.1. Divisome positioning: MinCDE

The MinCDE system is composed of three components and is required to ensure proper positioning of the bacterial cell division site by regulating the assembly of the divisome at the mid-cell. It is composed of three components (Ramm et al., 2019): MinD, a Walker-type ATPase, dimerizes, binds to the cell membrane, and interacts with MinC and MinE to localize the FtsZ ring at mid-cell; MinC, which binds to MinD and once bound, inhibits the formation of the FtsZ ring; MinE, which stimulates the ATPase activity of MinD, causing MinD to detach from the membrane once the ATP is

hydrolyzed. FtsZ is essential for recruiting other proteins that produce a functional septum at the division site. In the absence or dysfunction of the MinCDE system, an increased number of anucleate cells are observed (de Boer et al., 1989).

The cell membrane acts as a binding site for MinD, which localizes to one cell pole and, once membrane bound, polymerizes, resulting in growing clusters of MinD. This cluster is then disassembled by MinE, which stimulates MinD's ATPase activity and releases it back into the cytosol. MinD then traverses to the other cell pole, where the cycle repeats. This causes all components of the system to oscillate inside of the cell. Over time, this results in an intracellular gradient of MinC and MinD, peaking at the poles and reaching its nadir at the mid-cell. Since MinC inhibits the formation of the FtsZ ring, the FtsZ ring forms where the concentration of MinC is the lowest, which is at the mid-cell. This allows a symmetric division of the cell.

It has been inferred that some form of MinD-MinD recruitment (membrane-bound MinD helping cytosolic MinD to bind in close proximity) is required for the mechanism to work (Halatek & Frey, 2012; Huang et al., 2003).

2.2.2. Carboxysomes positioning in Cyanobacteria

Carboxysomes are protein-based bacterial organelles that encapsulate key enzymes of the Calvin-Benson-Bassham cycle, playing a vital role in carbon fixation and photosynthesis (Turmo et al., 2017). Cyanobacteria, in particular, harness solar energy to power their cells, relying on carboxysomes for efficient energy production. Since there are only a handful of carboxysomes per cell, they need an active partitioning system, the McdAB system (MacCready et al., 2018; MacCready & Vecchiarelli, 2021).

Carboxysomes are positioned at regular intervals along the long axis of the cell, ensuring equal distribution to both daughter cells upon division. The McdA protein oscillates from one end of the cell to the other, guaranteeing proper positioning of carboxysomes. McdB, a partner for McdA, is present on the surface of carboxysomes and binds to McdA, which in turn attaches to the nucleoid. This interaction, similar to the ParABS system, leads to the formation of tethers between the nucleoid and cargo and causes McdA to detach from the nucleoid faster.

Carboxysomes constantly seek the highest concentrations of McdA bound to nearby DNA, further emphasizing the resemblance to the ParABS system. The McdA and McdB proteins organize themselves using the nucleoid as a scaffold, which subsequently results in the even distribution of carboxysomes along the long axis of the cell. The remarkable resemblance to the ParABS system becomes even more apparent when investigating genetic modifications of McdA and McdB. Deletion

of either protein leads to carboxysomes clustering at the poles, a phenotype similar to what is observed in the ParABS system.

In summary, the remarkable parallels between the ParABS system and carboxysome distribution mechanisms underscore the existence of unifying strategies utilized across various cellular processes for positioning cargo, which is not limited to DNA.

2.2.3. Chemosensory array positioning

Bacterial chemosensory arrays are large hexagonal structures that play an essential role in chemotactic signaling in bacteria like *Rhodobacter sphaeroides* (Yang & Briegel, 2020). These arrays are necessary for bacteria to sense and respond to changes in their chemical environment, and to move towards or away from specific chemical signals, a process known as chemotaxis.

In *R. sphaeroides*, two spatially distinct chemosensory arrays, one transmembrane and the other cytoplasmic, control a single flagellum. The number and positioning of cytosolic arrays are controlled by two proteins sharing similarities with ParA and ParB of the ParABS system (Thompson et al., 2006): TlpT, a chemoreceptor, and PpfA, an ATPase. Typically, either one array (located at mid-cell) or two arrays (positioned at quarter positions) are present in a cell. TlpT, a chemoreceptor (Jones & Armitage, 2017), is crucial for cytoplasmic array formation, as its deletion leads to a loss of cytoplasmic array formation. Conversely, PpfA, a ParA homologue (Roberts et al., 2012), is necessary for positioning and segregating the cytoplasmic array; deleting PpfA results in slower and fewer arrays forming. The number of arrays formed depends on cell length, suggesting spatial limitations. Newly formed arrays are believed to be positioned through stochastic assembly (Thiem & Sourjik, 2008) rather than active positioning, as they form at their target positions without moving away from them.

Filamentous cells elongated by cephalexin treatment also exhibit a dependency on cell length for the number of arrays formed, with longer cells forming more arrays and PpfA positively influencing the number of arrays formed at a given cell length. In conclusion, TlpT and PpfA are vital components in the formation, positioning, and segregation of cytosolic chemosensory arrays, ensuring proper assembly and functionality, which is critical for chemotactic signaling in bacteria.

2.3. Important features of the ParABS system

This section contains a brief summary of the important features relevant to this thesis of the type Ia ParABS system. Both as a graphic and as a listing:

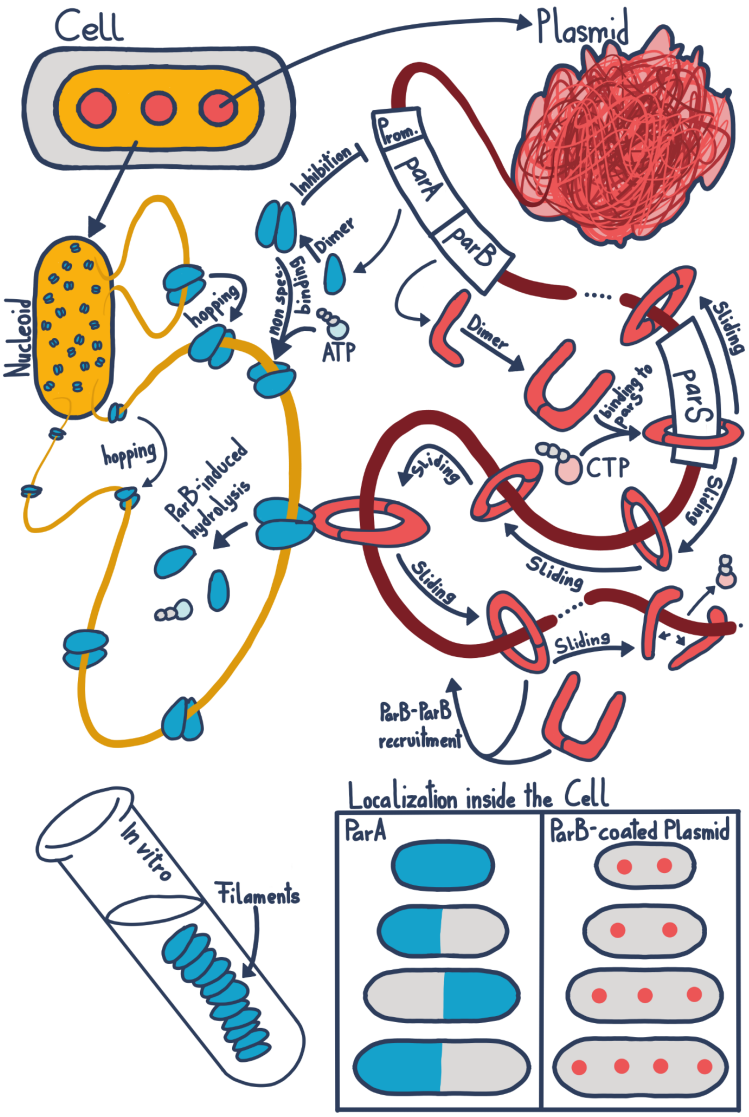

Figure 2: Cartoon of the ParABS system. **Top left:** cell with 3 plasmids (red) and nucleoid (yellow). Top right: plasmid with a ParABS system. The par-operon contains the genes *parA* and *parB* from which ParA (blue) and ParB (red) are expressed. **Left side:** ParA forms dimers, and these dimers auto-regulate the expression of both ParA and ParB. ParA binds ATP and, in this state, is competent to bind DNA non-specifically. While DNA-bound, it can transiently unbind and hop between DNA strands. The majority of ParA is bound to the nucleoid. ParA can fall off the DNA through basal hydrolysis or ParB-induced hydrolysis. ParA forms filaments *in vivo* in the presence of ATP. **Right side:** ParB also forms dimers, and in its dimeric state, it binds to *parS*. Furthermore, DNA-bound ParB dimers slide along the DNA strand (type Ia systems require CTP), freeing up the *parS* site so more ParB can bind to it. DNA-bound ParB stimulates the ATPase activity of DNA-bound ParA. In its DNA-bound state, ParB recruits cytosolic ParB dimers to bind to the DNA in close proximity. ParB can slide several kilobases away from the *parS* site before the dimer falls apart. **Bottom right:** Localization of ParA and ParB-coated plasmids inside the cell. ParA is either uniformly distributed or present in either half of the cell. ParB-coated plasmids are regularly positioned along the long axis of the cell.

(1) The ParABS system is an active partitioning system for low copy number plasmids, ensuring their stability and it is found on multiple different plasmids including the F-plasmid, P1-plasmid and the pB171-plasmid (Austin et al., 1982; Ebersbach & Gerdes, 2001; Mori et al., 1986).

(2) This system positions plasmids regularly along the cell's long axis, ensuring even distribution during cell division (Erdmann et al., 1999; Gordon et al., 1997; Köhler et al., 2022).

(3) This system consists of a binding site named *parS* and two proteins, ParA and ParB.

(4) ParA is a walker-type ATPase that forms dimers upon binding ATP (Motallebi-Veshareh et al., 1990).

(5) In its dimeric state, ParA can bind DNA non-specifically (Davis et al., 1992; Motallebi-Veshareh et al., 1990; Watanabe et al., 1992).

(6) After ATP hydrolysis, ParA detaches from the DNA.

(7) Most of the ParA in a cell is nucleoid-bound (Ebersbach & Gerdes, 2001) and can diffuse on the nucleoid through transient unbinding and hopping between DNA strands (Hwang et al., 2013).

(8) Nucleoid-bound ParA exhibits oscillatory movement, collectively transitioning between cell halves (Ah-Seng et al., 2013; Ebersbach & Gerdes, 2001).

(9) ParB form dimers (Funnell, 1991).

(10) In its dimeric, CTP-bound state, ParB can bind to *parS* (Osorio-Valeriano et al., 2019; Taylor et al., 2021).

(11) The *parS* site is located on the plasmid.

(12) Once bound to DNA, ParB can slide along the DNA strand, enabling significant loading of ParB onto the plasmid (Soh et al., 2019).

(13) After CTP hydrolysis, ParB detaches from the DNA.

(14) Nucleoid-bound ParA and plasmid-bound ParB form tethers, anchoring the plasmid to the nucleoid (Hwang et al., 2013).

(15) While interacting, ParB stimulates ParA's hydrolysis activity, causing it to release faster from the nucleoid (Davis et al., 1992; Watanabe et al., 1992).

(16) The DNA-relay model is the first to propose that plasmid movement is powered by chromosomal fluctuations of the plasmid, harnessed through ParA-ParB tethers (H. C. Lim et al., 2014; Surovtsev et al., 2016).

3. Paper I

The first paper of this thesis investigates the underlying mechanisms of plasmid positioning of the ParABS system. Previous studies have identified two distinct modes of plasmid movement within cells: consistent positioning and oscillatory motion. However, which is the dominant mode of movement has remained elusive due to a scarcity of quantitative data. In this paper, we employ a microfluidic device known as the "mother machine" to analyze thousands of cell cycles. By integrating this quantitative data with computational modeling, we unveil a single, unifying model capable of reconciling these discrepancies of past observations and illuminating the true dynamics of plasmid movement within cells.

My contribution to this paper includes all aspects of this paper apart from the layout of the microfluidic device used which was developed by Seán M. Murray & Eugen Kaganovitch and the writing which was done by Seán M. Murray and myself.

This paper was published in eLife on the 14th of november 2022: https://doi.org/10.7554/eLife.78743

3.1. High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS

Robin Köhler, Eugen Kaganovitch and Seán M. Murray*

Max Planck Institute for Terrestrial Microbiology and LOEWE Centre for Synthetic Microbiology (SYNMIKRO), Karl-von-Frisch Straße 14, 35043 Marburg, Germany

*Email: sean.murray@synmikro.mpi-marburg.mpg.de

3.1.1. Abstract

The faithful segregation and inheritance of bacterial chromosomes and low-copy number plasmids requires dedicated partitioning systems. The most common of these, ParABS, consists of ParA, a DNA-binding ATPase and ParB, a protein that binds to centromeric-like *parS* sequences on the DNA cargo. The resulting nucleoprotein complexes are believed to move up a self-generated gradient of nucleoid-associated ParA. However, it remains unclear how this leads to the observed cargo positioning and dynamics. In particular, the evaluation of models of plasmid positioning has been hindered by the lack of quantitative measurements of plasmid dynamics. Here, we use high-throughput imaging, analysis and modelling to determine the dynamical nature of these systems. We find that F plasmid is actively brought to specific subcellular home positions within the cell with dynamics akin to an over-damped spring. We develop a unified stochastic model that quantitatively explains this behaviour and predicts that cells with the lowest plasmid concentration transition to oscillatory dynamics. We confirm this prediction for F plasmid as well as a distantly-related ParABS system. Our results indicate that ParABS regularly positions plasmids across the nucleoid but operates just below the threshold of an oscillatory instability, which according to our model, minimises ATP consumption. Our work also clarifies how various plasmid dynamics are achievable in a single unified stochastic model. Overall, this work uncovers the dynamical nature of plasmid positioning by ParABS and provides insights relevant for chromosome-based systems.

3.1.2. Introduction

To ensure that their genetic material is faithfully partitioned to daughter cells upon cell division, low-copy plasmids and bacteria employ dedicated partitioning (*par*) systems, of which ParABS is the most common (Badrinarayanan et al., 2015; Kawalek et al., 2020). The system consists of three components: i) a centromeric-like region *parS*, ii) a Walker-type ATPase ParA and iii) the protein ParB. ParA binds DNA non-specifically in its ATP-dependent dimer state and therefore coats the nucleoid. ParB dimers bind to, and spread out several kilobases from, consensus sequences within *parS* to form a condensed nucleoprotein complex, called the partition complex, that is clearly visible using fluorescent microscopy. ParB also stimulates the ATPase activity of ParA releasing it from the nucleoid and generating a gradient of bound ParA around it. Partition complexes are believed to move up this gradient mediated by ParB-ParA bonds connecting the cargo (e.g. a plasmid or chromosomal origin) and the underlying nucleoid. In particular, the most recent molecular-scale models argue that the elastic fluctuations of the chromosome and/or ParA-ParB tethers power the directed movement of the plasmid cargo, while the self-generated ParA gradient specifies the direction (Hu et al., 2015; Lim et al., 2014).

ParABS systems fall into three main types based on their phylogeny (Gerdes et al., 2000). Types 1a and 1b are found on plasmids and are distinguished by their genetic organisation and the size of their genes. In particular, type 1b systems encode much smaller ParA and ParB proteins than their type 1a counterparts and have a broader host range, being found in both Gram-negative and Gram-positive bacteria. The third type is a diverse family consisting of chromosomal ParABS systems. In terms of size, they are similar to type 1a but their ParA sequences form a separate phylogenetic cluster. Perhaps the most significant distinction among the types comes from the recent result that ParB from F plasmid (a type 1a system) and several bacterial species are CTPases. ParB dimers form a DNA clamp that loads onto the DNA at *parS* sites before sliding (diffusing) along the DNA in a CTP-binding dependent manner (Osorio-Valeriano et al., 2019; Soh et al., 2019). While this has clarified the mechanism of ParB spreading, the role of its CTPase activity in partition complex formation and positioning remains unclear. Furthermore, the much smaller ParB of the type 1b systems do not have the same CTP binding pocket, yet still confer stability to low-copy plasmids.

ParABS systems function by segregating and positioning their DNA cargo to specific positions within the cell. Typically, cargos are located symmetrically and at equally-spaced intervals across the the

nucleoid i.e. at the mid, quarter or (1/6, 3/6, 5/6) positions for one, two or three cargos respectively[1]. This pattern of 'home' positions is known as regular positioning and has also been observed in related ParA-like systems that position non-DNA cargo (MacCready et al., 2018; Roberts et al., 2012; Schumacher et al., 2017). However, the dynamics of plasmid positioning have not been characterised. So it is unclear whether the observed position distributions arise through true regular positioning in which the plasmid 'senses' the geometry of the nucleoid and positions itself accordingly (subject to stochastic variation) or through a more approximate mechanism. Indeed, both F plasmid and pB171 have been described as exhibiting oscillatory dynamics as they follow corresponding changes in the ParA gradient, which may also lead to regular positioning as a time-averaged effect (Hatano et al., 2007; Ringgaard et al., 2009; Surovtsev et al., 2016a). While there have been several modelling studies of plasmid positioning (Adachi et al., 2006; Ietswaart et al., 2014; Jindal and Emberly, 2019; Ringgaard et al., 2009; Sugawara and Kaneko, 2011; Walter et al., 2017) and, in particular, two recent stochastic models that incorporate the molecular mechanism of force generation (Hu et al., 2017; Surovtsev et al., 2016a), the lack of quantitative measurements of plasmid dynamics has hindered their evaluation. This is especially important as it may be challenging to distinguish noisy true positioning from approximate positioning or noisy low-amplitude oscillations.

Here, we uncover the nature of plasmid dynamics and positioning through a combination of high-throughput imaging and analysis and comparison to a minimal molecular-level computational model. We find unambiguously that the type 1a F plasmid exhibits true regular positioning as if pulled to its home positions by an over-damped spring-like force and we quantitatively reproduce its positioning and segregation behaviour in a unifying stochastic model. Furthermore, our model, an extension of the previous DNA relay model (Surovtsev et al., 2016a), suggests that the fraction of the nucleoid that each ParA-ATP dimer explores during its lifetime is a critical determinant of the dynamics and we confirm its prediction that single plasmids in longer cells transition to oscillatory dynamics. We also identify the ratio of the ParB-induced and the basal rate of ATP hydrolysis by ParA as a second critical model parameter. Together, these two parameters map out the entire space of plasmid dynamics and give a physical understanding of all possible dynamics including oscillations, regular positioning, static, diffusive, as well as whether the ParA distribution has a maximum or minimum at the cargo. Thus, our model, though similarly based on elastic chromosome fluctuations, unifies the existing molecular-level stochastic models (Hu et al., 2017; Surovtsev et al., 2016a) by producing all the various possible plasmid dynamics in a single model. Finally, we examine the type 1b system of pB171 and find clearer oscillatory dynamics but again dependent on the number of plasmids and cell length. Our results show that both F plasmid and pB171 operate just below the

[1] In the case of n cargos, their relative positions are (i-1/2)/n for i = 1, 2, …, n.

threshold for oscillations to occur, with pB171 crossing the threshold in cells containing a single plasmid and F plasmid doing so only in cells that are additionally longer than average. Overall, our work resolves the nature of plasmid positioning and dynamics by ParABS and presents a unified stochastic model that explains the full range of behaviours in terms of well-defined system properties.

3.1.3. Results

3.1.3.1. The F plasmid is regularly positioned by a spring-like force

To clarify the nature of plasmid dynamics, we turned to a high-throughput microfluidics approach based on a custom-fabricated ‘mother machine’ device coupled with a segmentation, tracking and foci detection pipeline (Figure 1A,B). Using this approach, we tracked, at 1 min resolution, the dynamics of mini-F plasmids during many thousands of cell cycles using a fully functional ParB-mVenus fusion (Sanchez et al., 2015). Under our conditions, cells had a median of two ParB foci at birth and four at division (Figure 1–figure supplement 1) and divided approximately every 100 min. Since ParB foci separate within 5 min of plasmid replication (Onogi et al., 2002; Walter et al., 2020) and there are only a few replication events per cell cycle, in the following we will assume each ParB focus consists of a single plasmid.

Consistent with many previous works (Adachi et al., 2006; Hatano et al., 2007; Niki and Hiraga, 1997; Sanchez et al., 2015), we found that F plasmid is, irrespective of length, approximately located at mid-cell in cells with a single plasmid and close to the quarter positions in cells with two plasmids (Figure 1C,D). In the latter case, their positions have been more accurately specified as the quarter positions of the nucleoid (Le Gall et al., 2016). We found that the precision of positioning for the single plasmid case was independent of cell length, while for two plasmids it decreased weakly for cell lengths greater than 3 μm, perhaps due to variation in nucleoid segregation (Figure 1–figure supplement 1).

While the average position of plasmids was unambiguous, the nature of the positioning dynamics was not. In particular, it was unclear whether plasmids were consistently biased towards their average positions (‘true positioning’) or if they exhibited diffusive or oscillatory motion within a confined area around these positions (‘approximate positioning’). Note that we are not referring here to the stochastic noisiness of positioning but rather to the nature of the positioning itself (a system with true

positioning may still be noisy). As discussed above, oscillations, typically of ParA but also of the plasmid itself, have been suggested to underlie positioning in ParABS systems (Hatano et al., 2007; Ringgaard et al., 2009; Surovtsev et al., 2016a). In this direction, we observed, albeit very infrequently, oscillatory-like back-and-forth plasmid movements, reminiscent of some previous observations of F plasmid (Hatano et al., 2007). We will return to this below.

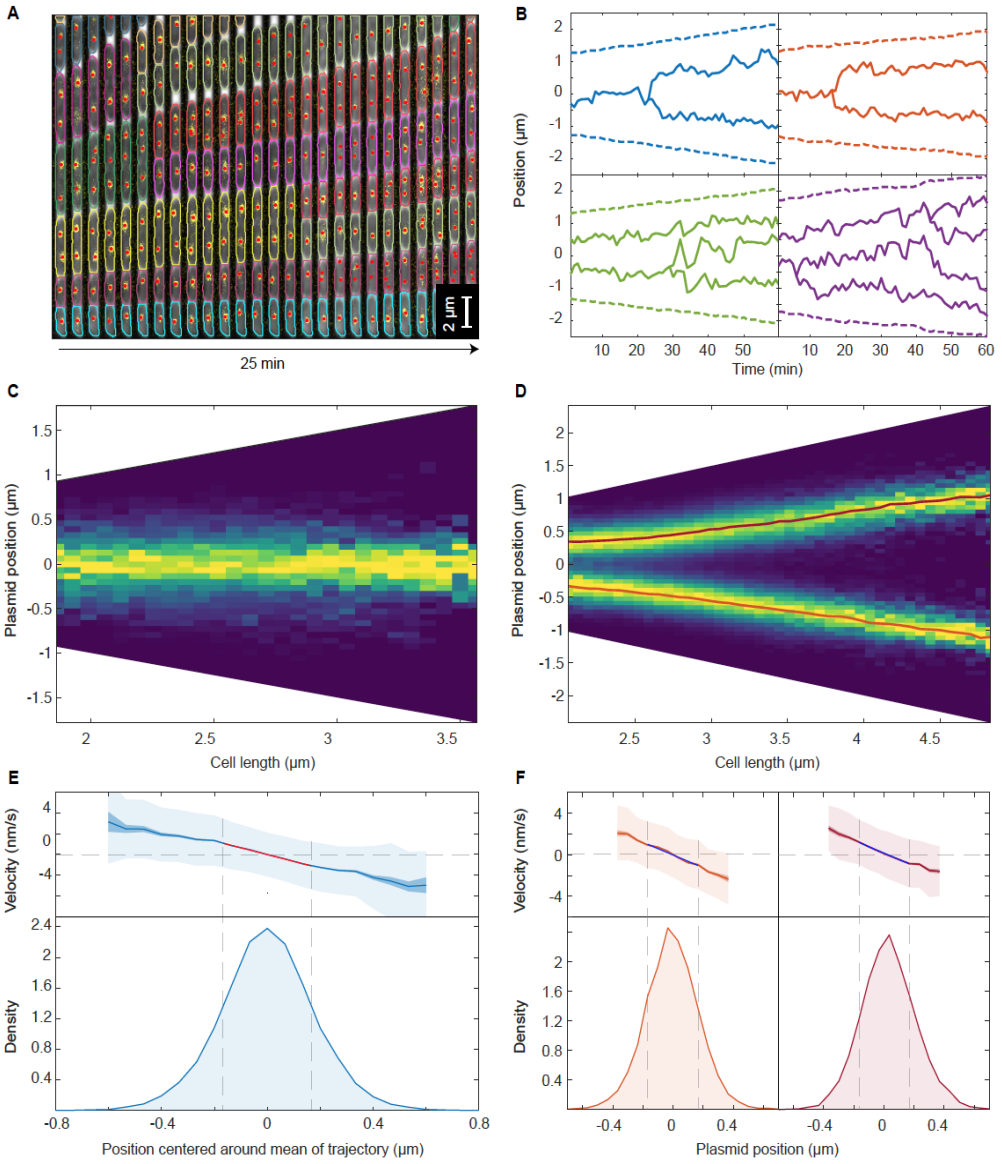

Figure 1. F plasmid exhibits true regularly positioning along the long axis of the cell (A) Timelapse of a single mother machine growth channel (*E. coli* strain DLT3125, phase contrast overlayed with ParB-mVenus fluorescence signal). Segmentation and tracking is indicated by coloured outlines. ParB-mVenus foci are highlighted by red dots. Time interval is 1 min. (B) Four example trajectories of tracked ParB-mVenus foci from different cells. Dashed lines indicate cell boundaries. (C) Distribution of foci positions as a function of cell length in cells containing one ParB-mVenus focus. Data from 879 cell cycles. (D) As in (C) but for cells containing two ParB-mVenus foci. Data from 5044 cell cycles. Red lines indicate the position of each peak as obtained by fitting to the sum of two Gaussian functions. (E) Top: Mean velocity of plasmids as a function of position relative to the trajectory mean in cells containing one plasmid. The velocity is measured over two consecutive frames, taken 1 min apart. Light and dark shading indicate standard deviation and standard error respectively. The red line indicates a linear fit. Note that the standard deviation of the velocity does not depend on position. Bottom: Probability density of plasmid position relative to mean of trajectory. Standard deviation 0.182 µm. Dashed lines indicate the region used for fitting which includes at least 68.27% of all data points. (F) As in (E) but for cells containing two ParB-mVenus foci and the position is relative to the indicated lines in (D). Standard deviations are 0.175 µm (old pole proximal) and 0.181 µm (new pole proximal). In (B-F), positions and velocities are measured along the

long cell axis. Position values are negative towards the old pole. See also Figure 1–figure supplement 1 and Figure 1–figure supplement 2.

To quantitatively examine the nature of plasmid positioning, we first measured the spatial dependence of plasmid velocity (measured between two consecutive frames) as a function of long-axis position within the cell. Analysing cells containing a single plasmid, we found a clear linear dependence of the mean of the velocity on position, while its variance was constant (Figure 1E). Furthermore, the position and velocity autocorrelation functions showed no population level evidence of oscillatory behaviour (Figure 1–figure supplement 2). Rather, the velocity autocorrelation was negative at a lag equal to the sampling time, a characteristic of elastic motion. We also analysed the trajectories of cells containing two plasmids. We found a similar linear dependence of the mean velocity around the mean positions (Figure 1F) and no evidence of oscillatory behaviour (Figure 1–figure supplement 2).

These results demonstrate that F plasmid exhibits true positioning. If this was not the case, we would expect a flattening of the velocity profile around the target positions and/or evidence of oscillations in the auto-correlation of position or velocity. Altogether the observed properties are characteristic of an over-damped spring-like force, similar to that observed for the chromosomal origins of *E. coli* (Hofmann et al., 2019; Kuwada et al., 2013). Under this model, the slope of the velocity profile is the reciprocal of the characteristic timescale, τ, at which elastic fluctuations act and we found this to be about 2 mins. Comparable values, given the 1 min frame rate, were found by fitting to the position and velocity autocorrelation (Figure 1–figure supplement 2A,B). On timescales much shorter than this, plasmid dynamics are dominated by diffusion, whereas on longer timescales, the effective spring-like force dominates. As our temporal resolution is on the same order as τ, we can obtain estimates, under the over-damped spring model, for both the intrinsic diffusion coefficient of the plasmid D and the spring constant k_{eff} of the effective force by fitting to the mean and variance of the velocity profile (see methods). We find $D=(2.27\pm0.24) \times 10^{-4}$ $\mu m^2 s^{-1}$ and $k_{eff}/(k_B T)=36.8\pm4.1$ μm^{-2} (bounds are the 95% confidence intervals). The latter implies a characteristic force of about 0.02 pN acting on the plasmid. Note that this estimate of the diffusion coefficient is not necessarily that of a plasmid lacking the ParABS system but rather describes the diffusive component of the dynamics in the presence of the system. To test this estimate, we tracked plasmid dynamics on a much shorter time-scale (1s frame rate) at which diffusion is expected to dominate and measured the mean square displacement (MSD) of the plasmid. Unlike at the longer timescale, we found a linear dependence on time, and a diffusion coefficient of $(2.01\pm0.14) \times 10^{-4}$ $\mu m^2 s^{-1}$ consistent with, and in support of, the over-damped spring model (Figure 1–figure supplement 2F,G).

3.1.3.2. Hopping of ParA-ATP on the nucleoid as an explanation of regular positioning

There have been two stochastic molecular-level models of plasmid positioning to date. Though different in some details, both models propose that elastic fluctuations of DNA and/or protein bonds power movement of plasmids up a gradient of DNA bound ParA-ATP. However, neither model exhibits true regular positioning as we observed for F plasmid. In the DNA relay model (Surovtsev et al., 2016a), plasmids oscillate across the nucleoid, reversing direction upon reaching either a pole or another plasmid, with regular positioning emerging from these oscillations only as a time-averaging effect. In the Brownian Ratchet model (Hu et al., 2021, 2017) on the other hand, plasmids exhibit 'local excursions' around home positions that are determined by the distance they segregate upon replication. In a narrow region of parameter space, this scheme leads to equi-positioning rather than regular positioning, i.e. plasmids maintain a particular inter-plasmid spacing along the long axis of the nucleoid rather than being positioned at particular locations. Given that previous coarse-grained models have displayed regular positioning (Ietswaart et al., 2014; Sugawara and Kaneko, 2011; Walter et al., 2017), we wondered how we could modify or extend these molecular models to exhibit true regular positioning.

Ietswaart et al. have previously shown that regular positioning can theoretically be achieved, independently of the particular mechanism of force generation, through the balancing of the diffusive fluxes of nucleoid-bound ParA-ATP into the plasmid from each longitudinal direction. If plasmids, which act as sinks for ParA-ATP, move in the direction of greatest incoming flux, then they will move toward the regularly positioned configuration since this is the unique configuration in which the diffusive fluxes balance. This 'flux balance' mechanism has since been argued to underlie positioning in several other systems (Hofmann et al., 2019; Murray and Sourjik, 2017; Schumacher et al., 2017). It was also realised that a critical component of the mechanism is that the ParA-ATP dimers must diffuse on the nucleoid sufficiently far before hydrolysing ATP and unbinding (Murray and Howard, 2019; Subramanian and Murray, 2021). If the associated length-scale, s, is too short then only ParA-ATP dimers that first bind the nucleoid close to the plasmid will have the opportunity to interact with it. As a result, the fluxes of ParA into the plasmid balance across most of the cell and the plasmid does not receive any positional information (Figure 2A (i-iii)). As s increases the plasmid receives more positional information through the disparity in the diffusive ParA flux and beyond a threshold of half the nucleoid length true regular positioning is possible (Figure 2A (iv-v)). Note that the threshold decreases with the number of plasmids - with each additional plasmid, a shorter distance needs to be 'sensed' per plasmid. Sensing between plasmids occurs through competition for the same ParA-ATP

dimers (Figure 2B). The diffusion of ParA dimers on the nucleoid referred to above could occur through hopping of dimers between DNA strands during transient unbinding events or through the direct contact of DNA strands. Indeed, this has been argued to be essential for ParA gradient formation in *Caulobacter crescentus* (Surovtsev et al., 2016b) and was observed *in vitro* using single-particle microscopy (Vecchiarelli et al., 2013). Note that formally there still exists a non-zero disparity in the incoming fluxes into the plasmid in the low s regime, however it becomes infinitesimal as s decreases below the threshold (Subramanian and Murray, 2021).

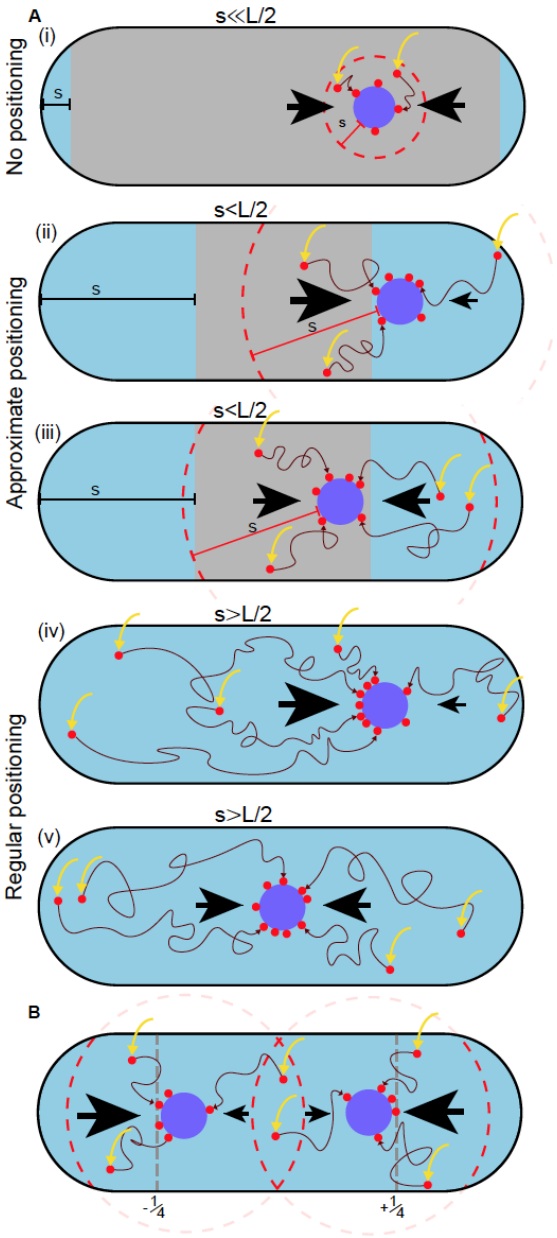

Figure 2. A difference in ParA-ATP flux can provide positional information if the diffusive length-scale is sufficiently long

The plasmid is biased away from locations at which there is a disparity in the incoming flux of ParA-ATP from either side (black arrows). However, the diffusion of ParA-ATP dimers on the nucleoid introduces a length-scale s, defined as the average distance dimers diffuse between association (yellow arrow) and dissociation due to hydrolysis. As a result the plasmid receives only ParA-ATP dimers that associate at most a distance s from it (red dashed line). (A) When $s \ll L/2$ (i), where L is the nucleoid length, a disparity in the fluxes into the plasmid only exists when the plasmid is very close to the poles (blue region). In the interior region (grey), the fluxes balance and the plasmid obtains no positional information. As s increases (ii-iii), the region in which the plasmid receives no positional information shrinks leading to approximate mid-nucleoid positioning. When $s \gtrsim L/2$ (iv-v), ParA-ATP dimers can explore the entire nucleoid before reaching the plasmid. Hence the fluxes of ParA-ATP into the plasmid are balanced only at the mid-position. True regular positioning is achievable. (B) A similar argument applies to a cell with multiple plasmids but with threshold $L/(2n)$, where n is the number of plasmids. Here, sensing between plasmids occurs through competition for the same ParA-ATP dimers (the overlap between the two dashed circles). Quarter-positions are highlighted by grey dashed lines.

The above argument explains why regular positioning was not observed in the DNA relay model (Surovtsev et al., 2016a). The key insight of that model was that bound ParA-ATP dimers experience the elastic fluctuations of the chromosomal DNA to which they bind and that these fluctuations can

power the movement of the partition complex across the cell. If the partition complex has more tethers to the nucleoid in one direction, then the elastic pull of the chromosome will lead to a net force in this direction and a corresponding directed movement. However, in the model the 'home' position of each DNA-bound ParA-ATP dimer remains fixed. ParA dimers were assumed to remain bound until they interact with a (ParB-coated) plasmid i.e. dimers do not diffuse (hop) on the nucleoid. Hence, the diffusive length-scale, s, is zero and regular positioning cannot occur. On the other hand, in the Brownian Ratchet model (Hu et al., 2017) diffusion of ParA-ATP dimers was included but with a length-scale four times shorter than the nucleoid length. The model was therefore also not inside the regular positioning regime.

3.1.3.3. A unifying stochastic model explains all plasmid behaviours in terms of physical parameters

Motivated by the previous discussion, we decided to develop our own minimal molecular model of ParABS positioning (Figure 3A). We take the DNA relay model as a starting point due to its relative simplicity (the Brownian Ratchet model explicitly models the ParA-ADP state and implements the force-dependent breakage of bonds and so has several more parameters).

The original DNA-relay scheme is as follows (Surovtsev et al., 2016a). The nucleoid is considered as a two-dimensional surface to which dimers of ParA-ATP can bind (at rate k_a). Upon association, dimers exhibit elastic fluctuations around their binding 'home' positions. If a dimer contacts the partition complex, itself modelled as a ParB-coated disk, it immediately binds, forming a tether between the PC and the nucleoid. The PC experiences the elastic force resulting from all attached tethers and moves as a Brownian particle under this force. Tethers are broken by ParB-induced hydrolysis (rate k_h), with ParA returning to a diffuse pool in the cytosol. Since the transition back to its DNA-binding competent state is slow (Vecchiarelli et al., 2010), the cytosolic pool of ParA-ATP is assumed to be well mixed.

Our model supplements this scheme with two additional components: diffusion of DNA-bound ParA-ATP dimers across the nucleoid (with diffusion coefficient D_h, where the subscript indicates diffusion of the home position) and plasmid-independent ATP hydrolysis and dissociation (with rate k_d). See Material and Methods for further details of the model. In the original model, dimers only unbind from the DNA due to interaction with ParB on the plasmid. However, ParA exhibits basal ATP activity (Ah-Seng et al., 2009). Together with diffusion on the nucleoid, plasmid-independent hydrolysis introduces a finite diffusive length-scale to the system, namely the distance a ParA dimer diffuses on the nucleoid before dissociating.

While, theoretical models with emergent behaviours are, in some sense, more than the sum of their parts, insight can be gained by identifying which physical properties of the model, typically describable by a set of dimensionless quantities, are responsible for a given behaviour. Identifying these informative quantities is critical since exploring the entire parameter space is often unfeasible. In this direction, we sought to identify the most important dimensionless quantities that characterise the behaviours of the system:

$\lambda = \frac{s}{L/2} = \frac{\sqrt{2D_h/k_d}}{L/2}$: This is the average distance, relative to half the nucleoid length, L, that each ParA-ATP dimer would theoretically diffuse on the nucleoid unhindered along each direction before unbinding due to basal ATP hydrolysis (Figure 3A). As discussed above, we expect that regular positioning is only possible when $\lambda \gtrsim 1$ and we confirm this below, justifying our identification of this quantity as important for the system dynamics.

$\varepsilon = \frac{k_h}{k_d}$: As λ is the ratio of the diffusive timescale to the timescale of basal hydrolysis, we reasoned that a second quantity describing the ratio of the timescale of ParB-induced hydrolysis (k_h) to the timescale of basal hydrolysis would also be informative in specifying the dynamics. We expect that when this ratio, ε, is sufficiently large, the concentration of ParA-ATP at the plasmid will be less than that away from the plasmid (Figure 3B, Figure 3–figure supplement 2) and the opposite when ε is small. This will allow us to probe the corresponding variation found experimentally.

Since the force on the plasmid is generated by the tethers between it and nucleoid-associated ParA, we reasoned that the number of nucleoid-associated ParA should also affect the dynamics of the system. Thus, we introduce a third quantity, θ, the steady state number of DNA-bound ParA dimers in the absence of ParB-induced hydrolysis, given by $\theta = \frac{k_a}{k_a + k_d} n_A$, with n_A being the total number of dimers in the system. Note that this involves the ratio of the third reaction rate of the system, the association rate of ParA to the nucleoid, k_a, relative to, once again, the basal hydrolysis rate k_d.

We can independently vary λ, ε and θ through the parameters D_h, k_h and n_A respectively. However, we found that while n_A had, unsurprisingly, a strong effect on the degree of stochasticity in the system, it had little effect on the nature of the dynamics (Figure 3–figure supplement 3). The different regimes were clearly detectable from at least $n_A = 50$ dimers. We therefore focused on λ and ε.

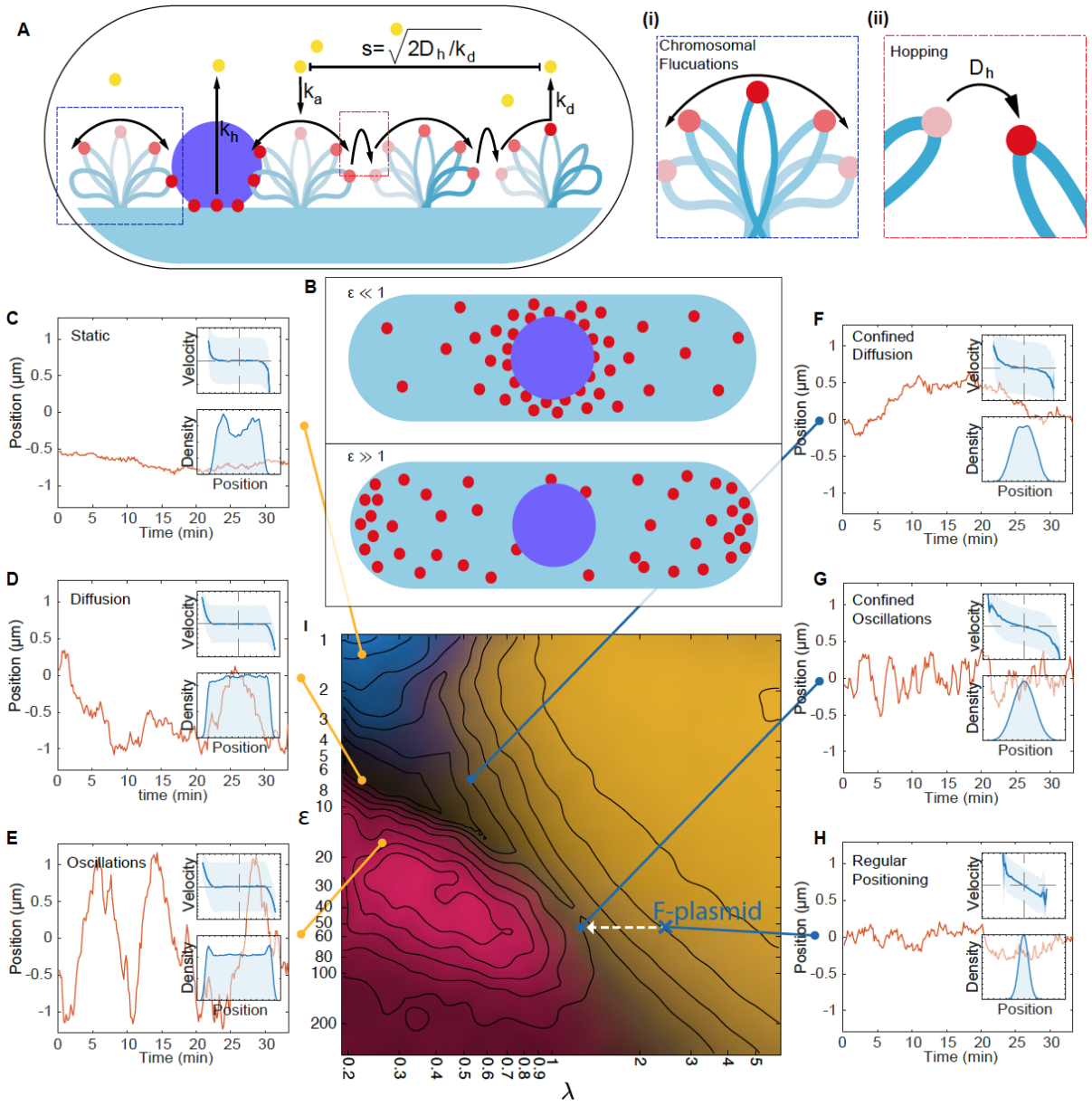

Figure 3. A minimal model of the ParABS system. (A) Schematic of the model. Light blue shading: nucleoid; light blue stroke: DNA-strand; red: nucleoid bound ParA; yellow: cytosolic ParA; purple: plasmid; arrows indicate binding and dynamics of the system; k_a: nucleoid binding rate of ParA; k_d: basal hydrolysis rate of ParA; k_{off}: hydrolysis rate of plasmid bound ParA. Insets: (i) elastic fluctuations of the chromosome, (ii) hopping or transfer of DNA-bound ParA-ATP dimers leads to an effective diffusion coefficient D_h. (B) A cartoon depicting low (<10) and high (>10) epsilon conditions. Low leads to a sink of ParA at the plasmid, high leads to a peak of ParA at the plasmid. (C) - (H) Example trajectories from different regimes form the phase diagram. Insets: top, velocity profile; bottom, position histogram; data from 1000 simulations. (I) Phase diagram obtained by varying D_h and k_{off}. Shown in terms of the dimensionless parameters λ and ε. The colour is based on an analysis of simulated trajectories as follows. Light brown: Regular positioning (confined and average position at mid-cell); blue: Static (confined and average position not at mid-cell); pink: Oscillations (highest peak in the position autocorrelation at non-zero lag); black: Diffusion (none of the previous). See Materials and Methods for details. Location of the F-plasmid is marked by a cross (Figure 3–figure supplement 1). Number of ParA-ParB tethers and plasmid mobility can be found in Figure 3–figure supplement 2).

We first considered the case of a single plasmid and performed simulations of the model over a range of values of these two quantities. The other parameters were fixed at estimated values (see Table 1) and the length of the simulated nucleoid was chosen to match the average length of cells with one F plasmid (Figure 1–figure supplement 1). These simulations produced a range of plasmid behaviours with a clear dependence on the position in this (λ, ε) phase space (Figure 3C to I). In particular we observed an interface at approximately $\lambda \approx 1$ separating two regimes at small λ, with a single regime at large λ.

We first consider small λ. The plasmid was found to move diffusively for $\varepsilon < 10$ as evidenced by its zero mean velocity across the nucleoid and flat-topped position distribution. However, the diffusivity of the plasmid decreases with decreasing ε so that at the lowest values of ε studied, at which the hydrolysis rate k_h at the plasmid is comparable to the rate k_d away from it, the plasmid moves so slowly that it is effectively static on the timescale of our simulations (35 min) and remains approximately at its initial position.

When ε is increased beyond 10 (i.e. when the hydrolysis rate at the plasmid is much greater than that away from it), we observed clear oscillatory behaviour, similar to that observed in previous models. This transition from static to diffusive to oscillatory can be understood in the terms of the differing timescale of tether dissociation on the one hand and the ParA repletion rate on the other (Hu et al., 2015; Walter et al., 2017). In the oscillatory regime, tethers break (due to ATP hydrolysis) faster than they can be replaced. This leads to a ParA depletion zone behind the plasmid that reinforces its movement in the same direction. The result is directed motion until the nucleoid edge, at which point the depletion zone fills, causing the plasmid to change direction. This turnaround time is apparent in the boundary peaks in the position distribution (Figure 3E). At even higher values of ε, tethers are so short lived that the dynamics become once again diffusive.

At the interface region $\lambda \approx 1$, the plasmid becomes confined to the centre region of the nucleoid where it exhibits either diffusive or oscillatory motion depending on the position along the interface. As λ is increased further, the positioning becomes more precise, the confined region shrinks and the plasmid exhibits true regular positioning. This is consistent with our hypothesis of the importance of the diffusive length-scale for the functioning of the flux-balance mechanism. Within this large regular positioning regime the position distribution, velocity profile and autocorrelation (Figure 3–figure supplement 1) have qualitatively the same form as we observed experimentally for cells containing a single plasmid and we found excellent quantitative agreement for $(\lambda, \varepsilon) = (2.66, 56.42)$ (Figure 4 and Figure 3H, blue cross in Figure 3I). Interestingly, these parameter values suggest that while the

dynamics of a single F plasmid sits within the regular positioning regime, it is not far from the interfacial region of confined oscillations.

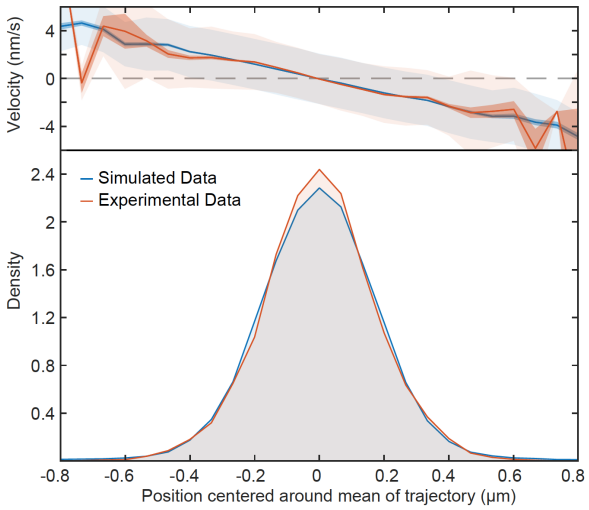

Figure 4: Fitted position of the F-Plasmid

Simulated data (blue) at $(\lambda, \varepsilon) = (2.66, 56.42)$ compared to the experimental data (red) from Figure 1E. Top: Mean velocity of plasmids as a function of position relative to the trajectory mean. Light and dark shading indicate standard deviation and standard error respectively. Bottom: Probability density of plasmid position relative to mean of trajectory.

We next measured how many ParA tethers were associated with the plasmid as the parameters were varied. We found the numbers of tethers varies positively with λ and negatively with ε, consistent with an increase in the flux of ParA dimers into the plasmid and longer tether lifetimes respectively (Figure Figure 3–figure supplement 2A). Interestingly we found a clear relationship between the mobility of the plasmid and the number of ParA tethers, with the oscillatory regime having the fewest number of (simultaneous) tethers and the greatest mobility and the regularly-positioning regime at low ε, having the most tethers and the slowest movement (Figure Figure 3–figure supplement 2B). This was also apparent from kymographs of the ParA distribution (Figure 3–figure supplement 2C). Note however, that the latter regime does exhibit regular positioning - it simply takes a very long time for the plasmid to move to mid-position. Similarly, in the 'static' regime the plasmid actually exhibits very slow diffusive motion. In this sense there are really only three regimes (diffusive, oscillatory and regular positioning) and their interfaces.

Finally, we explored how the other parameters of the model affect the dynamics. We varied the main parameters across four orders of magnitude centred on the set used in Figure 4 (which lies in the regular positioning regime). We found that only the two parameters varied in our sweep were able to push the system into the static or oscillatory regimes (Figure 3–figure supplement 4). Starting from the diffusive regime, this could also be achieved by changing the basal hydrolysis rate k_d consistent with how the dimensionless quantities λ and ε depend on it (changing k_d should move the system diagonally in the phase diagram). To confirm the role of these dimensionless quantities in determining the dynamics, we varied D_h, k_d, k_h and k_a simultaneously over two orders of magnitude. This has the effect of modulating the turnover rate of ParA tethers while keeping λ, ε and θ fixed. We found no

change in the nature of the dynamics beyond an expected increase in the frequency of the fluctuations in the plasmid position as the tether turnover rate is increased (Figure 3–figure supplement 5).

We also explored if regular positioning is achievable in the absence of ParA-ATP diffusion on the nucleoid (i.e. $D_h = 0$). However, we found that it only occurs if the length scale of chromosome fluctuations $\sigma_{x,y}$ is increased far beyond its measured value of about 0.1 μm to 1 μm (Figure 3–figure supplement 4A). At this unphysically high value, each DNA-bound ParA dimer can, through the fluctuations of the underlying DNA, interact with the plasmid over long distances and from across the cell. The plasmid is therefore positioned at mid-cell because this is the only position where the net force from all ParA dimers balances. However, based on the measurements of the chromosome fluctuations (Lim et al., 2014; Surovtsev et al., 2016a; Wiggins et al., 2010), we believe this regime is not biologically relevant.

3.1.3.4. Regular positioning of two plasmids

We next considered the case of cells having two and more plasmids. We found that our model could reproduce the same quarter positioning as observed for F plasmid (Figure 5A). Importantly, regular positioning was achieved irrespective of where the two plasmids were initially positioned. This is in contrast to the model of Hu et al., in which plasmids move apart a fixed distance. We also simulated plasmid replication by duplicating one plasmid during the simulation. We found that the replicated plasmids moved apart rapidly towards the quarter positions in a qualitatively similar way as we observed in our experimental data (Figure 5B-E). We expect that better knowledge of the biochemical parameters would further improve this comparison.

Examining the phase diagram for more than one plasmid, we found the boundary of the regularly positioning regime expands to lower λ values (Figure 5–figure supplement 1B). This is consistent with s, the distance ParA-ATP dimers diffuse on the nucleoid, needing to be greater than $\frac{L}{2n}$ for regular positioning to occur (Figure 2). When we displayed the phase diagrams in terms of $\lambda_n = n\lambda = \frac{s}{L/2n}$, we found that they all collapsed onto each other, with regularly positioning only occurring for $\lambda_n \gtrsim 1$, further confirming the importance of this parameter (Figure 5–figure supplement 1C).

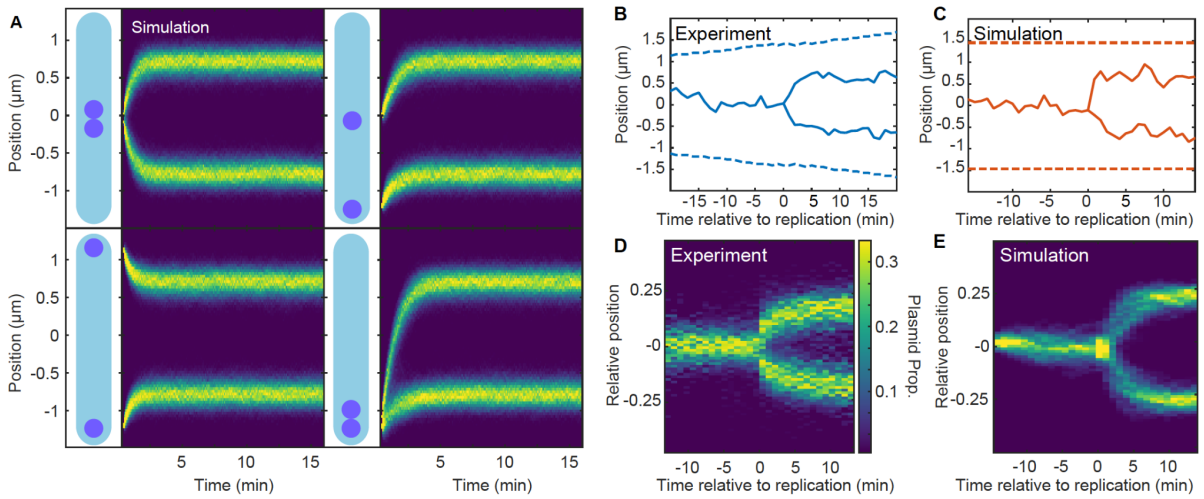

Figure 5. Regular positioning of two simulated plasmids

(A) Kymographs showing the distribution of plasmid positions starting from different initial positions along the long axis. Data is from 1000 simulations each. Nucleoid length is chosen to match our experimental data (see Table 2). (B) Example of F plasmid replication (splitting ParB-mVenus focus) event. (C) Example simulated replication event. (D) Kymograph of F plasmid splitting events as in (B). Data from 500 cell cycles were combined according to the time of focus splitting. (E) Kymograph of simulated plasmid replication. Upon replication, both plasmids occupy the same position but only one inherits the ParA-ParB tethers. This does not affect the result as the number of tethers equilibrates rapidly. Data from 1000 simulations. Note that in (D) position is relative to cell length, rather than nucleoid length as in the simulations (D). See also Figure 5–figure supplement 1A.

3.1.3.5. Length dependent transition to the confined oscillatory regime

While F plasmid operates within the regular positioning regime, our model predicts that its dynamics can become oscillatory by decreasing λ (Figure 3I, white arrow). Since λ depends inversely on the nucleoid length, L, we wondered whether oscillations would appear in longer cells. When we used the same model parameters determined above but with different lengths, we found that the system could indeed enter the (confined) oscillatory regime, commensurate with the length-induced decrease in λ (Figure 6A).

Motivated by these results, we went back to our F plasmid data and examined cells harbouring one plasmid with greater than average cell length. Consistent with our simulations, we found multiple cell cycles in which the plasmid was initially stably positioned at mid-cell but as the length of the cell increased, appeared to display low-amplitude oscillations (Figure 6B). To investigate if this transition was reproducible, we developed a method to classify segments of trajectories as oscillatory (or processive), regularly positioned or undetermined based on the velocity autocorrelation between consecutive frames (Figure 6–figure supplement 1). Binning the individual time-points from these classified segments according to cell length revealed the relative abundance of the two populations (Figure 6C). We found a marked increase in the proportion of oscillatory segments from cell lengths of about 3 μm, with up to 50% of timepoints being classified as oscillatory, consistent with our

prediction. This also confirms a previous rough estimate that F plasmid operates not far below the threshold of oscillatory instability (Walter et al., 2017).

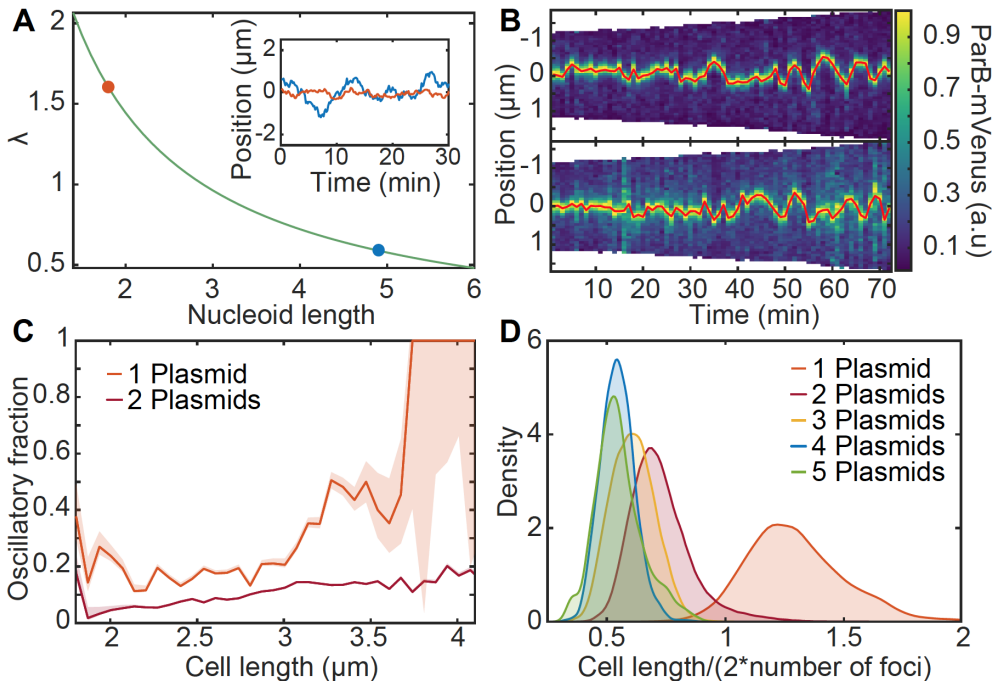

Figure 6. Length dependent effect on the transition from regular positioning to oscillations

(A) Inverse dependence of λ on nucleoid length. Inset: simulated trajectories at the highlighted lengths/λ values. (B) Kymographs of ParB-mVenus signal along the long axis of two cells depicting the transition from regular position to low-amplitude oscillations. The plasmid trajectory is highlighted in red. (C) The fraction of the oscillating population plotted against cell length for cells with one (879 cell cycles) or two plasmids (5044 cell cycles) (see Figure 6–figure supplement 1). (D) The distribution of threshold length-scale (L/2n) for cells containing different numbers of plasmids. Data from 16346 cell cycles.

Interestingly, the same analysis on cells containing two plasmids revealed a significantly smaller proportion of oscillatory segments and a weaker length dependence. However, this is again consistent with our prediction that, in a cell with n plasmids, ParA-ATP dimers need to diffuse on the nucleoid an average distance of at least $L/(2n)$ in order for the plasmids to sense each other and regular positioning to occur (Figure 2B). Since, within the population, this threshold distance is greatest for cells containing a single plasmid (Figure 6D), it is in these cells that we are most likely to observe a transition to oscillatory behaviour. More specifically, these results suggest that ParA-ATP dimers diffuse a distance of about 1.5 μm before dissociating. We also examined how oscillations are affected by changes in plasmid number within individual cells i.e. upon plasmid replication (Figure 6–figure supplement 2). We found that oscillatory behaviour appeared to decrease in that a classification of oscillatory dynamics before replication was not a reliable indicator of oscillatory dynamics

afterwards. The same was true across generations - we observed a rapid decay in the autocorrelation for containing an oscillatory trajectory segment (Figure 6–figure supplement 3.

3.1.3.6. pB171 operates closer to the oscillatory regime than F Plasmid

We have shown above that F plasmid is, for the most part, regularly positioned within cells, with a transition towards oscillatory behaviour only occurring in those cells with greatest sensing threshold $L/(2n)$, i.e. in cells with the lowest plasmid concentration. Might other ParABS systems exhibit more pronounced oscillatory dynamics? To explore this, we examined the dynamics induced by the ParABS system of the plasmid pB171. We chose this system as it has previously been described as oscillatory (Ringgaard et al., 2009) and it belongs to the other family of ParABS systems, namely type 1b (F plasmid is type 1a).

Using a previously constructed TetR/*tetO* labelling system, we first determined the copy number of this system and found it to be comparable to F plasmid (Figure 7–figure supplement 1). We then examined plasmid dynamics in cells containing a single plasmid and found clear unambiguous oscillatory behaviour in ~80% of such cells (Figure 7A), in stark contrast to F plasmid (Figure 7–figure supplement 2). This was reflected in the flat-topped plasmid position distribution (Figure 7B), which was very different from that of F plasmid and more similar to what we obtained in the oscillatory regime of our model (Figure 3–figure supplement 1). More importantly, the oscillatory nature of the dynamics was reflected in the position and velocity autocorrelations (Figure 7C,D), including a positive velocity autocorrelation between consecutive frames, which is a signature of processive motion (Figure 1–figure supplement 2). These curves qualitatively matched those obtained in our model within the oscillatory regime (Figure 3–figure supplement 1).

We also found that oscillatory dynamics were more likely in longer cells (Figure 7E), consistent with our prediction of the importance of nucleoid length in determining the dynamical regime (through λ). Overall oscillations were almost four times as likely for pB171 as for F plasmid (Figure 7F). However, this was much reduced when we considered cells with two plasmids, for which oscillations were much less apparent (Figure 7E,F, Figure 7–figure supplement 3). This suggests that, similar to F plasmid, the ParABS system of pB171 does not lie entirely within the oscillatory regime, but only enters it for cells containing a single plasmid, in which the sensing distance required for regular positioning is longest (see above).

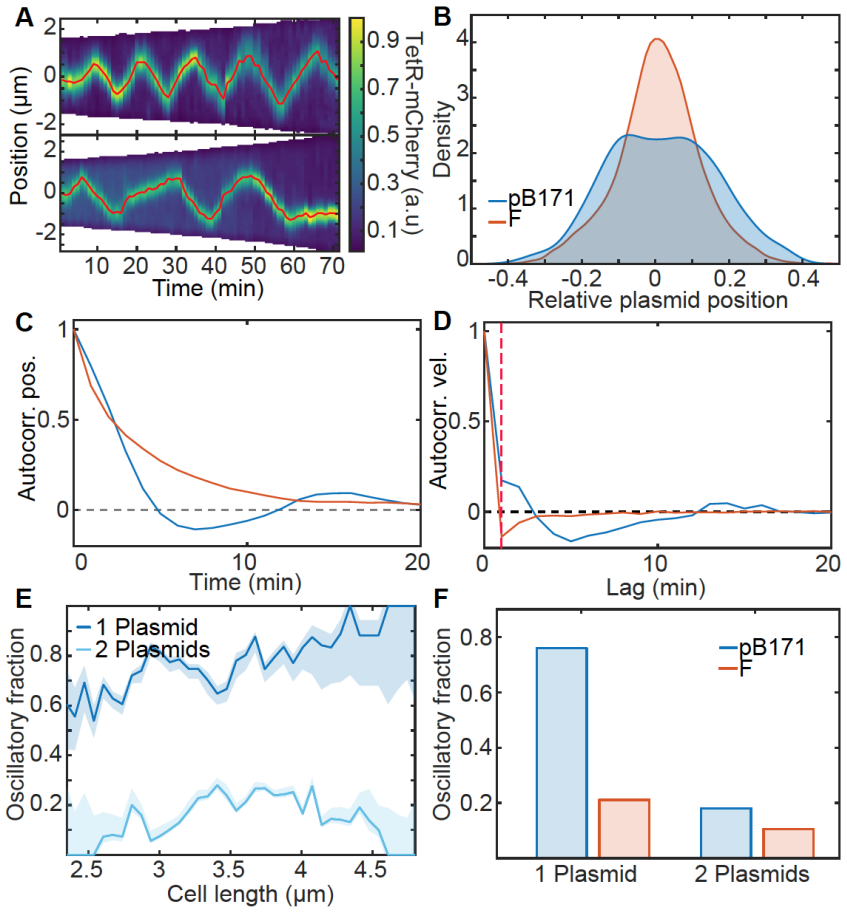

Figure 7. pB171 and its differences to F

(A) Two kymographs (pB171) of TetR-mCherry signal along the long axis of two cells with one plasmid. Red line indicates the trajectory of the plasmid (more examples of trajectories can be found in Figure 7–figure supplement 2). (B) Relative position position distribution for pB171 (blue, 68 cell cycles) and F-Plasmid (red, as in Figure 6) for cells with one plasmid. (C,D) Position and velocity autocorrelation for pB171 (blue) and F plasmid (red). Positive velocity autocorrelation at 1 min (dashed red line) indicates processive dynamics (see Figure 1–figure supplement 2). (E) The proportion of trajectory time points classified as oscillatory from pB171 plotted against cell length for cells containing one or two plasmids. Data from 68 and 117 cell cycles respectively. (F) Comparison between oscillating population of pB171 and F-Plasmid with one and two plasmids on whole population level.

3.1.4. Discussion

ParABS systems have become a paradigm of self-organisation within bacterial cells. Yet, it is still unclear how these systems function. Three main research questions can be identified: 1) How does ParB spread over the centromeric region to form the nucleoprotein partition complex (PC), 2) What is the nature of the force underlying directed movement of the PC, and 3) How is the directionality and positioning of the PC specified? The recent discovery (Jalal et al., 2020; Osorio-Valeriano et al., 2019; Soh et al., 2019) that ParB dimers are CTP-dependent DNA clamps that load onto, and slide away from, *parS* sites has shed light on the first question, at least for the type 1a and chromosomal ParABS systems. While the force-generating mechanism underlying directed movement has yet to be definitively identified, the current proposal, supported by modelling, is that movement is powered by the elastic fluctuations of the chromosome and/or ParA-ParB tethers (Hu et al., 2015; Lim et al., 2014).

The nature of partition complex positioning on the other hand has yet to be resolved. On the experimental side, quantitative measurements of plasmid dynamics have been lacking. It has therefore

not been possible to determine if plasmids are truly regularly positioned, which implies some method of geometry sensing or if they are rather recruited by, for example, regions of higher DNA density (Le Gall et al., 2016). Furthermore, while oscillatory dynamics have been observed, it was not clear whether this is representative of how these systems operate. On the modelling side, there have been several studies, both deterministic (Adachi et al., 2006; Ietswaart et al., 2014; Jindal and Emberly, 2019; Sugawara and Kaneko, 2011; Walter et al., 2017) and stochastic (Hu et al., 2021, 2017; Ietswaart et al., 2014; Surovtsev et al., 2016a), each producing some set of dynamical behaviours. However, the lack of quantitative dynamical measurements has meant that none of these models has been quantitatively compared or tested against experimental observations. It has therefore been unclear which model or, more specifically, which model ingredients, best describe plasmid positioning by ParABS.

In this work, we addressed this deficiency by first performing a high-throughput quantitative analysis of plasmid dynamics. We determined that the motion of F plasmid, which hosts a type 1a ParABS system, is consistently biased towards specific home positions (mid-cell in cells containing a single plasmid, approximate quarter positions in cells with two plasmids) as if pulled by a spring-like force. The precision of this 'regular positioning' was seen in the consistently spatially-varying average velocity of the plasmid and supports the presence of a geometry sensing mechanism. However, such positioning was not observed in either of the two existing molecular-level models of plasmid positioning (Brownian Ratchet (Hu et al., 2017) and DNA relay (Surovtsev et al., 2016a)). Our results therefore indicate that neither model is consistent with the dynamics of F plasmid (and indeed of pB171, see below).

Our model lies between the DNA relay and Brownian Ratchet models in terms of the model ingredients but encompases, in terms of qualitative outputs, the previous stochastic models as specific cases according to λ, the ratio of the length scale of ParA-ATP dimer diffusion on the nucleoid and the nucleoid length (Figure 8A). At one extreme the DNA relay model (Surovtsev et al., 2016a) does not incorporate ParA-ATP diffusion ($\lambda = 0$) and therefore oscillations are the only non-trivial behaviour. On the other extreme, the model of Ietswaart et al. (Ietswaart et al., 2014) includes it but without basal ATP hydrolysis ($\lambda = \infty$) such that only regular positioning is obtained. A model of the PomXYZ system of *Myxococcus xanthus* (Schumacher et al., 2017), which positions a protein rather than plasmid cargo, also lies here. The Brownian Ratchet model (Hu et al., 2017) on the other hand incorporates dimer diffusion but with a length scale (giving $\lambda = 0.5$) that places the system just outside of the regular positioning regime for the case of a single plasmid. This explains why this study found single plasmids to exhibit 'local excursions' around mid-cell i.e. approximate rather than regular positioning (Figure 3F). Further comparison of the different stochastic models is given in the Materials and Methods section and in Table 3.

A recent deterministic model deviates from this scheme (Jindal and Emberly, 2019) and requires some explanation. This model does not include ParA diffusion on the nucleoid (hence $\lambda = 0$) but it nonetheless produces regular positioning. This is in contrast to our model, for which we otherwise could not obtain, with biologically relevant parameters, regular positioning (Figure 3–figure supplement 4), as well as the DNA relay model. We believe this disparity is due to the continuous nature of the Jindal and Emberly model. The plasmid velocity is determined by the entire ParA dimer concentration but weighted according to the distance from the plasmid. Therefore, there is a regime (in which the plasmid movement is slower than ParA un-/binding) for which the mid-cell position is the stable configuration since at that location the weighted sum of ParA dimers on either side balances. This holds even if the ParA gradient is locally symmetric around an off-centre plasmid and because the model is continuous, even small differences produce an effect. In contrast, in the deterministic models with ParA-ATP diffusion in the nucleoid, the geometry of the cell is encoded in the local ParA gradient around the plasmid (when λ is sufficiently large) and plasmid positioning does not rely on interactions with distant ParA. In our stochastic model without such diffusion ($D_h = 0$), the relatively low concentration of ParA dimers means that the effect of rare long-distance interactions between the plasmid and ParA dimers is not sufficient to affect the dynamics due to the inherent stochasticity of the system. The Jindal and Emberly model also predicts that oscillations occur for intermediate plasmid concentrations i.e. the system transitions from regular positioning to oscillations to regular positioning with increasing plasmid concentration. However, we only observe the latter transition experimentally.

We also note that a previous deterministic model (Walter et al., 2017) implemented an alternative scheme in which the finite diffusive length-scale is of cytosolic ParA rather than the nucleoid-associated dimers. From a mathematical viewpoint, this system has very similar dynamics and is capable of both oscillations and regular positioning. However, given the rapid diffusion of small cytosolic proteins and the slow transition of ParA to its DNA-binding competent state (Vecchiarelli et al., 2010), we expect that ParA in the cytosol is well-mixed.

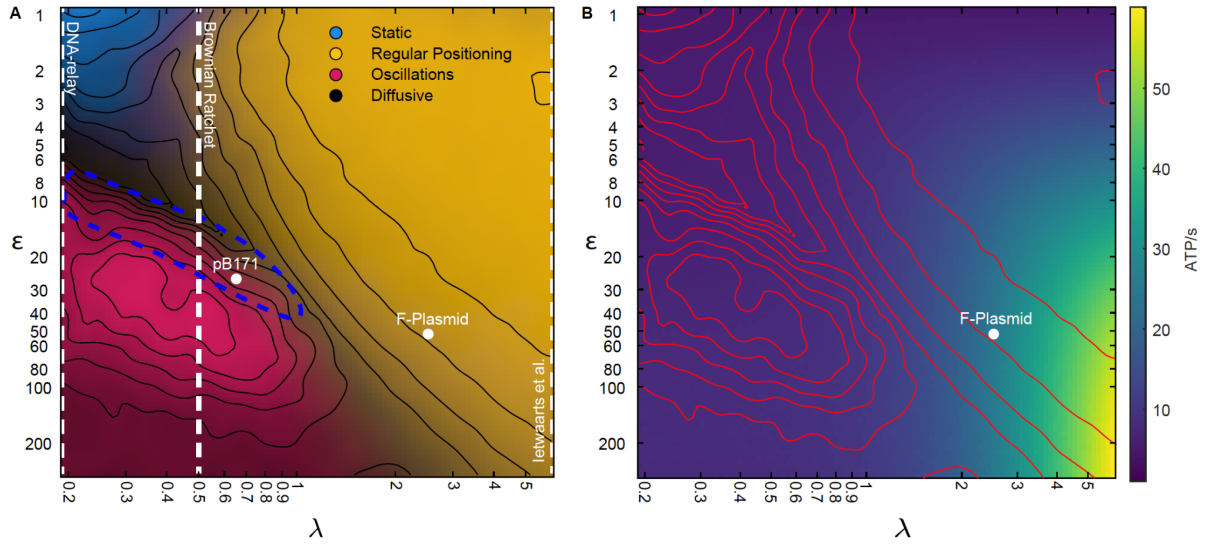

Figure 8. The model encompasses existing stochastic models as limiting cases

(A) The phase diagram for a single plasmid from Figure 3I with the conceptual location of existing stochastic models indicated. In the DNA relay model, bound ParA-ATP dimers do not diffuse on the nucleoid and so $\lambda = 0$. The model cannot produce regular positioning. In the model of Ietswaart et al., ParA dimers diffuse on nucleoid but only dissociate by interacting with the plasmid, therefore $\lambda = \infty$ and the model lies entirely in the regular positioning regime. For both models, the y-axis represents k_h the hydrolysis rate at the plasmid. The Brownian Ratchet lies between these two extremes. The length scale associated to ParA diffusion is finite but its value was fixed at $s = 0.5$ μm (given $\lambda = 0.5$ for a 2 μm nucleoid), so that the regular positioning regime is not explored. The locations of the ParABS systems of F plasmid and pB171 in cells containing a single plasmid are shown. The location of pB171 is an estimate based on a qualitative comparison of its dynamics. The blue dashed line marks the region in which the period of the oscillations resembles the experimental observations. (B) ATP consumption rate. Red lines are the contours from (A).

In Figure 8A, we indicate the location of F plasmid in the one-plasmid phase diagram of our model based on the fit to the subpopulation of cells containing a single plasmid (Figure 4). While on average it lies in the regular positioning regime, we have seen that in the longest cells it exhibits confined oscillations. This was predicted by our model since longer cells have lower λ. We also examined the dynamics of the type 1b ParABS system of pB171 and found clear oscillatory dynamics in the majority of cells carrying a single plasmid. While we do not have enough data for an accurate fitting, the 10-15 min period of the oscillations places pB171 within the indicated region, suggesting that both ε and λ are lower than for F plasmid. While adjusting the model parameters can change the nature (period etc) of the dynamics within the regions, we have found that their locations within the phase diagram are relatively robust (see e.g Figure 3–figure supplement 3 and 4). Thus, while additional system parameters may be involved, we speculate that that λ, and hence the diffusive length scale of ParA dimers, is lower for pB171 than for F plasmid. It remains to be seen if this is the case. Nevertheless, our results indicate that both F plasmid and especially pB171 lie close to the dynamical transition between regular positioning and oscillations, with the systems only crossing into the oscillatory regime for the subpopulation of cells with the lowest plasmid concentration. This was

previously suggested but not experimentally demonstrated for F plasmid (Walter et al., 2017). Importantly, both systems can be explained by the same model.

To understand why this might be the case, we used our model to measure the consumption of ATP throughout the explored phase space. Interestingly, we found that the oscillatory regime consumed the least ATP (Figure 8B). This is because this regime has the least flux of ParA-ATP dimers into the plasmid due to the short distance ParA dimers diffuse before dissociating from the DNA despite the fact that the plasmid moves back and forth across the nucleoid i.e. the directed movement of the plasmids cannot compensate for the reduced incoming flux of ParA dimers. Consistent with this, the oscillatory regime has the fewest simultaneous ParA-plasmid tethers (Figure 3–figure supplement 2). That the dynamics due to ParABS lie just below the onset of oscillation may therefore be due to achieving regular positioning while at the same time minimising energy consumption.

Overall, our results uncover the dynamical nature of ParABS systems and propose a unified stochastic model that accurately explains the observed plasmid dynamics and dynamical transitions. This model and the insights gained from it will further our understanding of chromosomal ParABS systems, which share many similarities with their plasmid-based relatives (especially those of type 1a). In particular, having a clear picture of partition complex dynamics will be useful to untangle the unknown role of CTP in partition complex positioning and segregation.

3.1.5. Acknowledgements

We thank Jean-Yves Bouet (Toulouse), Christine Jacobs-Wagner (Stanford), Yong Zhang (Copenhagen), Kenn Gerdes (Copenhagen) for providing strains and plasmids.

Table 1 Model Parameters

Parameter	**Brief description**	**Value**	**Source**
k_a	Association rate to the nucleoid of cytosolic ParA	0.19 s^{-1}	As in Surovtsev et al., 2016a. Based on *in vitro* measurement from Vecchiarelli et al., 2010. Results in 95% ParA nucleoid association in the absence of a plasmid.
k_d	Dissociation due to basal hydrolysis rate of ParA	0.01 s^{-1}	Based on *in vitro* measurement from Hwang et al., 2013; Vecchiarelli et al., 2013.
k_h	Tether dissociation due to plasmid stimulated hydrolysis of ParA	0.01 to 3 s^{-1}	Sweeped over in this study.
D_p	Diffusion coefficient of the plasmid	3 x 10^{-3} $\mu m^2 s^{-1}$	As in Surovtsev et al., 2016a. Based on MSD of a Δ*par* plasmid.
D_h	Diffusion coefficient of ParA home position on the nucleoid	3.22 x 10^{-4} to 0.29 $\mu m^2 s^{-1}$	Sweeped over in this study.
D_A	Diffusion coefficient of DNA-bound ParA due to chromosomal fluctuations	0.01 $\mu m^2 s^{-1}$	As in Surovtsev et al., 2016a;
W	Width of the cell	0.95 μm	This study
L	Length of the cell	2.5 to 4.34 μm	This study
dt	Simulation time step	0.001 s	This study
R_p	Radius Plasmid	0.05 μm	As in Surovtsev et al., 2016a.

			Estimate from Sanchez et al., 2015.
R_A	Radius ParA	0.002 μm	As in Surovtsev et al., 2016a. Based on ParA crystal structure from Leonard et al., 2005.
σ_x	Width of elastic fluctuations of the chromosome along long cell axis	0.1 μm	As in Surovtsev et al., 2016a.
σ_y	Width of elastic fluctuations of the chromosome along short cell axis	0.05 μm	As in Surovtsev et al., 2016a.
n_A	Number of ParA dimers	500	Midrange estimate from Adachi et al., 2006; Bouet et al., 2005; Lim et al., 2014.
n_p	Number of Plasmids	1 to 5	This study

Table 2 Simulation parameters used in figures

Figure	$k_h\,(1/s)$	$D_h\,(\mu m^2/s)$	$L\,(\mu m)$	n_p
Figure 3C, Figure 3–figure supplement 1B, Figure 3–figure supplement 4A, Figure 3-figure supplement 5B	0.0133	0.000440	2.53	1
Figure 3D, Figure 3–figure supplement 1C, Figure 3–figure supplement 4B, Figure 3-figure supplement 5B	0.0769	0.000440	2.53	1
Figure 3E, Figure 3–figure supplement 1D, Figure 3–figure supplement 4C, Figure 3-figure supplement	0.1785	0.000623	2.53	1

5B				
Figure 3F, Figure 3–figure supplement 1E	0.0752	0.002497	2.53	1
Figure 3G, Figure 3–figure supplement 1F	0.5642	0.014162	2.53	1
Figure 3H, Figure 3–figure supplement 1G, Figure 3–figure supplement 4D, Figure 3-figure supplement 5B	0.5642	0.056760	2.53	1
Figure 3I, Figure 8, Figure 3–figure supplement 1A, Figure 3–figure supplement 2, Figure 3–figure supplement 3[2], Figure 3–figure supplement 4E	0.01 to 3	0.000322 to 0.29	2.53	1
Figure 4	0.5642	0.056760	2.53	1
Figure 5A	0.5642	0.056760	2.91	2
Figure 5C,E[3]	0.5642	0.007179	2.91	1 -> 2
Figure 6A (inset: orange, blue)	0.5642	0.056760	1.82,4.93	1
Figure 5–figure supplement 1A	0.5642	0.056760	3.54	1 -> 5
Figure 5–figure supplement 1B,C	0.01 to 3	0.000322 to 0.29	2.53, 2.91, 3.67, 4.34	1,2,3,4

Table 3 Model comparison

	Hopping and relay (This study)	DNA-Relay (Surovtsev et al., 2016a)	Brownian-Ratchet (Hu et al., 2017)	Model of (Ietswaart et al., 2014)	Model of (Schumacher et al., 2017) [4]
Elastic fluctuations	✓	✓	✓	X	✓

[2] n_A was changed in the range of 5 to 1000
[3] k_d was changed to 0.001
[4] In this model of the PomXYZ of *Myxococcus xanthus*, PomZ is the analog to ParA and the PomXY cluster is the analog of the partition complex

Basal-hydrolysis	✓	X	✓	X	X
ParA diffusion on nucleoid	✓	X	✓	✓	✓
Length-scale of ParA diffusion on the nucleoid	finite	0	finite	infinite	infinite
ParA diffusion on partition complex	X	X	X	X	✓
Hydrolyzed nucleoid bound ParA state	X	X	✓	X	X
Limited binding of ParA to partition complex	X	X	✓	✓	X
Limit on tether length	X	X	✓	X	X
Cytosolic ParA pool	Well mixed	Well mixed	Unlimited pool, well mixed	Well mixed	Well mixed
Observed behaviour	Diffusion, regular positioning, static, oscillations	Oscillations	Diffusion, local excursion[5], static, oscillations	Regular positioning[6]	Regular positioning

[5] The Brownian ratchet model has all the necessary components/mechanisms to produce regular positioning. However, the parameters used in the study were such that the length-scale was not sufficiently high for regular positioning and therefore only 'local excursion' (approximate positioning in our terminology) instead of regular positioning was observed in cells with one plasmid.

[6] There are two stochastic models (with and without ParA filaments) presented in Ietswaart et al., 2014. The model including filaments is an extension of the other. Both models are capable of producing regular positioning.

3.1.6. Materials and Methods

3.1.6.1. Strains and growth condition

F plasmid experiments use strain DLT3125 (Sanchez et al., 2015), a derivative of the *E. coli* K-12 strain DLT1215 (Bouet et al., 2005) containing the mini-F plasmid derivative pJYB234. This plasmid carries a functional ParB-mVenus fusion. Overnight cultures were grown at 37°C in LB-Media containing 10 µg/ml thymine + 10 µg/ml chloramphenicol.

Experiments on plasmid pB171, use strain SR1 (Ringgaard et al., 2009), a derivative of the *E. coli* K-12 strain containing a *ΔpcnB* mutation which reduces the copy number of the hosted pB171-derived plasmids. SR1 carries plasmids pSR233 and pSR124 (Ringgaard et al., 2009). Plasmid pSR233 is a miniR1 plasmid carrying the parABS system (*par2)* of pB171 in addition to a *tetO* array. Plasmid pSR124 encodes an inducible *tetR-mCherry* fusion under the control of a P_{BAD} promoter. TetR binds to *tetO* and allows to track the motion of pSR233. Overnight cultures were grown at 37 °C in LB-Media containing 1 µg/ml thiamine + 50 µg/ml kanamycin + 100 µg/ml ampicillin.

3.1.6.2. Microfluidics

Like the original mother machine (Wang et al., 2010), our design consists of a main channel through which nutrient media flows and narrow growth-channels in which cells are trapped. However, we follow Baltekin et al. (Baltekin et al., 2017) and include i) a small opening at the end of each growth channel ii) a waste channel connected to that opening to allow a continuous flow of nutrients through the growth channels iii) an inverted growth-channel that is used to remove the background from fluorescence and phase contrast. We used a silicon wafer with this design to create the mother machine. We poured a polydimethylsiloxane (PDMS) mixture composed of a ratio of 1:7 (curing agent:base) over the wafer and let it rest at low pressure in a degasser for ~30 minutes to remove air bubbles inside. The PDMS was then baked at 80°C overnight (~16h). The cured PDMS was peeled off the wafer. Before imaging, the chip is bonded to a glass slide using a plasma generator (30 seconds at 75W) and subsequently baked for a further 30 minutes at 80°C, while the microscope is prepared.

3.1.6.3. Microscopy

We used a Nikon Ti microscope with a 100x/1.45 oil objective and a Hamamatsu Photonics camera for all imaging. For imaging cells of strain DLT3125 we used a mother machine. Overnight cultures were inoculated into fresh media (M9 + 0.5% glycerol + 0.2% casamino acids + 0.04 mg/mL thymine + 0.2 mg/mL leucine + 10 µg/mL chloramphenicol) for 4 hours at 30°C before imaging. Cells were loaded into the chip through the main channel and the chip was placed into a preheated microscope at 30°C. The cells were constantly supplied with fresh media by pumping 2 µL/min of M9 + 0.5% glycerol + 0.2% casamino + 0.04 mg/mL thymine + 0.2 mg/mL leucine through the microfluidic chip. Cells were grown for 2 hours inside the microscope before imaging. Cells were imaged at 1 minute intervals for approximately 72 hours. Both phase contrast and YFP-signal were captured. Imaging was repeated independently with similar results.

For imaging cells of strain SR1 we used agar pads. Overnight cultures were inoculated into fresh media (M9 + 0.5% glycerol + 0.2% casamino + 1 μg/ml thiamine + 10 μg/ml arabinose + 50 μg/ml kanamycin + 100 μg/ml ampicillin) for 2 hours at 30°C before imaging. The arabinose was added to induce synthesis of *tetR-mCherry*. Longer or continuous induction of arabinose leads to replication defects. Cells were placed on an 1% agar pad made from M9 + 0.5% glycerol + 0.2% casamino acids + 1 μg/mL thiamine and they were imaged at 1 minute intervals for 4 hours. Both phase contrast and RFP-signal were captured. Imaging was performed twice and the data combined.

3.1.6.4. Image processing

Our image processing pipeline for mother-machine experiments consists of three parts: I) preprocessing, II) segmentation and foci finding, and III) cell and foci tracking. While Parts I and III use custom Matlab scripts, Part II is based on SuperSegger (Stylianidou et al., 2016), a Matlab-based package for segmenting and tracking bacteria within microcolonies (original code is available at https://github.com/wiggins-lab/SuperSegger), that we modified to better handle high-throughput data. SuperSegger employs pre-trained neural networks to segment cells by identifying their boundaries. It comes with a pre-trained model for *E. coli* which worked very well with our data. Therefore there was no need to train our own neural network. SuperSegger is capable of tracking cells however the tracking did not work properly with mother-machine images and so we developed our own method. Nevertheless, acknowledging that one of the main components of our pipeline, the segmentation, uses SuperSegger we refer to the entire pipeline as MotherSegger (code is available at https://gitlab.gwdg.de/murray-group/MotherSegger/-/tree/PaperParABS).

In Part I each frame of an acquired image stack is aligned (the offset between frames in x and y is removed). Afterwards the image stack is rotated so the growth channels are vertical. A mask of the mother machine layout is fitted to the phase contrast, using cross-correlation, to identify where the growth channels are located. Each growth channel is extracted from the image stack and the flipped inverted channel is subtracted to remove the background from both the fluorescence signal and phase contrast. The images are then segmented and fluorescent foci are identified using Supersegger.

In Part III both foci and cells are tracked. Since cells cannot change their order inside the growth channel, they can be tracked by matching similar cell length between frames (starting from the bottom of the growth channels). Once individual cell cycles are identified, the foci positions found by Supersegger are re-specified relative to the bounding box of the cell (the smallest rectangular image containing the cell mask) on each frame. Since cells are vertical in the channels without any significant tilting, the bounding box is aligned with the cell axes. Within each cell cycle, foci are tracked between frames by finding the closest focus on the next frame inside the same cell cycle. The effect of growth on foci position was neglectable since cells grew on average much less than one pixel per frame at the 1 min frame rate and 100 min doubling time used here. Finally, half the cell length was subtracted from the foci positions along the long cell axis (vertical direction) so that 0 corresponds to the middle of the cell. The sign of the positions was also adjusted so that negative positions refer to the old-pole proximal side of the cell.

To filter out potential segmentation errors, cell cycles that do not have exactly 1 parent and 2 daughters are excluded from analysis along with their immediate relatives (with the exception of those who are pushed out of the growth channel). For the analysis of foci trajectories, we considered only trajectories coming from at least 12 consecutive frames with the same number of foci. For pB171, we used (unmodified) SuperSegger to process images of cells growing on agarose pads.

3.1.6.5. Over-damped spring

The distribution $p\left(x, \delta t|x_0\right)$ describes the probability that a Brownian particle, initially at position x_0, experiencing a spring-like force (harmonic potential) towards 0 is found at position x at a time δt later (Doi and Edwards, 1988):

$p\left(x, \delta t|x_0\right) = \sqrt{\frac{f/k_BT}{2\pi S}} exp\left[-\frac{f/k_BT}{2S}\left(x - x_0 e^{-\delta t/\tau}\right)^2\right]$ where $S = 1 - e^{-2\delta t/\tau}$, $\tau = \frac{k_BT}{fD}$, k_B is Boltzmann's constant and T is the absolute temperature. The stiffness of the spring is f/k_BT and D is the intrinsic diffusion coefficient. From this, it is straightforward to calculate the expected value and variance of the step-wise velocity $v := \frac{x - x_0}{\delta t}$ to be $E[v] = \frac{e^{-\delta t/\tau} - 1}{\delta t} x_0$ and $Var[v] = \frac{D\tau}{\delta t^2}(1 - e^{-2\delta t/\tau})$. Note the two properties characteristic of an (over-damped) spring-like force: The expected value $E[v]$ linearly scales with x_0 while the variance $Var[v]$ is independent of the initial position. We observed the same properties in our experimental data. We determined D and τ via $D = \frac{Var[v]Ln(\delta tm+1)}{\delta tm^2+2m}$, $\tau = -\frac{\delta t}{ln(\delta tm+1)}$ where m is the slope of the velocity profile.

The distribution $p\left(x, \delta t|x_0\right)$ can also be used to calculate the position and velocity autocorrelations:

$E[x(t_0)x(t_0 + t)]/E[x(t_0)^2] = e^{-t/\tau}$ and $E[v(t_0)v(t_0 + t)]/E[v(t_0)^2] = \frac{2e^{-t/\tau} - e^{-|t-\delta t|/\tau} - e^{-(t+\delta t)/\tau}}{2 - 2e^{-\delta t/\tau}}$ respectively. Finally, a characteristic force can be defined as the force on the particle at an extension of one standard deviation of the equilibrium distribution i.e. at $x = \sqrt{k_BT/k}$. For F plasmid, this gives a force of $F = k\sqrt{k_BT/k} = 0.019\, pN$ at $T = 30°C$.

3.1.6.6. Model

Our model is an extension of the previous DNA-relay model (Surovtsev et al., 2016a) that incorporates diffusion on the nucleoid (hopping) and basal hydrolysis of ParA-ATP and uses analytic expressions for the fluctuations rather than a second order approximation. Like the DNA relay it is a 2D off-lattice stochastic model and updates positions in discrete time steps dt. The implementation was written in C++ (code is available at https://gitlab.gwdg.de/murray-group/hopping_and_relay/-/tree/PaperParABS). It consists of the following components. ParA associates to the DNA non-specifically in its ATP-dependent dimer state with the rate k_a. Once associated, ParA (i.e. ParA-ATP dimers) moves in two distinct ways: (i) Diffusive motion on the nucleoid with the diffusion coefficient D_h. This is an effective description of the movement of dimers due to transient

unbinding events that allow them to 'hop' between DNA-strands. We do not consider the alternative scenario in which dimers transfer between DNA strands when the latter come into contact. In this scenario the effective diffusion coefficient would depend on the parameters describing the DNA fluctuations (D_A and $\sigma_{x,y}$). (ii) Between hopping events, each bound ParA dimer experiences the elastic fluctuations of the DNA strand it is bound to. This is implemented as elastic (spring-like) fluctuations around its initial position. Dimers dissociate from the nucleoid due to either basal ATP hydrolyse at a rate k_d or due to hydrolysis stimulated by ParB on the plasmid. The latter is modelled as a ParB-coated disc and ParB-ParA tethers form whenever the disk comes in contact with a ParA dimer. ParB-stimulated hydrolysis then breaks these tethers at a rate k_h, returning ParA to the cytosolic pool. The plasmid experiences the elastic force of every tethered ParA and moves according the its intrinsic diffusion coefficient D_p and the resultant force of all tethers. An overview of this scheme is shown in Figure 3A.

As in the DNA relay model we have made some simplifications that we next make explicit. First, we only modelled three states of ParA: 'nucleoid associated' and 'cytosolic' and 'tethered'. Second, cytosolic ParA are assumed to be well mixed. This is justified based on the slow conformation changes needed to return it to a state competent for DNA-binding (Vecchiarelli et al., 2010). Third, no individual ParB molecules were modelled, rather the plasmid is treated as a disk coated with enough ParB that each nucleoid bound ParA that makes contact with the plasmid instantaneously finds a ParB partner, therefore removing the need to model individual ParB. This is justified by the substantially higher local concentration of ParB compared to ParA at the plasmid.

The nucleoid is modelled as a rectangle with the dimensions $L \times W$. The positions of ParA and the plasmid(s), are updated every time step dt as follows. Between hopping events, each nucleoid associated ParA dimer fluctuates about a home position x_h. The new position $x(t + dt)$ of each dimer is given by $x(t + dt) = x_h + \delta x$, where δx is drawn with probability $p(\delta x, dt \mid x(t) - x_h)$ where x(t) is its original position (see section 'Over-damped spring') and the normalised spring constant (f/k_BT above) along each dimension is $1/\sigma_{x,y}^2$ and the diffusion coefficient D_A. During hopping events $x(t)$ and x_h are both offset by a value drawn from a Gaussian distribution with $\mu = 0$ and $\sigma = \sqrt{2D_h dt}$ for both dimensions. The displacement of the plasmid is determined similar to each ParA dimer but according to the resultant force acting on it. This resultant force vector has an effective spring constant equal to the spring constant of a single tether times the number of tethers and acts towards an equilibrium position $x_p(t) + \sum_{tethers} (x_h - x(t))/n$, where $x_p(t)$ is the plasmid position and the sum is over all (n) tethers. We ignore the effects of Torque. The intrinsic diffusion coefficient of the plasmid is D_p. If the plasmid has no tethers attached then it moves by normal diffusion, with displacements drawn from a Gaussian distribution with $\mu = 0$ and $\sigma = \sqrt{2D_p dt}$. The x and y components of all positions are updated independently and all simulations in this paper were run until the system reached equilibrium before acquiring data used for analysis.

3.1.6.7. Comparison of stochastic ParABS models

The most recent stochastic models of positioning by ParABS and ParABS-like systems explicitly incorporate earlier proposals for the mechanism of force generation, namely, that the elastic fluctuations of the DNA and/or ParA-ParB protein tethers can power the movement of cargo up the gradient of DNA-bound ParA-ATP dimers (Hu et al., 2015; Lim et al., 2014). However, the models differ in other ways (see Table 3).

The DNA relay model (Surovtsev et al., 2016a) does not allow DNA-bound ParA dimers to diffuse (hop) on the nucleoid. They fluctuate around a home position due to the elastic fluctuations of the underlying chromosomal locus. In our terminology, this model therefore has $\lambda = 0$, where λ is the ratio of the ParA dimer diffusive length scale to the nucleoid length (see main text). Oscillations were the only non-trivial behaviour found in this model.

The Brownian Ratchet model (Hu et al., 2017) on the other hand includes diffusion of dimers on the nucleoid as well as several other details such as explicit modelling of the transient DNA-bound ParA-ADP state, limited binding to cargo and the force and length-dependent dissociation of ParA-ParB tethers. It also includes basal (plasmid independent) ParA hydrolysis. Together with diffusion on the nucleoid, this results in λ being finite. However, through parameters analyses, its value was fixed at $\lambda = 0.5$. As a result, 'local excursions' (confined diffusion in our terminology) were observed for single plasmids rather than regular positioning. This led the authors to conclude that the biological system lies in a narrow regime of the model parameter space in which two or more plasmids are regularly positioned due to newly replicated plasmids moving apart a fixed distance ('directed segregation').

While not incorporating an explicit mechanism of force generation, the earlier model of Ietswaart et al. (Ietswaart et al., 2014) is worth mentioning here. This stochastic model was based on the formation of short DNA bound ParA-ATP filaments. It included diffusion of ParA dimers on the nucleoid but without basal hydrolysis ($k_d = 0$). Hence, the diffusion of each ParA dimer on the nucleoid is interrupted only upon interaction with a plasmid and $\lambda = \infty$. This model gave regular positioning, as did a variant without ParA filament formation, as the only non-trivial behaviour. The authors explained the emergence of regular positioning by a 'flux balance' argument. Plasmids are positioned because that is the unique configuration in which the diffusive flux of ParA dimers into each plasmid from either side balances (see Figure 2 and main text). They demonstrated this mathematically using a simplified deterministic model in which plasmids both act as sinks for ParA-ATP and move up the gradient of ParA-ATP on the nucleoid.

The above models have also been applied to ParA-like systems. In the PomXYZ system of *Myxococcus xanthus*, the ParA-like protein PomZ positions a large protein cluster formed by PomX and PomY at the middle of the cell. Schumacher et al. (Schumacher et al., 2017) explained this behaviour using the same elastic DNA/protein bond fluctuations as the models above, combined with the flux-balance mechanism of Ietswaart et al. Like the latter, their model did not include basal hydrolysis of PomZ and therefore $\lambda = \infty$. More recently, the Brownian Ratchet model has also been used to explain the positioning of carboxysomes in the cyanobacterium *Synechococcus elongatus* (MacCready et al., 2018).

3.1.6.8. Phase Space

To generate the phase space (Figure 3I, 8A, Figure 3–figure supplement 1A) of our model we chose 100 values of λ and ε, resulting in a 100 by 100 grid of different parameter combinations. This was done by varying D_h for λ and k_h for ε. To visualise the behaviour of each parameter combination we considered 3 quantities (i) φ, (ii) ψ and (ii) χ. (i) φ describes the goodness of regular positioning. The mean position of a plasmid-trajectory along the long axis is used as an input to a triangle wave function: $\varphi = f_n(x) = 1 - |2(nx/L - \lfloor nx/L \rfloor - 0.5)|$, where x is the mean position of the trajectory, L is the cell length and n is the number of plasmids. If the mean of a trajectory is equal to the position defined by regular positioning for that number of plasmids, then φ=1. If a trajectory is positioned at a pole or exactly between two regular positions it returns 0. (ii) ψ describes the mobility of a trajectory. ψ is equal to the standard deviation of the trajectory positions divided by $L/\left(\sqrt{12}n\right)$ (the standard deviation of the uniform distribution of width L/n). If a trajectory is oscillating, its distribution of position is roughly uniform (Figure 3E) resulting in a ψ close to 1. If a trajectory stays approximately at one position, ψ is much lower than 1. (iii) χ describes if a trajectory is oscillating. χ is equal to the highest positive maxima after the first negative minima in the normalised autocorrelation function of position. χ is equal to 0 if there is no negative minima or no positive maxima after a negative minima. From these three quantities, we calculated three descriptors with values between 0 and 1, describing the three regimes of oscillations, regular positioning and static. Regular Positioning: $(1 - \psi)\varphi$ is high if the plasmid is non-mobile and regularly positioned. Static: $(1 - \psi)(1 - \varphi)$ is high if non-mobile and not regularly positioned. Oscillations: χ is high if oscillating. Each regime is associated with a colour and this colour is scaled by its corresponding descriptor. The colours were chosen to be colourblind friendly (light brown RGB:[255 193 7], blue RGB:[30 136 229], pink RGB:[216 27 96]). Note that for diffusive trajectories, we expect $\varphi = 1$, $\psi = 1$ and $\chi = 0$ and hence all descriptors are 0 (visualised as black). To smoothen the phase diagram we used the morphological operation 'opening' followed by a 2-D Gaussian filter. The methodology above, while somewhat arbitrary, was found to describe the dynamics of the system well.

3.1.6.9. Classification of trajectories

Regular positioning and oscillations are distinguishable by calculating the velocity autocorrelation between adjacent frames (lag 1 min, Figure 1–figure supplement 2A). However, this method does not work for trajectories which change behaviour. Therefore, we developed a procedure to find segments inside a trajectory which are oscillatory or regularly positioned. A sliding window is moved across a trajectory and the velocity autocorrelation is calculated inside the window. If the autocorrelation at lag 1 is positive the point in the middle of the sliding window is annotated 'oscillatory'. Otherwise the point is labelled 'regularly positioned' (Figure 6–figure supplement 1A,B). Multiple points in a row with the same annotation form a segment (Figure 6–figure supplement 1B). With this procedure a trajectory can be broken down into multiple segments belonging to different behaviours (Figure 6–figure supplement 1C). To annotate our data we used a sliding window of size 12 and a requirement of 6 successive points of the same annotation to form a segment (Figure 6–figure supplement 1D). One weakness of this approach is the small window size which may result in false positives oscillatory segments.

3.1.7. References

Adachi S, Hori K, Hiraga S. 2006. Subcellular Positioning of F Plasmid Mediated by Dynamic Localization of SopA and SopB. *J Mol Biol* **356**:850–863. doi:10.1016/J.JMB.2005.11.088

Ah-Seng Y, Lopez F, Pasta F, Lane D, Bouet J-Y. 2009. Dual Role of DNA in Regulating ATP Hydrolysis by the SopA Partition Protein*. *J Biol Chem* **284**:30067–30075. doi:10.1074/jbc.M109.044800

Badrinarayanan A, Le TBK, Laub MT. 2015. Bacterial Chromosome Organization and Segregation. *Annu Rev Cell Dev Biol* **31**:171–199. doi:10.1146/annurev-cellbio-100814-125211

Baltekin Ö, Boucharin A, Tano E, Andersson DI, Elf J. 2017. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proc Natl Acad Sci* **114**:9170–9175. doi:10.1073/pnas.1708558114

Bouet JY, Rech J, Egloff S, Biek DP, Lane D. 2005. Probing plasmid partition with centromere-based incompatibility. *Mol Microbiol* **55**:511–525. doi:10.1111/j.1365-2958.2004.04396.x

Doi M, Edwards SF. 1988. The Theory of Polymer Dynamics. Clarendon Press.

Gerdes K, Moller-Jensen J, Jensen RB. 2000. Plasmid and chromosome partitioning: Surprises from phylogeny. *Mol Microbiol* **37**:455–466. doi:10.1046/j.1365-2958.2000.01975.x

Hatano T, Yamaichi Y, Niki H. 2007. Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Mol Microbiol* **64**:1198–1213. doi:10.1111/j.1365-2958.2007.05728.x

Hofmann A, Mäkelä J, Sherratt DJ, Heermann D, Murray SM. 2019. Self-organised segregation of bacterial chromosomal origins. *eLife* **8**:e46564. doi:10.7554/eLife.46564

Hu L, Rech J, Bouet J-Y, Liu J. 2021. Spatial control over near-critical-point operation ensures fidelity of ParABS-mediated bacterial genome segregation. *Biophys J* **120**:3911–3924. doi:10.1016/j.bpj.2021.08.022

Hu L, Vecchiarelli AG, Mizuuchi K, Neuman KC, Liu J. 2017. Brownian Ratchet Mechanism for Faithful Segregation of Low-Copy-Number Plasmids. *Biophys J* **112**:1489–1502. doi:10.1016/j.bpj.2017.02.039

Hu L, Vecchiarelli AG, Mizuuchi K, Neuman KC, Liu J. 2015. Directed and persistent movement arises from mechanochemistry of the ParA/ParB system. *Proc Natl Acad Sci* **112**:E7055–E7064. doi:10.1073/pnas.1505147112

Hwang LC, Vecchiarelli AG, Han Y-W, Mizuuchi M, Harada Y, Funnell BE, Mizuuchi K. 2013. ParA-mediated plasmid partition driven by protein pattern self-organization. *EMBO J* **32**:1238–1249. doi:10.1038/emboj.2013.34

Ietswaart R, Szardenings F, Gerdes K, Howard M. 2014. Competing ParA structures space bacterial plasmids equally over the nucleoid. *PLoS Comput Biol* **10**:e1004009. doi:10.1371/journal.pcbi.1004009

Jalal AS, Tran NT, Le TBK. 2020. ParB spreading on DNA requires cytidine triphosphate in vitro. *eLife* **9**. doi:10.7554/eLife.53515

Jindal L, Emberly E. 2019. DNA segregation under Par protein control. *PLOS ONE* **14**:e0218520. doi:10.1371/journal.pone.0218520

Kawalek A, Wawrzyniak P, Bartosik AA, Jagura-Burdzy G. 2020. Rules and Exceptions: The Role of Chromosomal ParB in DNA Segregation and Other Cellular Processes. *Microorganisms* **8**:105. doi:10.3390/microorganisms8010105

Kuwada NJ, Cheveralls KC, Traxler B, Wiggins PA. 2013. Mapping the driving forces of chromosome structure and segregation in Escherichia coli. *Nucleic Acids Res* **41**:7370–7377. doi:10.1093/nar/gkt468

Le Gall A, Cattoni DI, Guilhas B, Mathieu-Demazière C, Oudjedi L, Fiche J-B, Rech J, Abrahamsson S, Murray H, Bouet J-Y, Nollmann M. 2016. Bacterial partition complexes segregate within the volume of the nucleoid. *Nat Commun* **7**:12107. doi:10.1038/ncomms12107

Leonard TA, Butler PJ, Löwe J. 2005. Bacterial chromosome segregation: structure and DNA binding of the Soj dimer — a conserved biological switch. *EMBO J* **24**:270–282.

Lim HC, Surovtsev IV, Beltran BG, Huang F, Bewersdorf J, Jacobs-Wagner C. 2014. Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *eLife* **3**. doi:10.7554/eLife.02758

MacCready JS, Hakim P, Young EJ, Hu L, Liu J, Osteryoung KW, Vecchiarelli AG, Ducat DC. 2018. Protein gradients on

the nucleoid position the carbon-fixing organelles of cyanobacteria. *eLife* **7**. doi:10.7554/eLife.39723

Murray SM, Howard M. 2019. Center Finding in E. coli and the Role of Mathematical Modeling: Past, Present and Future. *J Mol Biol* **431**:928–938. doi:10.1016/j.jmb.2019.01.017

Murray SM, Sourjik V. 2017. Self-organization and positioning of bacterial protein clusters. *Nat Phys* **13**:1006–1013. doi:10.1038/nphys4155

Niki H, Hiraga S. 1997. Subcellular Distribution of Actively Partitioning F Plasmid during the Cell Division Cycle in E. coli. *Cell* **90**:951–957. doi:10.1016/S0092-8674(00)80359-1

Onogi T, Miki T, Hiraga S. 2002. Behavior of Sister Copies of Mini-F Plasmid after Synchronized Plasmid Replication in Escherichia coli Cells. *J Bacteriol* **184**:3142–3145. doi:10.1128/JB.184.11.3142-3145.2002

Osorio-Valeriano M, Altegoer F, Steinchen W, Urban S, Liu Y, Bange G, Thanbichler M. 2019. ParB-type DNA Segregation Proteins Are CTP-Dependent Molecular Switches. *Cell* **179**:1512-1524.e15. doi:10.1016/j.cell.2019.11.015

Ringgaard S, van Zon J, Howard M, Gerdes K. 2009. Movement and equipositioning of plasmids by ParA filament disassembly. *Proc Natl Acad Sci* **106**:19369–19374. doi:10.1073/pnas.0908347106

Roberts MAJ, Wadhams GH, Hadfield KA, Tickner S, Armitage JP. 2012. ParA-like protein uses nonspecific chromosomal DNA binding to partition protein complexes. *Proc Natl Acad Sci U S A* **109**:6698–703. doi:10.1073/pnas.1114000109

Sanchez A, Cattoni DI, Walter J-C, Rech J, Parmeggiani A, Nollmann M, Bouet J-Y. 2015. Stochastic Self-Assembly of ParB Proteins Builds the Bacterial DNA Segregation Apparatus. *Cell Syst* **1**:163–173. doi:10.1016/j.cels.2015.07.013

Schumacher D, Bergeler S, Harms A, Vonck J, Huneke-Vogt S, Frey E, Søgaard-Andersen L, Hoppert M, Kahnt J, Munoz-Dorado J, Søgaard-Andersen L. 2017. The PomXYZ Proteins Self-Organize on the Bacterial Nucleoid to Stimulate Cell Division. *Dev Cell* **41**:299-314.e13. doi:10.1016/j.devcel.2017.04.011

Soh Y-M, Davidson IF, Zamuner S, Basquin J, Bock FP, Taschner M, Veening J-W, De Los Rios P, Peters J-M, Gruber S. 2019. Self-organization of parS centromeres by the ParB CTP hydrolase. *Science*. doi:10.1126/science.aay3965

Stylianidou S, Brennan C, Nissen SB, Kuwada NJ, Wiggins PA. 2016. *SuperSegger* : robust image segmentation, analysis and lineage tracking of bacterial cells. *Mol Microbiol* **102**:690–700. doi:10.1111/mmi.13486

Subramanian S, Murray SM. 2021. Pattern selection in reaction diffusion systems. *Phys Rev E* **103**:12215. doi:10.1103/PhysRevE.103.012215

Sugawara T, Kaneko K. 2011. Chemophoresis as a driving force for intracellular organization: Theory and application to plasmid partitioning. *Biophysics* **7**:77–88. doi:10.2142/biophysics.7.77

Surovtsev IV, Campos M, Jacobs-wagner C. 2016a. DNA-relay mechanism is sufficient to explain ParA-dependent intracellular transport and patterning of single and multiple cargos. *Proc Natl Acad Sci* E7268–E7276. doi:10.1073/pnas.1616118113

Surovtsev IV, Lim HC, Jacobs-Wagner C. 2016b. The Slow Mobility of the ParA Partitioning Protein Underlies Its Steady-State Patterning in Caulobacter. *Biophys J* **110**:2790–2799. doi:10.1016/j.bpj.2016.05.014

Vecchiarelli AG, Han Y-W, Tan X, Mizuuchi M, Ghirlando R, Biertümpfel C, Funnell BE, Mizuuchi K. 2010. ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Mol Microbiol* **78**:no-no. doi:10.1111/j.1365-2958.2010.07314.x

Vecchiarelli AG, Hwang LC, Mizuuchi K. 2013. Cell-free study of F plasmid partition provides evidence for cargo transport by a diffusion-ratchet mechanism. *Proc Natl Acad Sci U S A* **110**:E1390-7. doi:10.1073/pnas.1302745110

Walter J-C, Dorignac J, Lorman V, Rech J, Bouet J-Y, Nollmann M, Palmeri J, Parmeggiani A, Geniet F. 2017. Surfing on Protein Waves: Proteophoresis as a Mechanism for Bacterial Genome Partitioning. *Phys Rev Lett* **119**:028101. doi:10.1103/PhysRevLett.119.028101

Walter JC, Rech J, Walliser NO, Dorignac J, Geniet F, Palmeri J, Parmeggiani A, Bouet JY. 2020. Physical Modeling of a Sliding Clamp Mechanism for the Spreading of ParB at Short Genomic Distance from Bacterial Centromere Sites. *iScience* **23**. doi:10.1016/j.isci.2020.101861

Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S. 2010. Robust growth of Escherichia coli. *Curr Biol* **20**:1099–103. doi:10.1016/j.cub.2010.04.045

Wiggins PA, Cheveralls KC, Martin JS, Lintner R, Kondev J. 2010. Strong intranucleoid interactions organize the Escherichia coli chromosome into a nucleoid filament. *Proc Natl Acad Sci U S A* **107**:4991–5. doi:10.1073/pnas.0912062107

3.1.8. Supplementary Figures

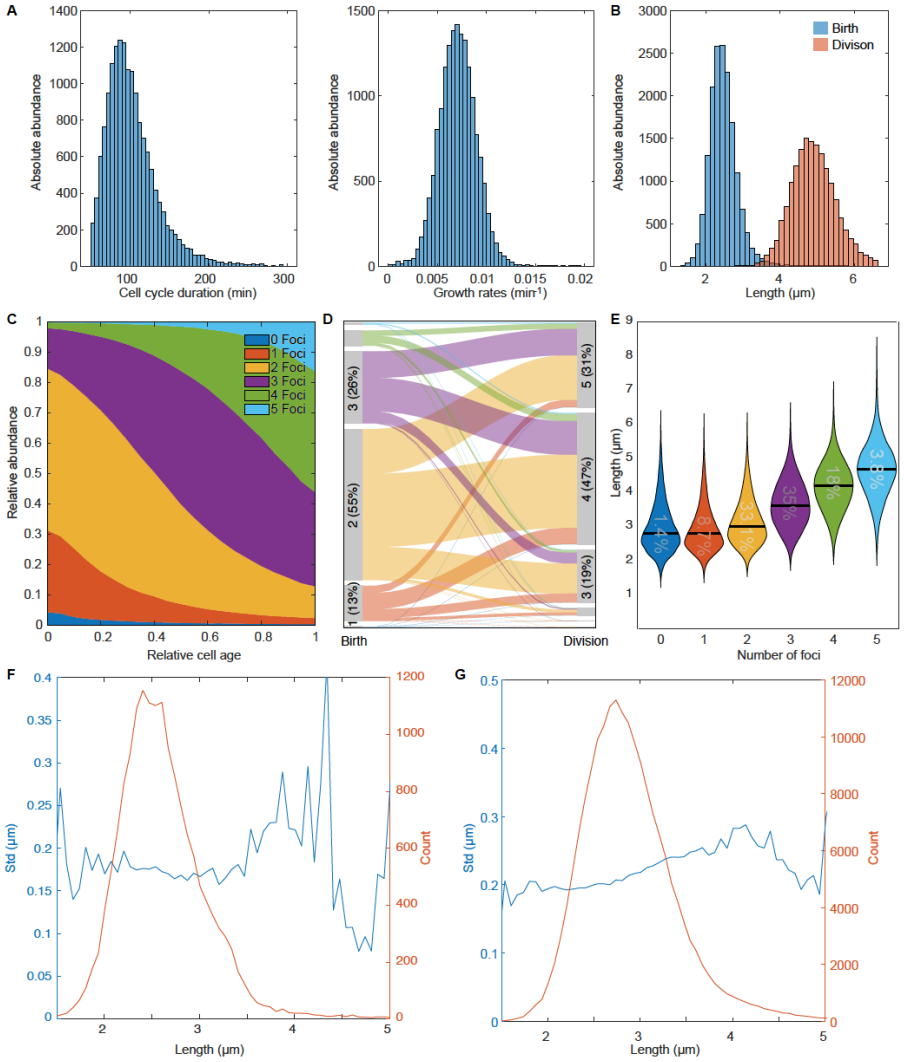

Figure 1–figure supplement 1. An overview of the F-plasmid

(A) Distribution of cell cycle duration and growth rate. The mean cell cycle duration is 101.17 ± 34.55 min (mean ± s.d.). (B) Distribution of cell length at birth and at division. The mean length is 2.46 ± 0.38 μm (mean ± s.d.) and 4.92 ± 0.63 μm (mean ± s.d.) respectively. (C) Distribution of ParB-mVenus foci number as a function of relative cell age. Cells have 1.9 ± 0.78 ParB foci (mean ± s.d.) at birth and 3.58 ± 0.96 ParB foci (mean ± s.d.) at division. (D) Alluvial plot showing how the number of foci changes from birth to division. The maximum number of plasmids of the first and last 5 frames were used to determine the number of plasmids at birth and division respectively. (E) Violin plot depicting the middle 99.9 percentile of the cell length distribution for cells with different numbers of plasmids. The number inside the violin plot indicates the relative abundance of each population. The black bar indicates the mean. (F) Standard deviation of position (blue) and plasmid count (orange) plotted against cell length for cells containing 1 foci. Plasmid position was centred around the mean of trajectory similar to Figure 1E. (G) Standard deviation of position (blue) and plasmid count (orange) plotted against cell length for cells containing 2 foci.

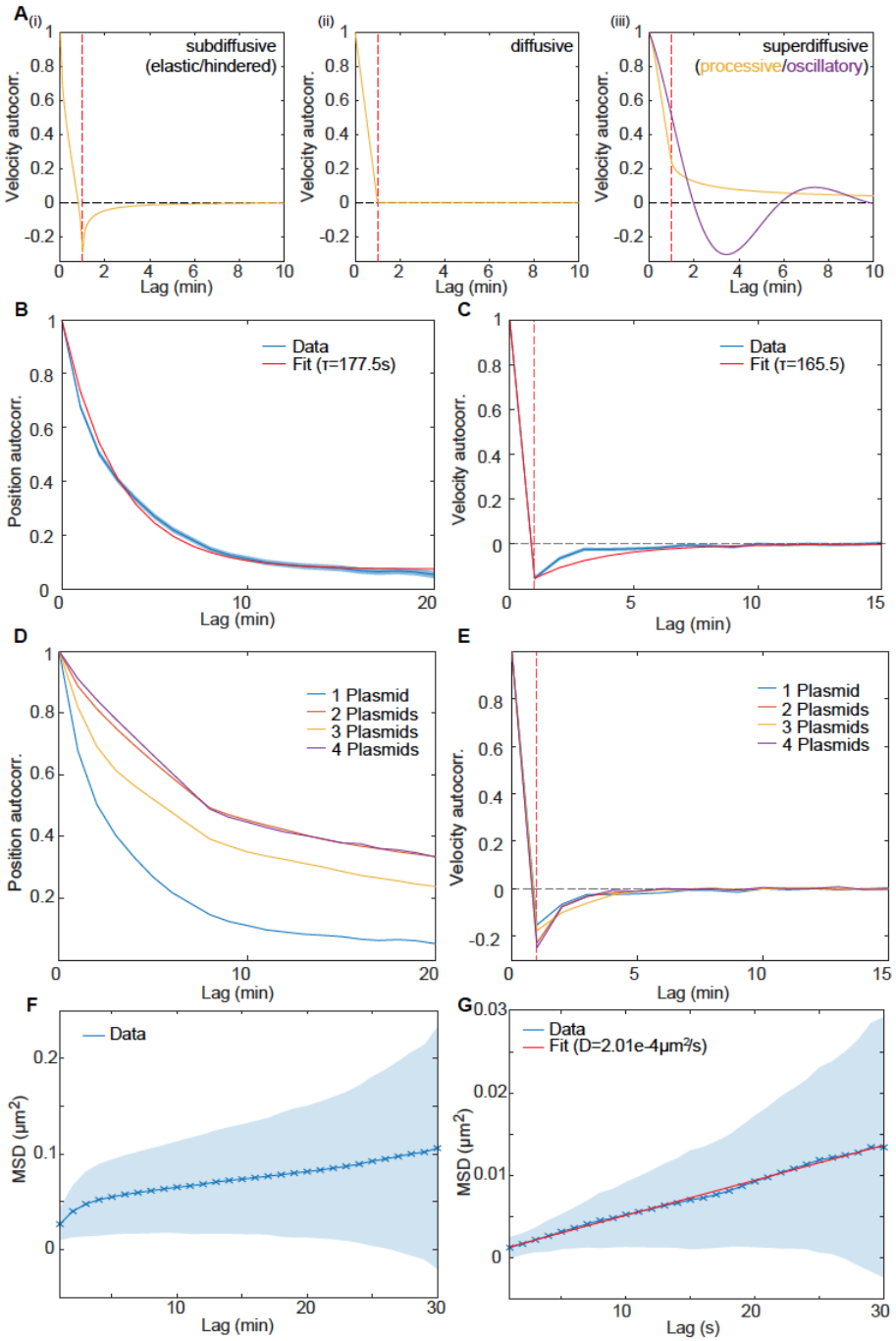

Figure 1–figure supplement 2. Dynamics of the F-plasmids are indicative of elastic/hindered diffusion

(A) The shape of the velocity autocorrelation (VAC) functions for (i) subdiffusive/elastic/hindered, (ii) diffusive and (iii) superdiffusive/processive dynamics. Red line indicates a lag equal to the sampling time (the time over which the velocity was measured). In subdiffusive/hindered processes, movement is biased towards the direction the particle came from; in a diffusive process there is no directional bias, while in superdiffusive processes (e.g. directed movement), the particle is biased to keep moving in the same direction. This results in the VAC between frames (i.e. at a lag equal to the sampling time) being negative, zero and positive, respectively

(indicated by the red line). The orange curves are generated using the normalised velocity autocorrelation $VAC = (|t-1|^{\alpha} + |t+1|^{\alpha} - 2|t|^{\alpha})/2$ for any process with $MSD(t) \sim t^{\alpha}$ and taking (i) $\alpha < 1$, (ii) $\alpha = 1$ and (iii) $\alpha > 1$. The purple line in (iii) is a sketch of what the VAC might look like for oscillatory dynamics. (B-C) Mean position autocorrelation (PAC) and VAC of trajectories from cells with one plasmid. Shaded area indicates standard deviation. The autocorrelation time-scale τ, in B and C, was obtained by fitting the PAC to $\frac{e^{-t/\tau}+c}{c+1}$ and the VAC to $\frac{2e^{-t/\tau}-e^{-|t-\delta t|/\tau}-e^{-(t+\delta t)/\tau}}{2-2e^{-\delta t/\tau}}$ (red lines), where δt is the time over which the velocity is calculated (1 frame = 60 s) and c is a constant accounting for any old-pole/new-pole bias. These functions arise from a model of over-damped spring-like motion. See the methods for further details. (D) Mean PAC of trajectories from different populations with 1,2,3 or 4 plasmids. (E) As in (D) but for the VAC. (F) MSD measured every 60 s (n=2558). (G) MSD measured every 1 s (n=75). The diffusion constant D=2.01±0.14e^{-4} μm^2/s (mean ± 95% CI) was obtained by fitting $2D * Lag + c$ to the MSD-curve (red line).

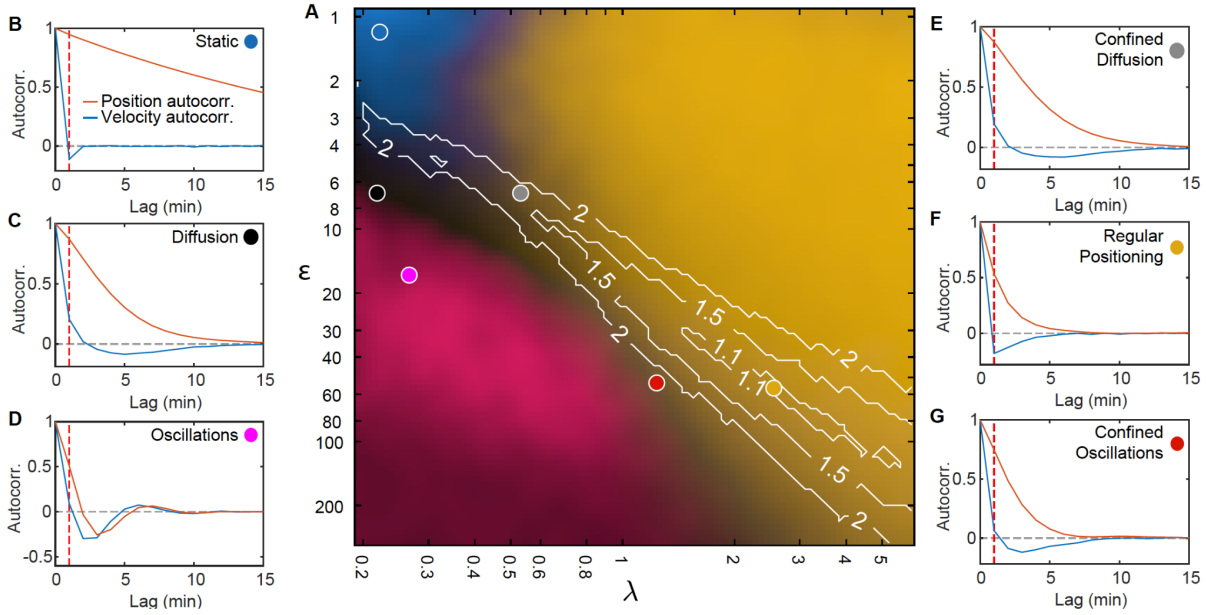

Figure 3–figure supplement 1. Fitting standard deviation of position and velocity place the F-plasmid inside the regular positioning regime

(A) Phase diagram of simulations with one plasmid (same as Figure 3I). Contours mark regions where the standard deviation of both position and velocity matches the experimental data, only by varying less than the fold change indicated by the number on the contour. (B-G) position (orange) and velocity (blue) autocorrelation at the indicated positions on the phase diagram (same location as in Figure 3I). Dashed red line indicates a lag of 1 min. Only the autocorrelations of the regular positioning regime match the experimental curves (Figure 1–figure supplement 2B,C)

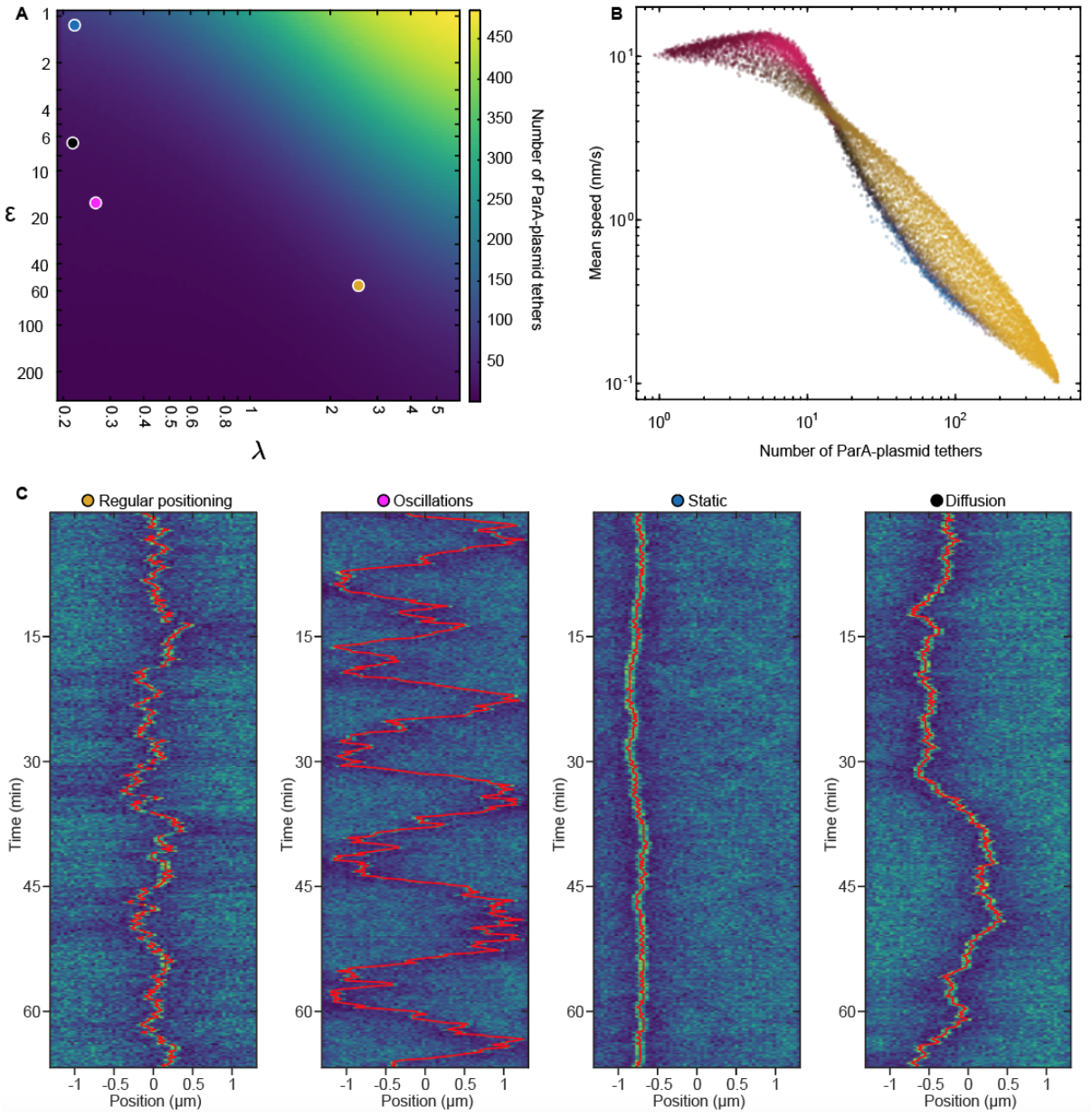

Figure 3–figure supplement 2. Number of ParA-plasmid tethers and their relation to plasmid mobility

(A) Heatmap depicting the number of ParA-plasmid tethers across the phase space for cells with on plasmid. (B) Data from Figure 3I re-plotted against the number of ParA-plasmid tethers and plasmid speed (measured over 1 minute intervals) instead of ε, λ. The colour of the data points is the same as in Figure 3I. (C) Kymographs of the ParA distribution of the marked points in (A), as also used in Figure 3. The position of the plasmid is indicated by the red line.

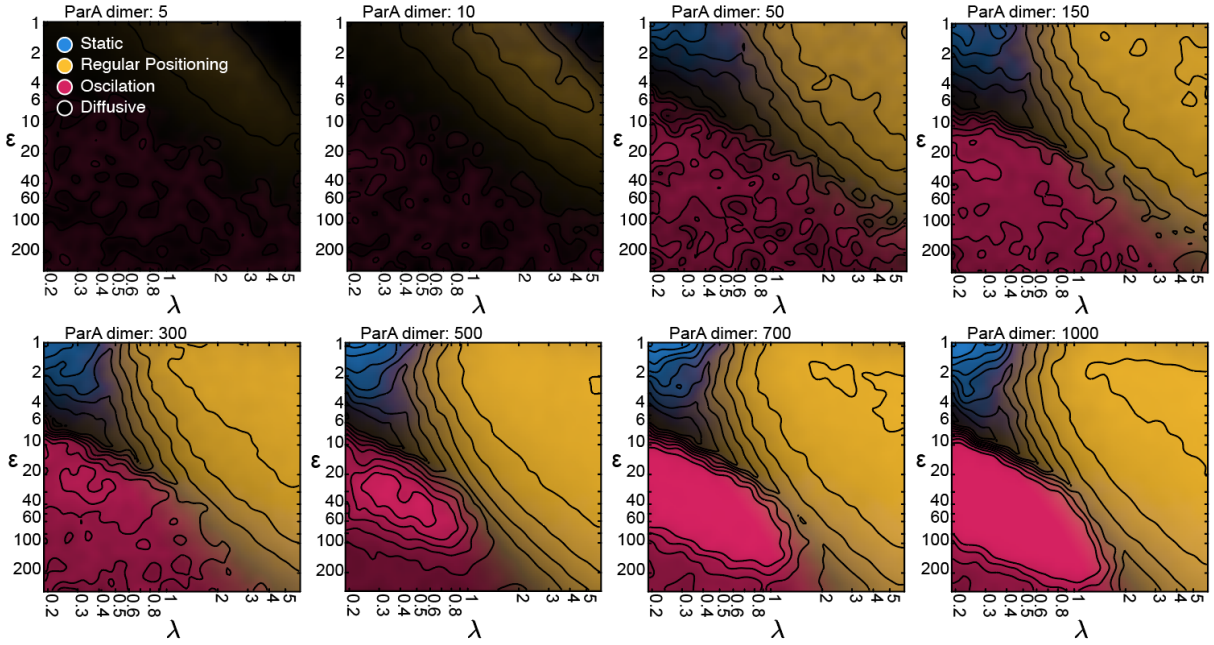

Figure 3–figure supplement 3. The system is robust against varying the total number of ParA dimers.

Same as in Figure 3I, but with varying numbers of ParA dimers. Unsurprisingly, diffusive dynamics dominate at low numbers. However, the regular positioning regime begins to appear from as little as 10 ParA dimers and all four regimes are detectable from 50 ParA dimers. As the number increases from 50 to 300, the borders between the regimes slightly shift. However, this saturates at around 300 ParA dimers, beyond which the number of ParA has little influence on the dynamical nature of the system.

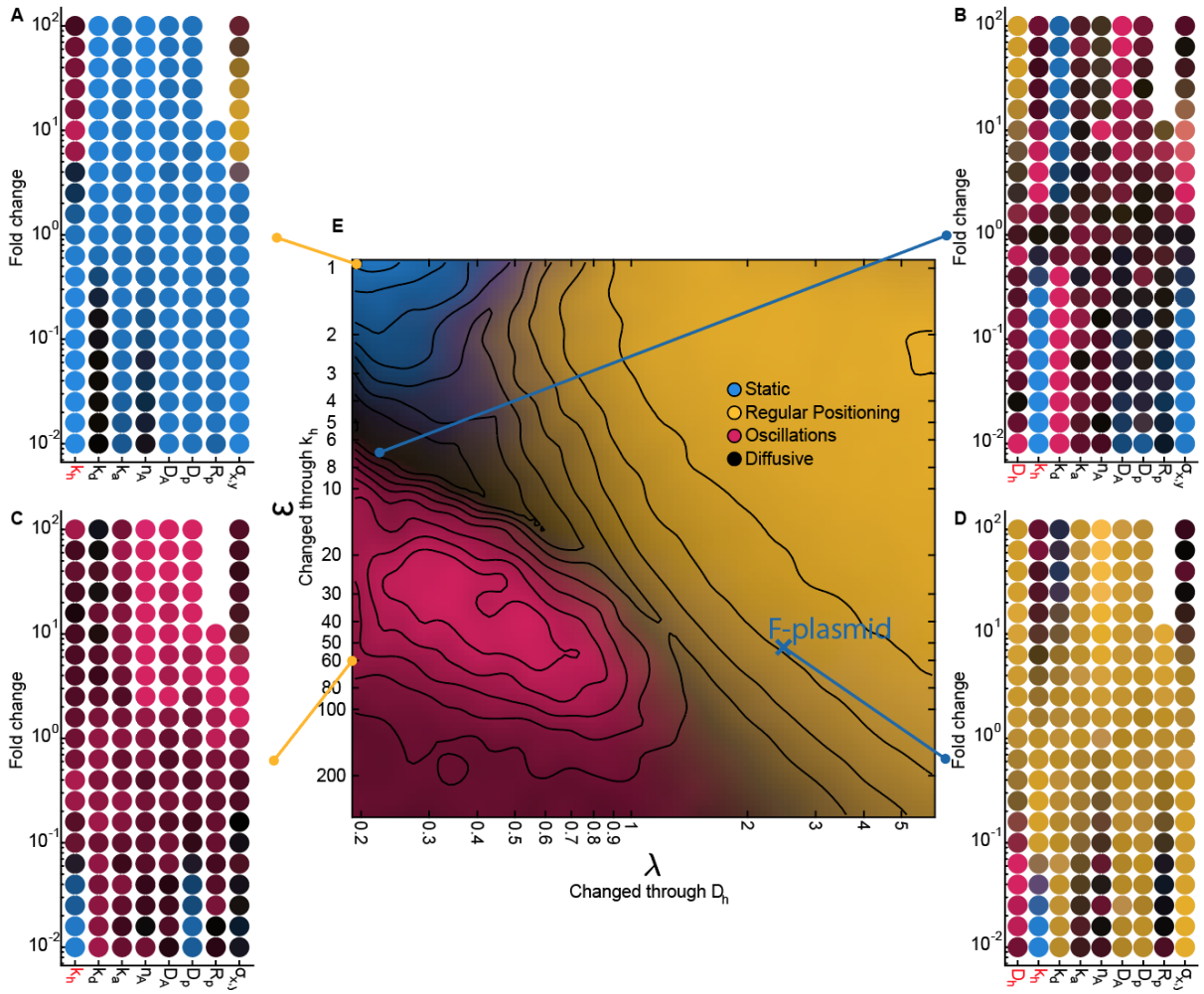

Figure 3-figure supplement 4. The effect of varying system parameters at characteristic locations in our phase diagram.

(A-D) Each scatter plot contains multiple 1D sweeps centred at the corresponding location in our phase diagram. Each 1D sweep (at its extreme) increases or decreases one parameter by a factor of 100 (fold change 10^{-2} - 10^{2}). At a fold change of 10^{0} the parameters are the same as at the indicated location. Labels in red indicate parameters which were used to change ε and λ in our phase diagram. The locations for (A) and (C) were chosen such that $D_h = 0$ (no diffusion on the nucleoid). (D) is located at our predicted parameters for F plasmid. The radius of the plasmid (R_p) could not be increased more than 10 fold since above that threshold the diameter of the plasmid was greater than the width of the cell. The colours in the scatter plot indicate the behaviour of the system as in Figure 3I. (E) Same as Figure 3I.

Notable transitions:

1) Decreasing the tether hydrolysis rate k_h results in longer-lived tethers and hence slower plasmid movement.

Beyond a point, the plasmid appears static on the timescale of the simulation. However, we have confirmed by performing longer simulations that it is not static but diffusive for $\lambda < 1$ and regularly positioned for $\lambda > 1$ (as explained in the text, the blue region in the top left of the phase diagram is technically diffusive) as predicted by flux balance. As k_h is increased in the $\lambda < 1$ regime, an oscillatory transition occurs when the plasmid begins to move faster than hydrolyzed ParA dimers can be replaced resulting in a depletion zone behind the plasmid and directed movement.

2) Decreasing D_h decreases the diffusive length-scale λ, moving the system out of the regular positioning regime and towards either oscillatory or diffusive dynamics. It also decreases the total flux of ParA into the plasmid leading to fewer tethers but this is not responsible for the nature of the dynamics as increasing n_A, the total amount of ParA, does not affect the nature of the dynamics (see also Figure 3-figure supplement 3).

3) Increasing k_d decreases both λ and ε (as well as the fraction of nucleoid-bound ParA dimers θ) and so moves the system approximately diagonally in the phase diagram.

4) The plasmid diffusion coefficient D_p is most relevant in the oscillatory regime, in which there are the fewest tethers. Oscillations rely on the plasmid moving faster than hydrolyzed ParA tethers can be replaced. Thus increasing plasmid mobility through D_p results in stronger directed movement and hence oscillations, while decreasing it moves the system towards more diffusive behaviour (C).

5) An additional requirement for non-diffusive dynamics is that the tether lifetime is longer that the timescale of the tether-induced 'pulling' ($\frac{1}{k_h} > \frac{\sigma_{x,y}^2}{D_p}$ for a single tether). This effect explains the darkening in the phase diagram at the bottom of the oscillatory regime. The same transition to diffusive dynamics occurs at high values of the spring constant $\sigma_{x.y}$. Note however that for the longest tether lifetime and high $\sigma_{x.y}$, regular positioned was observed at $D_h = 0$ ($\lambda = 0$; no diffusion on the nucleoid) i.e. outside of our claimed regular positioning regime (A). This occurs because at this unphysical value, $\sigma_{x.y}$ is comparable to the size of the cell and therefore DNA-bound ParA dimers can interact with the plasmid from every location within the cell. The plasmid is therefore positioned at mid-cell because this is where the net force from all tethered dimers balances. In other words geometry sensing occurs, not through the local detection of a disparity in incoming fluxes but through the global detection of all ParA dimers.

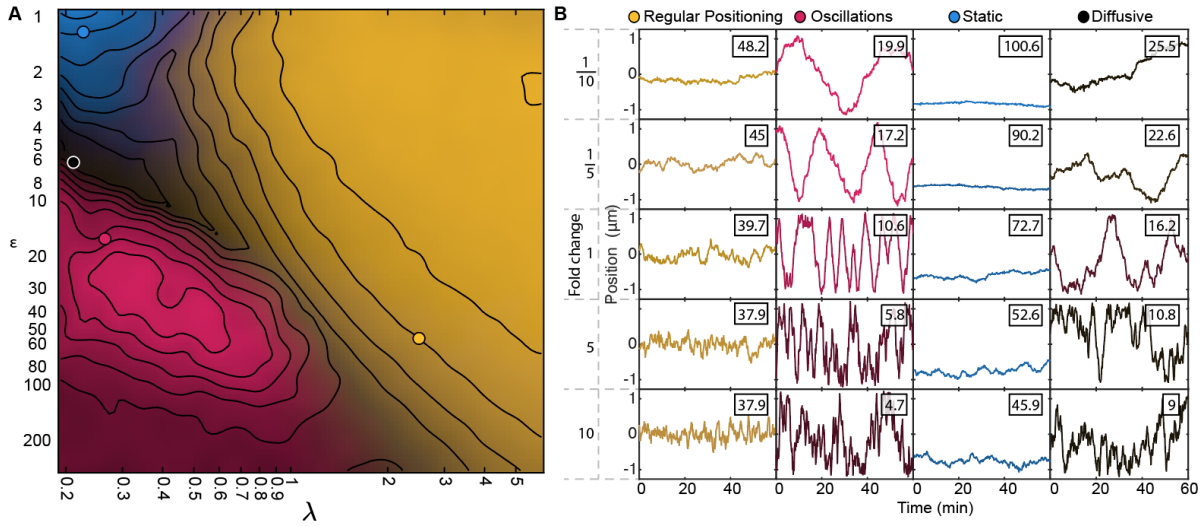

Figure 3-figure supplement 5. 1D sweeps orthogonal to the phase diagram support the role of λ and ε in defining the dynamics.

(A) Same as Figure 3I with four marked positions. (B) Example simulated trajectories at the marked positions. The parameters D_h, k_d, k_h and k_a were changed simultaneously by the indicated factor. This causes a change in the turnover rate of ParA-plasmid tethers while keeping the dimensionless quantities λ, ε and θ unchanged. The colour of each trajectory shows the classification of the dynamics at that fold change according to the colour scheme introduced in Figure 3I based on multiple long trajectories. Note the change in the frequency of fluctuations in plasmid position, consistent with changes in tether lifetime. The average number of tethers (indicated in each panel) does not remain constant because with increasing D_h each bound ParA dimer has less time between hopping events to explore its local neighbourhood through elastic fluctuations of the underlying DNA (controlled by the parameters D_h and $\sigma_{x,y}$), leading to a lower rate of tether formation. In the oscillatory regime, the increase in the plasmid speed results in a shortening of the period of the oscillations. However, at the shortest tether lifetimes, noise begins to dominate as the tether lifetime approaches the timescale of tether-induced pulling.

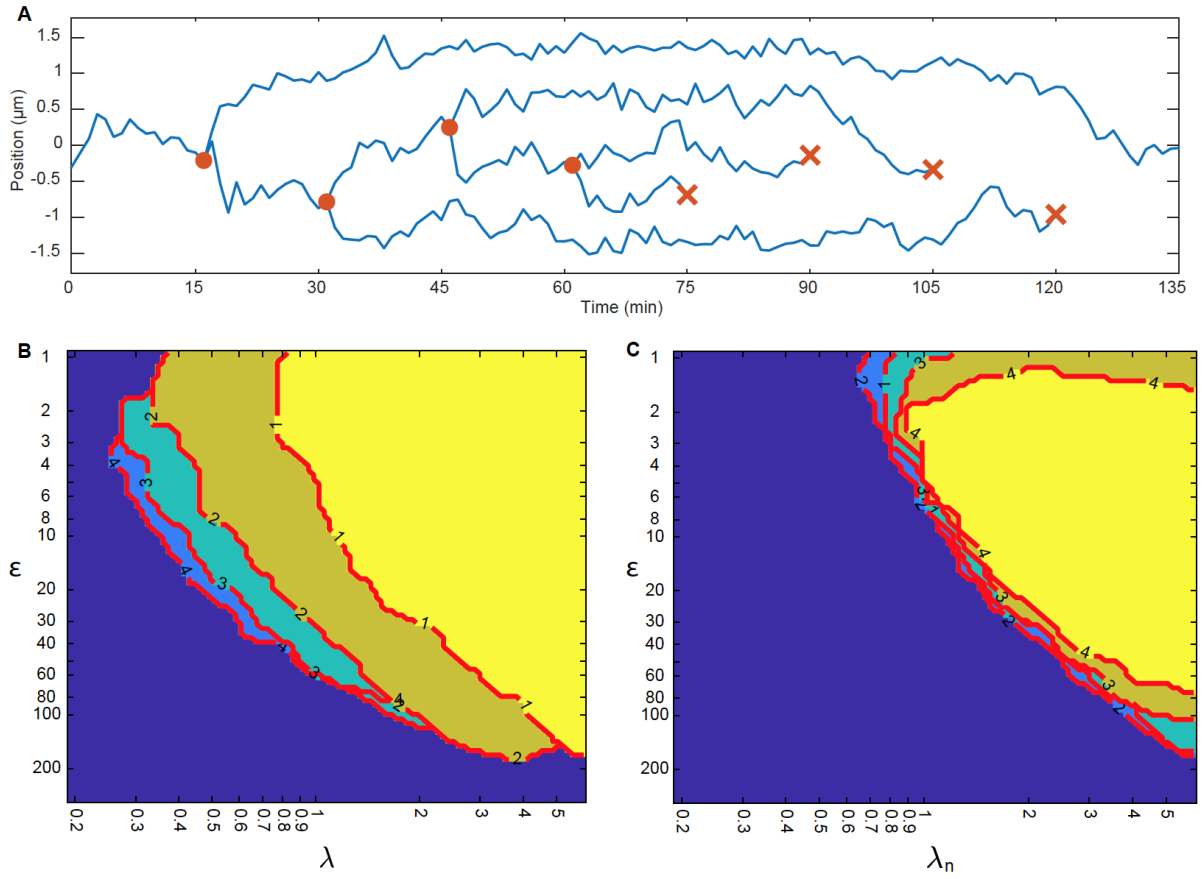

Figure 5–figure supplement 1 Regular positioning with multiple plasmids

(A) Example trajectory showing positioning relative to mid-nucleoid during plasmid replication/deletion. One plasmid is replicated every 15 minutes (orange circles) for 60 minutes. Afterwards, one plasmid is deleted every 15 minutes (orange cross). (B) Boundary of the regular positioning regime for 1-4 plasmids, determined by (1-ψ)*φ>0.7 (see methods). Number on the contour indicates the number of plasmids, n. The colours are arbitrary and for visual clarity (region on top is the brightest). (C) Same as in (B) but with a normalised λ to take account of the number of plasmids: $\lambda_n = n\lambda$. This way all regimes collapse into one location.

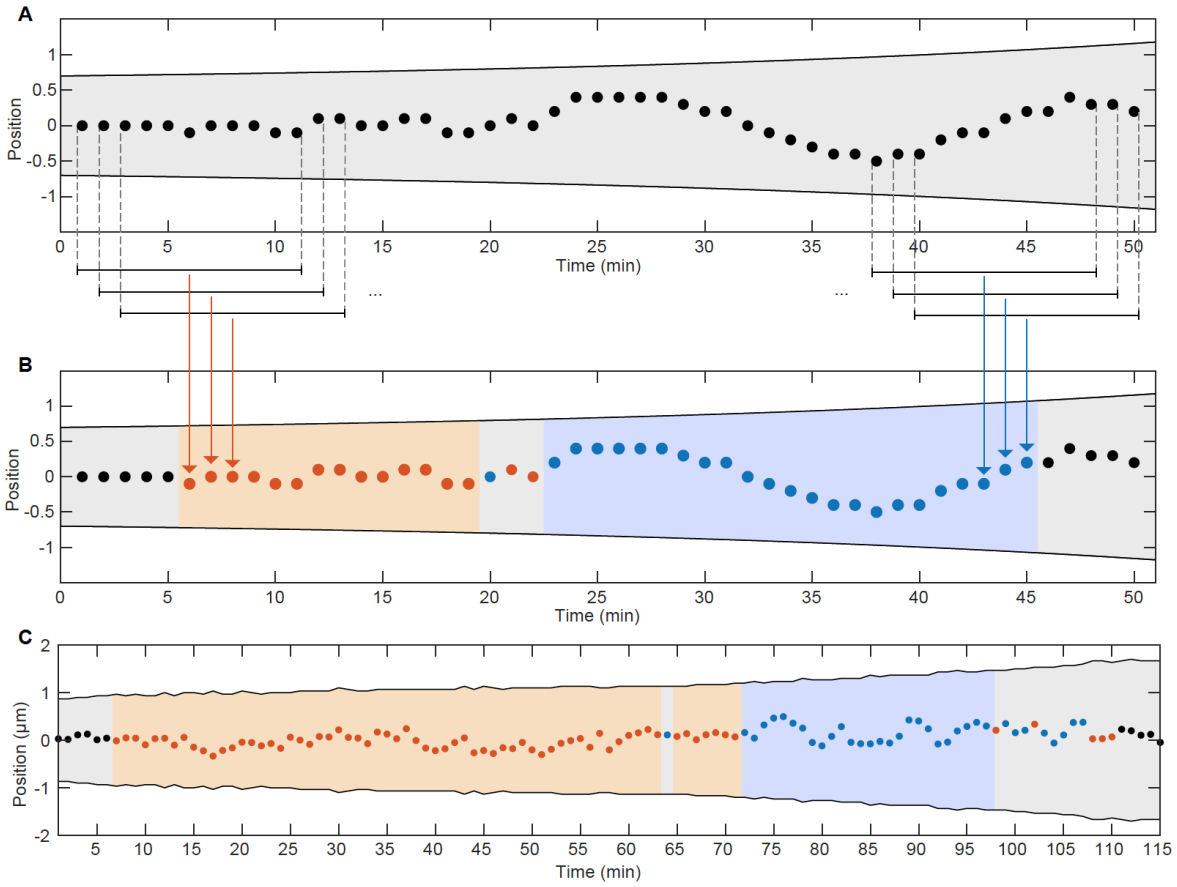

Figure 6–figure supplement 1. Classification of segments of trajectories

(A) Example trajectory (made up) of a cell containing one foci. Shaded areas indicate cell boundaries. Black bars represent a sliding window moving along the trajectory. Points which are captured by the sliding window are marked by dashed lines. (B) Same trajectory as in (A) but each point in the middle of a sliding window was colour-coded based on the sign of the velocity autocorrelation at a lag of 1 min (i.e. between consecutive frames) inside the sliding window. Orange indicates negative correlation/regular positioning and blue positive correlation/oscillations. Segments are classified (indicated by coloured shading) only if they consist of more than 6 points with the same annotation. These were used for subsequent analysis. (C) Example of an annotated/classified trajectory from experimental data.

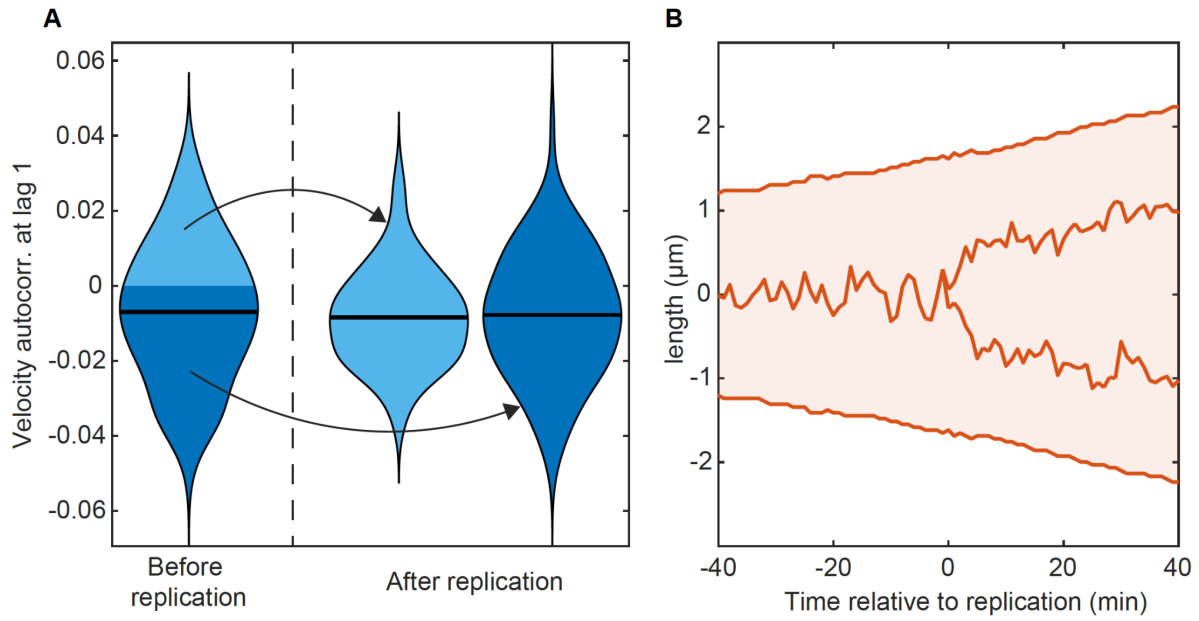

Figure 6–figure supplement 2 Oscillatory behaviour decreases upon plasmid replication

(A) A violin plot of the velocity autocorrelation at lag 1 min in the 12 frames before and after plasmid replication (one fluorescent focus becoming two). The black bar indicates the mean. Post-replication trajectories are separated according to the sign of the value before replication. Trajectories with persistent dynamics (positive velocity autocorrelation) do not have an increased bias to remain persistent after replication. (B) An example cell cycle in which oscillatory dynamics are detectable before plasmid replication but not afterwards.

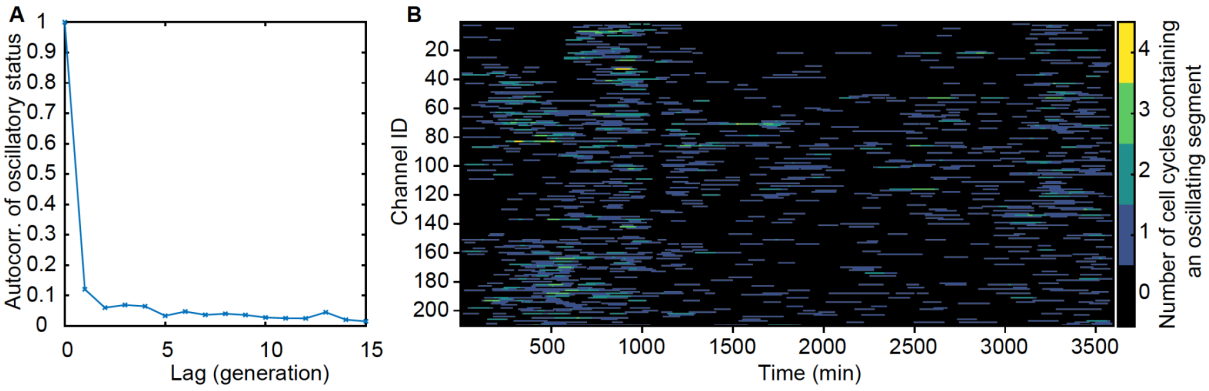

Figure 6–figure supplement 3 There is no significant inheritance of oscillatory behaviour

(A) 210 lineages of mother cells were followed over many generations to calculate the autocorrelation of oscillatory status (1 if the cycle contains an oscillatory segment, 0 otherwise) (B) The number of analysed cell cycles from each channel of the mother machine that contain an oscillatory segment.

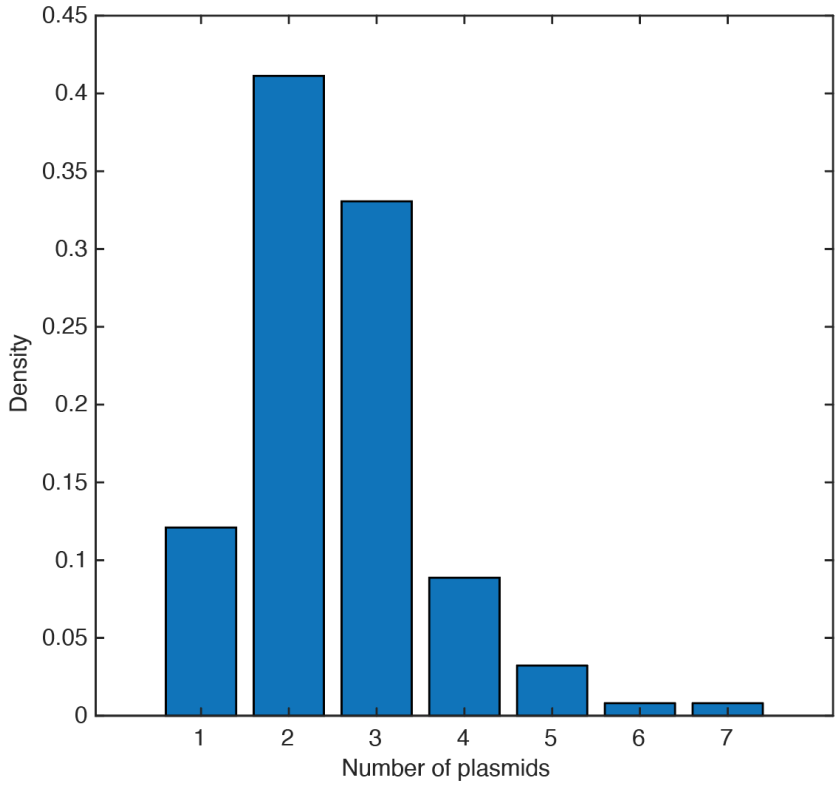

Figure 7–figure supplement 1. Distribution of copy number of pB171 as measured by the number of fluorescent foci within cells. Data from 124 cells. The mean is 2.56 plasmids per cell.

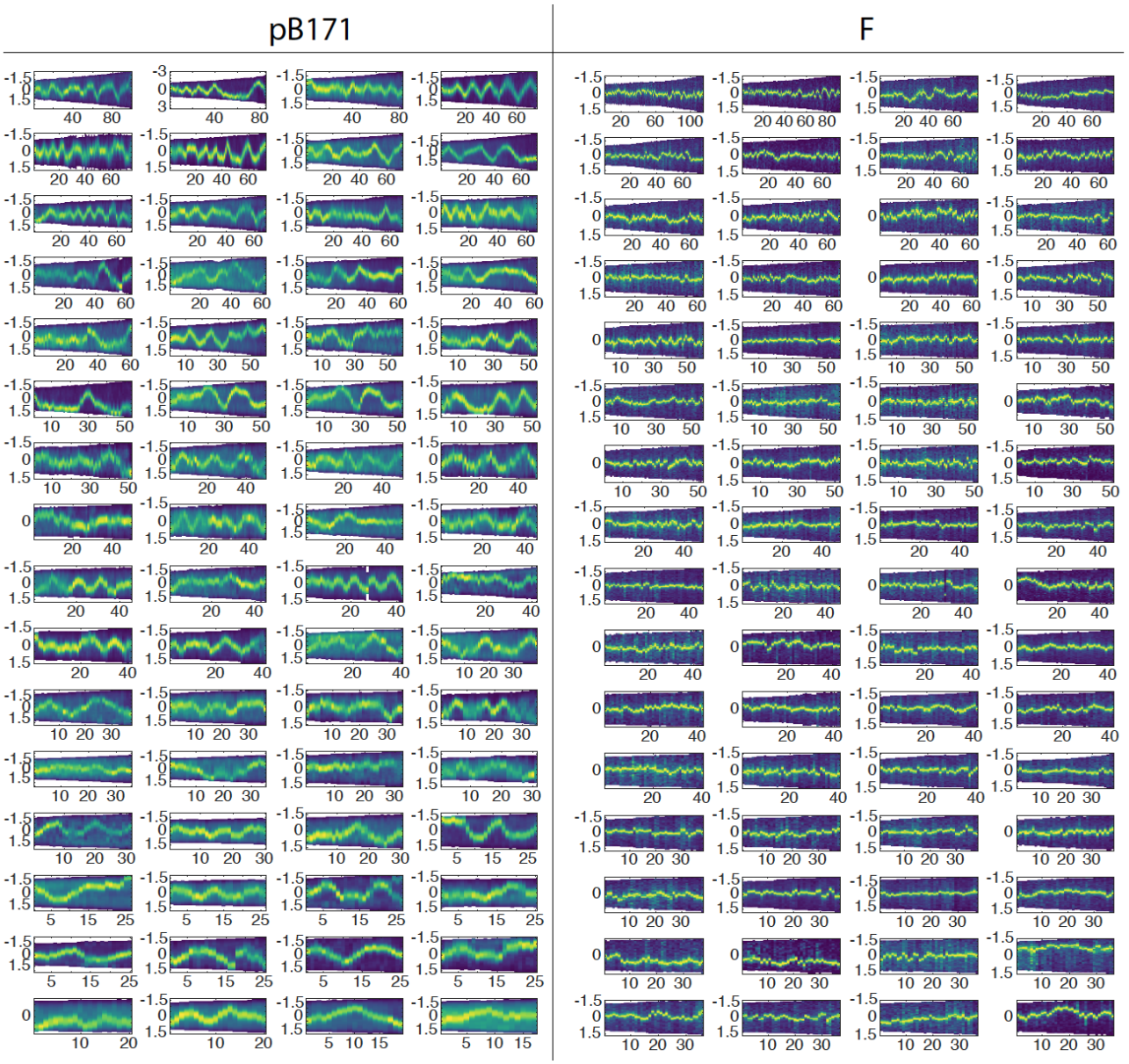

Figure 7–figure supplement 2. Comparison of pB171 and F-plasmid

Example kymographs of cells with one plasmid. Time is in minutes on the x-axis and position along the long axis in μm is on the y-axis. The intensity corresponds to TetR-mCherry (pB171) or ParB-mVenus (F).

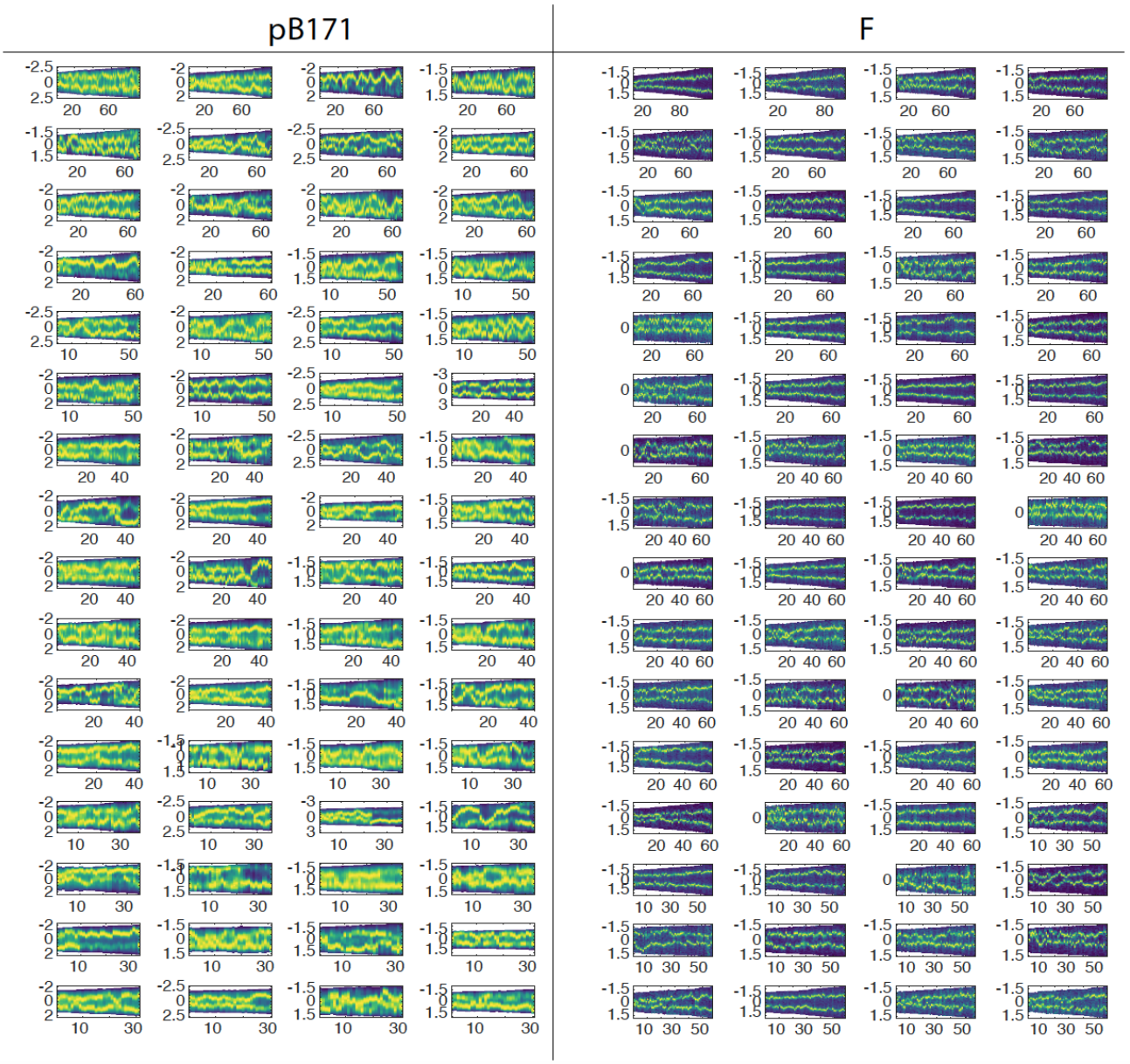

Figure 7–figure supplement 3. Comparison of pB171 and F-plasmid in cells containing two plasmids

As in Figure 7–figure supplement 2 but trajectories are from cells containing two plasmids. See Figure 7F for quantification.

4. Paper II

Microscopy, especially when labeling proteins like ParB with fluorophores, is a powerful tool for investigating systems with spatial and temporal patterns. However, tracking multiple foci simultaneously is challenging due to factors like photobleaching, high signal-to-noise ratios, and detection errors. While existing tracking software employs greedy algorithms to balance precision and computational efficiency, they often do not utilize additional system information to restrict the tracking problem (Chenouard et al., 2014; Manzo & Garcia-Parajo, 2015; Meijering et al., 2012; Shen et al., 2017). In the case of tracking plasmids, taking into account their increasing numbers during the cell cycle can significantly reduce potential combinations, simplifying the tracking process.

The second paper in this thesis introduces a tracking algorithm, called ★Track, that we developed to guarantee precision in examining the movement of plasmids in the ParABS system, especially in situations where cells contain multiple foci and traditional tracking methods are not accurate enough. The algorithm's core concept involves transforming the tracking problem into a pathfinding problem, solvable through an implementation of the A*-algorithm (Hart et al., 1968). By incorporating additional system information, such as the persistent nature of plasmids, the algorithm enables more accurate investigation not only of the ParABS system but also of replication behaviours for both plasmids and chromosomes, as precise tracking directly translates into reliable prediction of replication events.

My contribution to this paper includes all aspects of this paper apart from the experiments regarding the origin of replication of the chromosome which were performed by Ismath Sadhir and the writing which was done by Seán M. Murray and myself.

This paper was published in Biophysical Journal on the 2nd of May 2023: https://doi.org/10.1016/j.bpj.2023.03.033

4.1. ⋆Track: Inferred counting and tracking of replicating DNA loci

Robin Köhler, Ismath Sadhir, Seán M. Murray*

Max Planck Institute for Terrestrial Microbiology and LOEWE Centre for Synthetic Microbiology (SYNMIKRO), Karl-von-Frisch Straße 14, 35043 Marburg, Germany

*Email: sean.murray@synmikro.mpi-marburg.mpg.de

4.1.1. Abstract

Fluorescent microscopy is the primary method to study DNA organization within cells. However the variability and low signal-to-noise commonly associated with live-cell time lapse imaging challenges quantitative measurements. In particular, obtaining quantitative or mechanistic insight often depends on the accurate tracking of fluorescent particles. Here, we present ⋆Track, an inference method that determines the most likely temporal tracking of replicating intracellular particles such DNA loci while accounting for missing, merged and spurious detections. It allows the accurate prediction of particle copy numbers as well as the timing of replication events. We demonstrate ⋆Track's abilities and gain new insight into plasmid copy number control and the volume dependence of bacterial chromosome replication initiation. By enabling the accurate tracking of DNA loci, ⋆Track can help to uncover the mechanistic principles of chromosome organisation and dynamics across a range of systems.

4.1.2. Significance

Microscopy is one of the main tools for studying the intracellular organisation of cells. In particular, fluorescent proteins allow us to study the dynamics of many cellular processes. However, this requires the accurate tracking of fluorescent foci. Here, we present ⋆Track a tool tailored to the tracking of replicating persistent subcellular particles such as DNA loci. ⋆Track provides accurate predictions of particle copy number and replication timing even in the presence of substantial noise. The knowledge of these quantities are critical for uncovering the mechanisms behind many cell-cycle dependent processes, such the control of chromosome and plasmid replication initiation.

4.1.3. Introduction

Fluorescence live-cell microscopy is a powerful tool for the study of subcellular organisation and dynamics. However, quantitative analysis requires accurate tracking of the detected fluorescent foci,

which may visualize organelles, molecular complexes or even single molecules. The conceptually simplest approach to tracking is to sequentially join particle localisation on successive frames based on minimizing some distance metric, which could potentially account for the nature of the motion e.g. diffusive or directed. For each particle on a given frame, the distances to all particles on the subsequent frame are calculated and the pair giving the smallest distance is selected. However, at high densities particles come into proximity of one another, introducing ambiguities in the linking process. One can attempt to reconstruct all trajectories simultaneously (multi-particle tracking), potentially using information from multiple frames (multi-frame tracking) by minimising a cost function over all possible links between particles (to obtain the global solution for that cost function). Unfortunately, since this combinatorial problem scales factorially in the number of particles and exponentially in the number of frames, finding the global solution is computationally unfeasible for even tens of particles (there are $(10!)^9 \cong 10^{60}$ different ways to connect just 10 particles over 10 frames). As a result many tracking algorithms have been developed that introduce constraints and approximations to solve the particle linking problem locally and provide a best-guess prediction of the global solution[1–4]. Some of these have been released as publicly available software[5–10].

Further challenges come from permanent or transient appearance and disappearance of particles due to fluorescence photobleaching, foci moving out of the focal plane or simple detection failure. Some tracking algorithms therefore allow trajectories to terminate or begin mid-sequence under certain conditions as well as close gaps in a trajectory due to missing localisations. Merge and splitting events may also be allowed. Given these different approaches and the different implementations used, it is not surprising that no single algorithm excels at all particle tracking applications[4].

The algorithms referenced above focus on the relatively high-density regime in which there are many particles with overlapping trajectories. However, for systems with fewer but still overlapping trajectories, more global and less-approximative methods become more feasible. One such application is the study of the dynamics and replication of chromosomal loci or low-copy extrachromosomal DNA such as plasmids within living cells. Under normal circumstances, these ‘particles’ only increase in number during the cell cycle, typically by a factor of two and are persistent in that, while the fluorescent foci visualising them may appear or disappear, the particles themselves do not. This additional information places a significant constraint that can be used to optimize their tracking. At the same time, the typically low copy numbers involved mean that the problem lies towards the other extreme of the efficiency-accuracy tradeoff i.e. some of the computationally necessary simplifications required at higher numbers can be avoided. To our knowledge no tracking software is tailored to this scenario and we found that existing tools did not perform with sufficiently high accuracy. Note that despite the low copy numbers, generating the optimal tracking may not be trivial if there are many frames (e.g. there are $(2!)^{99} \cong 10^{30}$ ways to connect 2 particles over 100 frames). Furthermore false positive (i.e. spurious) and false negative detections (due to the foci moving out of the focal plane,

temporary merging or detection failure) confound the tracking and make it challenging to determine the time point at which foci are duplicated. This is especially true for bacteria due their small size and the smaller pool of fluorescent protein. However, since DNA loci are persistent, both false positive and false negative detections can be more confidently identified.

Here, we present ★Track (pronounced 'star-track')[7] a tool for the accurate tracking of replicating chromosomal and extrachromosomal DNA, and other low-copy number persistent particles, from time lapse images of live cells. The algorithm is transparent and user friendly and is completely specified, in the default case, by only four input parameters (there are no hidden internal parameters). Matlab and Python implementations are provided so that it can be easily integrated into existing spot detection pipelines. We use ★Track to study the timing of plasmid replication within single cells, something which has not previously been achieved[11], and gain insight into the mechanisms of plasmid copy number control. We also investigate chromosome replication initiation in *Escherichia coli*.

4.1.4. Materials and Methods

4.1.4.1. Strains and growth condition

F plasmid experiments use strain DLT3125[12], a derivative of the *E. coli* K-12 strain DLT1215[13] containing the mini-F plasmid derivative pJYB234. This plasmid carries a functional ParB-mVenus fusion. Overnight cultures were grown at 37°C in LB-Media supplemented with 10 μg/ml thymine + 10 μg/ml chloramphenicol. The strain IS130 was constructed by transduction of matP-YPet from RH3[14] and oriC-$parS_{P1}$ from strain RM3[15,16] into *Escherichia coli* K-12 MG1655 strain (lab collection) which was then transformed with pFHCP1-mTurquoise2 plasmid. The plasmid pFHCP1-mTurquoise2 was derived from plasmid pFHC2973 by deletion of ygfp-$parB_{pMT1}$ and replacement of CFP with mTurquoise2[15,17]. Overnight cultures were grown in M9 minimal media supplemented with 0.2% glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$ and 0.5 mg/mL BSA.

4.1.4.2. Microfluidics

A mother machine[18] is a microfluidic device consisting of long narrow growth channels, the width of a single bacterial cell, connected to a main channel through which nutrient media flows. Cells can be trapped in the growthing channels and monitored over many generations. Our device is broadly similar to the original. The growth channels are 35 μm long and are organized into blocks of 30. Channel widths vary between 0.8-1 μm. However, we follow Baltekin et al.[19] and also include i) a small opening at the end of each growth channel ii) a waste channel connected to that opening to allow a continuous flow of nutrients through the growth channels iii) an inverted growth-channel that is used to remove the background from fluorescence and phase contrast. Before imaging, the chip is bound to a glass slide using a plasma generator and baked for 30 minutes at 80°C.

[7] As well as referring to the A* algorithm, we are inspired by the quote of Captain Kirk from Star Trek: "*There's no such thing as the unknown, only things temporarily hidden*".

4.1.4.3. Microscopy

We used a Nikon Ti microscope with a 100x/1.45 oil objective and a Hamamatsu Photonics camera for all imaging. For imaging cells of strain DLT3125 and IS130 we used a mother machine. Overnight cultures of DLT3125 were inoculated into fresh media (M9 + 0.5% glycerol + 0.2% casamino acids + 0.04 mg/mL thymine + 0.2 mg/mL leucine + 10 μg/mL chloramphenicol) for 4 hours at 30°C before imaging. For the strain IS130, 50 μM IPTG (for induction of mTurquoise2-$ParB_{P1}$) was added to the media defined in the 'Strains and growth condition' section 2 hours before and during the experiment. Cells were loaded into the chip through the main channel and the chip was placed into a preheated microscope at 30°C. The cells were constantly supplied with fresh media by pumping 2 μl per minute through the microfluidic chip. Cells were grown for at least 2 hours inside the microscope before imaging. DLT3125 was imaged at 1 minute intervals and IS130 was imaged at 5 minute intervals for approximately 72 hours. Both phase contrast and fluorescent signal were captured.

4.1.4.4. Image processing

Our image processing pipeline, Mothersegger (https://gitlab.gwdg.de/murray-group/MotherSegger) has been described previously[20]. Briefly, it consists of four parts: I) preprocessing, II) segmentation III) cell tracking and IV) foci detection. Parts I, III and IV use custom Matlab scripts, while Part II is based on SuperSegger[21], a Matlab-based package for segmenting and tracking bacteria within microcolonies, that we modified to better handle high-throughput data. In Part I each frame of an acquired image stack is aligned (the offset between frames in x and y is removed). Afterwards the image stack is rotated so the growth channels are vertical. A mask of the mother machine layout is fitted to the phase contrast, using cross-correlation, to identify where the growth channels are located. Each growth channel is extracted from the image stack and the flipped inverted channel is subtracted to remove the background from both the fluorescence signal and phase contrast. In Part III the cells are tracked. Since cells cannot change their order inside the growth channel, they can be tracked by matching similar cell length between frames (starting from the bottom of the channel). Cell cycles that do not have exactly 1 parent and 2 daughters are excluded from analysis along with their immediate relatives (with the exception of who are pushed out of the growth channel). Part IV detects fluorescent foci within cells. It is based on the SpotFinderZ tool from Microbetracker[22].

4.1.4.5. Computer generated trajectories and manipulation

The trajectories used in Figure 1c-f were generated with our previously published model of plasmid positioning[20]. This is a stochastic model of plasmid positioning by the interaction of plasmid-bound ParB with nucleoid associated ParA. These trajectories were used as ground truth. As stated in Figure 1 we manipulated the ground truth by adding (false positives) and removing (false negatives) a percentage of foci. If the ground truth contains 100 foci and 10% false positives are added and 5% false negatives are removed the resulting modified data set contains 105 foci (100 + 10 fp - 5 fn).

4.1.4.6. A* algorithm

The first step of ★Track is based on the A* graph-traversal algorithm[23], which we first summarize. A graph consists of nodes connected by edges that have a weight corresponding to the cost of traversing that edge. The cost of a path on the graph is the cumulative cost of all its edges. Given a start and target node of a graph, the A*-algorithm finds the shortest path between them. A path is considered the shortest if there exists no other path connecting these two nodes with a lower cost. The algorithm begins by initializing a candidate list of partial paths consisting of only the first edge (the edges of the starting node). The key feature of the algorithm is the generation of a lower bound for the cost of completing a partial path. The determination of this bound is problem specific. The algorithm then selects the first edge that gives the lowest lower bound for the total cost (the cost of the first edge plus the lower bound for completing the path starting from that edge). In the second iteration, all possible choices of second edge (with the chosen first edge) are added to the candidate list and again the best second edge is selected based on the cost estimate. This process repeats until either the target node is reached or the estimated cost of the current best partial path exceeds the estimated cost of completing a partial path further up the graph, e.g. a path with a different first edge. In this case, the process then continues from this different first edge. In this way the algorithm eventually finds the path with global minimum cost.

4.1.4.7. Cost of linking two foci

Use of the A* algorithm requires the specification of a cost for linking focus i on one frame to focus j on a later frame. For this, we assume that foci move diffusively between frames and consider the probability that the 2D distance traveled would be at least as great (the cumulative distribution function) as that observed

$$P(d, t, D) = e^{-d^2/(4Dt)}$$

where d is the 2D spatial distance between the foci, t is the number of frames between foci and D is the diffusion coefficient. The movement cost is then $c_{i,j} = -\ log(P(d, t, D))$. An alternative model can be used if required. If the two foci are not from consecutive frames then we add a cost $c_{neg} = -\ log\left(p_{neg}\right)$ for each of the intermediate $t - 1 \leq m$ frames, where p_{neg} is the probability of a false negative and m is a user-provided upper bound for the number of frames on which the focus was missing. Below, we define a cost threshold for acceptable links that is used in the first part of the algorithm. Note that if $p_{neg} = 0$ tracking cannot contain false negatives as the tracking would otherwise fail due to having infinite cost.

4.1.4.8. Layered Graph

The second requirement of the A* algorithm is that the search is between start and target nodes. On the other hand, the initial tracking consists of time-directed connections between foci that need not form a connected path. We therefore developed a mapping of the temporal tracking problem to a radial layered graph structure on which the A* algorithm could be applied.

We illustrate this with an example in Figure S1. Panel S1a shows an example data set consisting of $n_{foci} = 6$ foci localisations across $t_{end} = 5$ frames. We first determine the set of acceptable links between these foci. Acceptable links are those with a cost below the threshold $c_{th} = min(-\ m\ log\left(p_{neg}\right),\ c_{max})$ based on the user provided maximum number of frame skips, m and the false negative probability. This threshold ensures no link is allowed between foci separated by more than m intermediate frames, as well as very spatially distant links across fewer frames. It corresponds to spatial-temporal search radius for linking foci. We also include an upper bound on the threshold $c_{max} = -\ log(0.0001)$ to account for when $p_{neg} = 0$ (no false negatives allowed) in which case the threshold would otherwise be infinite. Note that this threshold is only used in this step of the algorithm and not in the subsequent steps (stitching/shuffling).

Given these acceptable links, we can then create the layered graph. We first label foci sequentially starting from one of the potentially multiple foci on the first frame (Fig. S1a). The corresponding graph contains one layer for each focus plus an additional outermost layer to determine when the tracking is complete (Fig. S1b). Every path from the first to the last layer corresponds to one possible tracking of the foci in Figure S1a. The relationship to a tracking is as follows. An edge from **any node** in **layer** *i* into a numbered **node** *j* in layer $i + 1$ corresponds to an acceptable link between focus i and focus j (Fig. S1 c & d). An outgoing edge into a node labelled 'x' corresponds to focus i having no outgoing link. Note that the layer into which an edge goes, does not matter for the interpretation, nor the number of the node it came from. A path from layer 1 to the outermost layer crosses each layer once, therefore each layer (except the outermost) has exactly one outgoing edge. This ensures that each focus has at most 1 outgoing link. Furthermore, each path contains each numbered node at most once. If a path traverses a numbered node i, node i will not appear again in the higher layers. This ensures that each focus has only one incoming link. The construction of the graph is illustrated in Supplementary Video 1.

Each edge in the graph has a weight (referred to but not shown in Fig. S1b) which corresponds to the cost of linking the foci associated with that edge, as described above. An edge going from layer i into a numbered node j has a cost $c_{i,j}$. An edge from layer i going into a node labelled 'x' has a cost equal to $c_{i,x} = min(c_{th}, -\ log\left(p_{neg}^{\ t_{end} - t_i}\right))$ where $t_{end} - t_i$ is the number of frames between that of focus i, t_i and the last frame t_{end}. The cost of a path is the cumulative cost of its edges and the optimal tracking is the path from the root node in layer 1 to the outermost layer with the lowest cost.

To use the A* algorithm we require a heuristic lower-bound estimate of the cost of reaching the outermost layer from each node in the graph. A path from layer i has to traverse all subsequent layers to reach the outermost layer. A lower bound for the cost is then the sum of all the lowest costing outgoing links $c_{est}(i) = \sum_{j=i}^{n_{foci}} min\left(\left\{c_{j,\cdot}\right\}\right)$ where the minimum is over the cost of all acceptable outgoing links from layer (focus) j (including to the 'x' node).

Since the number of nodes in each layer increases exponentially, it is infeasible, and also unnecessary, to pre-generate the entire graph when tracking more than a few foci over a small number of frames. Therefore the

graph is generated dynamically as the A* algorithm traverses it. An example can be seen in Supplementary Video 2. Note that the optimal tracking is found without generating the entire graph. To further reduce the complexity, the candidate list is filtered once a new layer is reached. All candidate paths ending $1 + max\left(\left\{j - i : t_j - t_i \leq m\right\}\right)$ layers below the newly reached layer are removed from the candidate list. Therefore, the greater the number m of allowed frame skips (consecutive false negatives), the more of the graph is retained. Additionally, if the candidate list exceeds 10^6 entries, the 10% of partial paths with the highest cost estimates are removed. While formally making the algorithm approximative, in practice this threshold has little effect on the accuracy of the results (Fig. S2a) and can be reduced to as low as 10^5 in order to reduce the memory usage of the program and reduce runtime (Fig. S2b).

4.1.4.9. Stitching

The goal of this procedure is to integrate the trajectory segments returned by the A* algorithm such that the number of foci monotonically increases. This is done in 3 steps. First, loose beginnings (foci which have no incoming link) are linked to earlier foci (Fig. S3b), either to a real focus according to the lowest cost of the link or to an imaginary focus before the first frame. The latter is for implementation reasons only.

In the next step, loose ends (foci which have no outgoing link) are integrated into the tracking (Fig. S3d). This occurs in three different ways: (1) a loose end can be linked to an imaginary focus after the last frame (Fig. S3c ii & iii), impling some number of missed foci (false negatives) when the loose end is not on the last frame; (2) all the foci of the segment containing the loose end up to any branching point are considered as false positives; (3) a loose end can be interwoven into the tracking by replacing an existing link, which can but does not have to be a link created by integrating a loose beginning (Fig. S3c iv & v). This may create a different loose end, which is treated as in (2). Which method occurs is determined by the cost of the entire tracking.

The cost of the tracking is calculated differently than in the A* step as it incorporates the concept of false positives and applies to the entire tracking and not just disconnected segments. Furthermore, the threshold cost (c_{th}) of the A* step is no longer used. Each false positive focus has a cost $c_{pos} = -\ log(p_{pos})$, where p_{pos} is the false positive probability. Each real focus then has a cost arising from its probability to not be a false positive $c_{real} = -\ log(1 - p_{pos})$. The cost of a tracking is then $c_t = n_{fp}c_{pos} + n_{real}c_{real} + n_{fn}c_{neg} + \sum_{links\ i \to j,\ for\ real\ foci\ i,j} c_{i,j}$, where n_{fp} is the number of false positives, n_{real} is the number of detected foci and n_{fn} is the number of false negative foci. Note that if $p_{pos} = 0$ false positives are not allowed and the $n_{fp}c_{pos}$ term is not included in the total cost.

Returning to the integration of loose ends, the change in the cost of the tracking is calculated for every possible way to integrate a loose end and the method with the lowest cost is chosen. The procedure is iterative starting at the last loose end. It continues until there are no loose ends left or the remaining loose ends consist of false positives.

The third and final step is simply the removal of the foci identified as false positives from the tracking (Fig. S3e).

The result of this procedure is a tracking in which every focus is path connected to the first and last frames and the number of foci never decreases. All the steps above were developed with this goal in mind. Connecting loose beginnings first and then handling loose ends was found to be the best approach for stitching together trajectories produced by the A* algorithm, compared to connecting loose ends first or doing both simultaneously. This is because of disconnected fragments, which possess both a loose end and loose beginning. Multiple of these fragments might be a part of a long branch but remain unlinked due to the cost threshold during the A* step. This is problematic because these fragments might be labeled as false positives and removed, instead of being incorporated as a single long branch. To circumvent this issue, loose beginnings are connected first, which links all disconnected fragments to the tracking and provides the necessary context for making the most appropriate decisions how and where to incorporate loose ends.

4.1.4.10. Shuffling

Shuffling resolves remaining inaccuracies in the tracking after the stitching procedure. Inaccuracies sometimes appear in the output of the A*-algorithm due to false negative foci. For example, the crossing links in the third panel of Figure 1b is due an inaccurate trajectory segment from the A* step, which in turn is due to multiple consecutive missed foci causing the upper segment in the top panel to terminate. The most straightforward way to resolve these types of inaccuracies is by shuffling links as follows. Two links connecting four foci, focus i linked with focus j and focus n linked with focus m, have a combined movement cost of $c_{i,j} + c_{n,m}$. The links between the foci can be shuffled, subject to $max(i, m) < min(j, n)$, such that focus i is linked with focus m and focus n is linked with focus j. If the new cost of $c_{i,m} + c_{n,j}$ is lower than $c_{i,j} + c_{n,m}$ then the new configuration is accepted. Foci are shuffled until there are no pairs of links left that can be shuffled to reduce the cost of the tracking (Fig. 1b bottom). A limitation of this procedure is that it can only find an improved tracking that is obtainable by shuffling two links. Improved trackings obtainable through more complicated rearrangements cannot be identified.

4.1.5. Results and Discussion

★Track operates on a provided list of foci localisations (t, x, y), consisting of the frame number and location of the fluorescent foci within a cell or region of interest (corrected for movement and growth as required). The algorithm requires a cost function that specifies, for any given focus, how likely it is that another focus on the next or some later frame is the next detection of the same underlying particle (a higher likelihood results in a lower cost). The latter case implies that the same focus was not detected on some intermediate frames and ★Track requires the user to specify the maximum number of consecutive frames for which this is allowed to occur (‘not detected’ includes foci that have merged to within the diffraction limit). This naturally defines a minimum spatio-temporal distance (cost threshold) between foci that greatly increases the computational efficiency of the first step of the

algorithm. In subsequent steps connections greater than the threshold are allowed provided they improve the overall tracking.

The default cost function is relatively simple and uses only the positions of foci. However, it could easily be modified to incorporate the spot intensity, goodness of fit or specific models of movement. We have found for the applications studied that spot intensity and shape can vary considerably between frames due to stochastic variation and movement out of the focal plane and we therefore base the cost function only on the foci positions. The default function assumes that foci move diffusively between frames (this is always true on a short enough timescale) and requires only the diffusion coefficient D. The user must also specify p_{fn}, the probability of a spot not being detected on a given frame (false negative) and p_{fp}, the probability of a detected spot being spurious (false positive). These three parameters can be estimated by analyzing cells with a sufficiently low number of foci and no splitting events (Fig. S4). Precise values are not required as the algorithm is robust with respect to these parameters (Fig. S2 c & d).

The pipeline consists of three steps (Fig. 1a). First, the tracking problem is formulated as a pathfinding problem on a layered graph (Fig. S1). The A* algorithm[23] then uses the cost function, with a threshold cost derived from the max number of frame skips, to find the optimal path on this graph and thereby generate candidate trajectory segments (Fig. 1b second panel, Fig. S1). The resulting trajectories are then stitched or interwoven together in such a way that the total cost of the tracking is minimised. Afterwards, each focus is either marked as a false positive and removed or is part of a trajectory that can be traced from the first frame to the last (Fig. 1b third panel, Fig. S3). Finally, a lower cost tracking is searched for by shuffling links in the tracking (Fig. 1b bottom). This helps to mitigate inaccuracies in the first step due to false negative/positive foci. The contributions made by each step of the algorithm are given in Figure S5. Note that ★Track does not predict the location of missing foci, but simply their existence. A detailed description of the A* algorithm and the stitching and shuffling steps can be found in the Methods section.

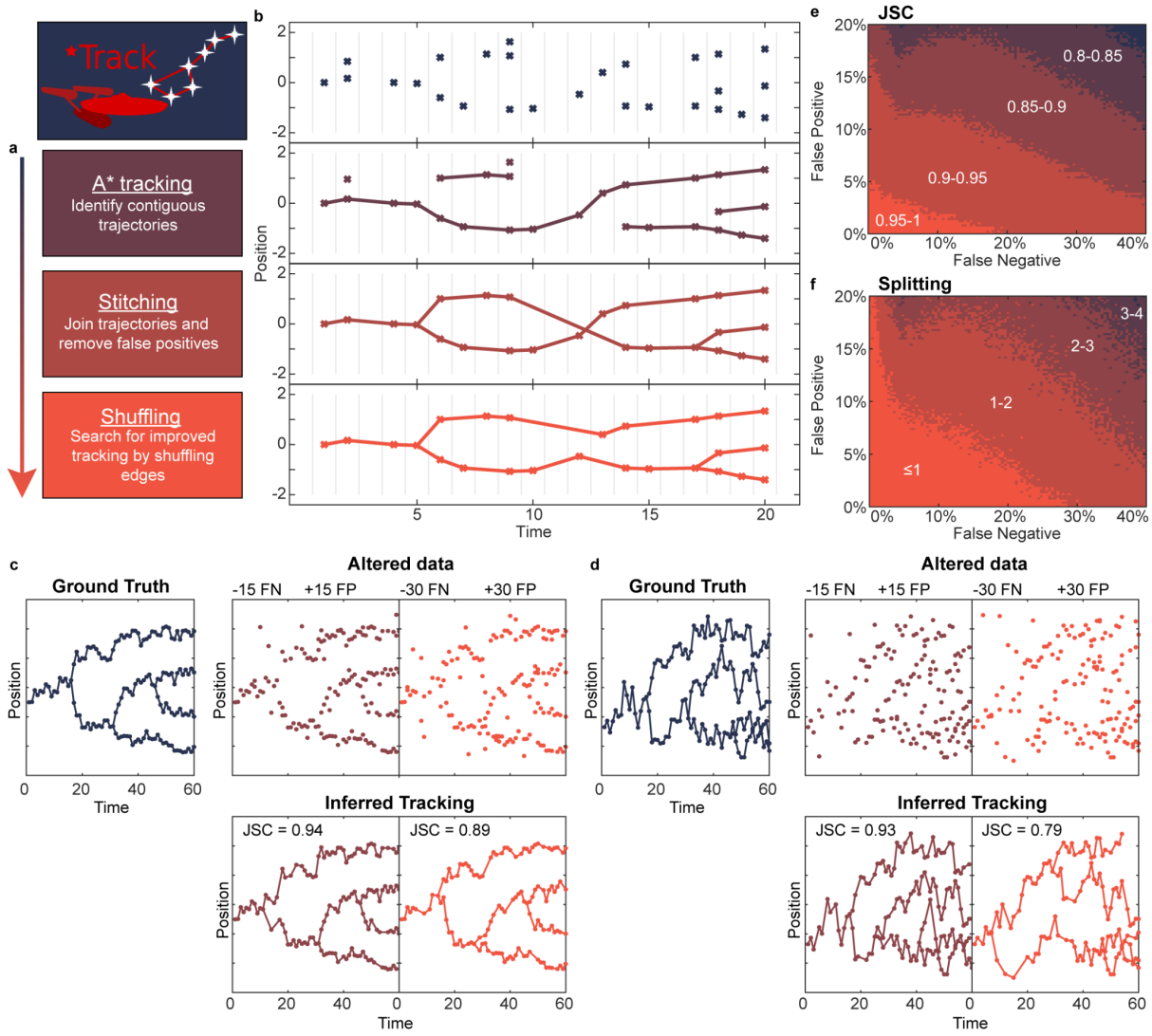

Fig. 1: Overview and accuracy of ★Track. **a**, All main steps of the algorithm with a short description. The arrow indicates the order. **b**, Example timelapse of localisations and how the data looks after each step in **a**. **c**, A simulated trajectory (150 foci) is used as ground truth to measure the accuracy of ★Track. The ground truth is altered by adding a number of spurious foci (false positive, FP) and removing a number of real foci (false negative, FN) (panels labeled Altered data). ★Track was then used to infer a tracking of the altered data (panels labeled Inferred). Accuracy is quantified using the Jaccard similarity coefficient (JSC) .The JSC is equal to the number of real foci included in the tracking divided by the sum of the number of spurious foci in the tracking and the total number of real foci in the altered data. The JSC is between 0 and 1, where 1 is a perfect score (the tracking contains all the real and no spurious foci) and 0 is the worst score (no real foci included in the tracking). **d**, Same as **c** but for more mobile particles. **e**, JSC for a wide range of false positive and false negative rates. At each point, 100 trajectories were generated as in **c**. **f**, Same as in **e** but for splitting accuracy. The number indicates the mean absolute deviation of the splitting event in frames from the ground truth.

To assess the accuracy of ★Track, we generated ground truth data using a stochastic model of plasmid positioning[20] and mimicked the presence of false positive and false negative foci by randomly adding and removing data points respectively. We found that even after removing a substantial fraction of

localisations and adding many false positives, ★Track could accurately reproduce the ground truth tracking including correctly identifying the timepoints of plasmid replication (Fig. 1 c,d). Performing a sweep over a range of false positive and false negative rates, we found that ★Track performs with consistently high precision (Fig. 1e,f) and outperformed the state of the art and widely employed method u-track[5,6,24] (Fig. S6).

To test ★Track on real data, we applied it to the study of plasmid copy number control. This has been a topic of active research for several decades[25–29] but progress has more recently slowed due, in part, to the inability to accurately measure plasmid copy numbers in individual cells[11]. Having recently performed a high-throughput study of the partitioning mechanism of the low copy number F plasmid using a microfluidic 'mother machine' device[20], we decided to revisit our data in the context of copy number control. F plasmid is a tractable system in this regard since duplicated plasmids segregate beyond the diffraction limit within about a minute of replication[30,31]. We expected that ★Track should therefore be able to accurately determine the temporal changes in copy number during the cell cycle.

In Figure 2a we present an example timelapse showing how the number of detected foci (of plasmid-labeling ParB-mVenus) changes substantially from frame to frame (see Fig. S7 for more examples). As discussed above this can occur due to foci moving out of the focal plane, merging together or simply due to stochastic fluctuations in the number of plasmid-bound ParB-mVenus. However, ★Track inferred a consistent tracking that produces a stepwise increasing copy number and predicts the number and frame of the replication events (Fig. 2a). We then applied the algorithm to the entire data set of 4096 cell cycles and examined how the copy number changes between birth and division. The raw data displayed several irregularities such as cells having no detected plasmids at birth and cell cycles in which the number of plasmids decreased (Fig. 2b). All of these inconsistencies were removed by ★Track (Fig. 2c), which also inferred a mean copy number at birth ~12% greater than that of the mean number of detected foci. We found a decreasing linear relationship between the number of plasmids gained during the cell cycle and the number at birth (Fig. 2d), indicating a 'sizer'-like mechanism for copy control[32]. Interestingly, the mean number of plasmids gained did not depend on the growth rate of the individual cell (Fig. 2e), indicating that plasmid production is coupled to the growth rate of the host so as to produce the same number of plasmids irrespective of the cycle duration.

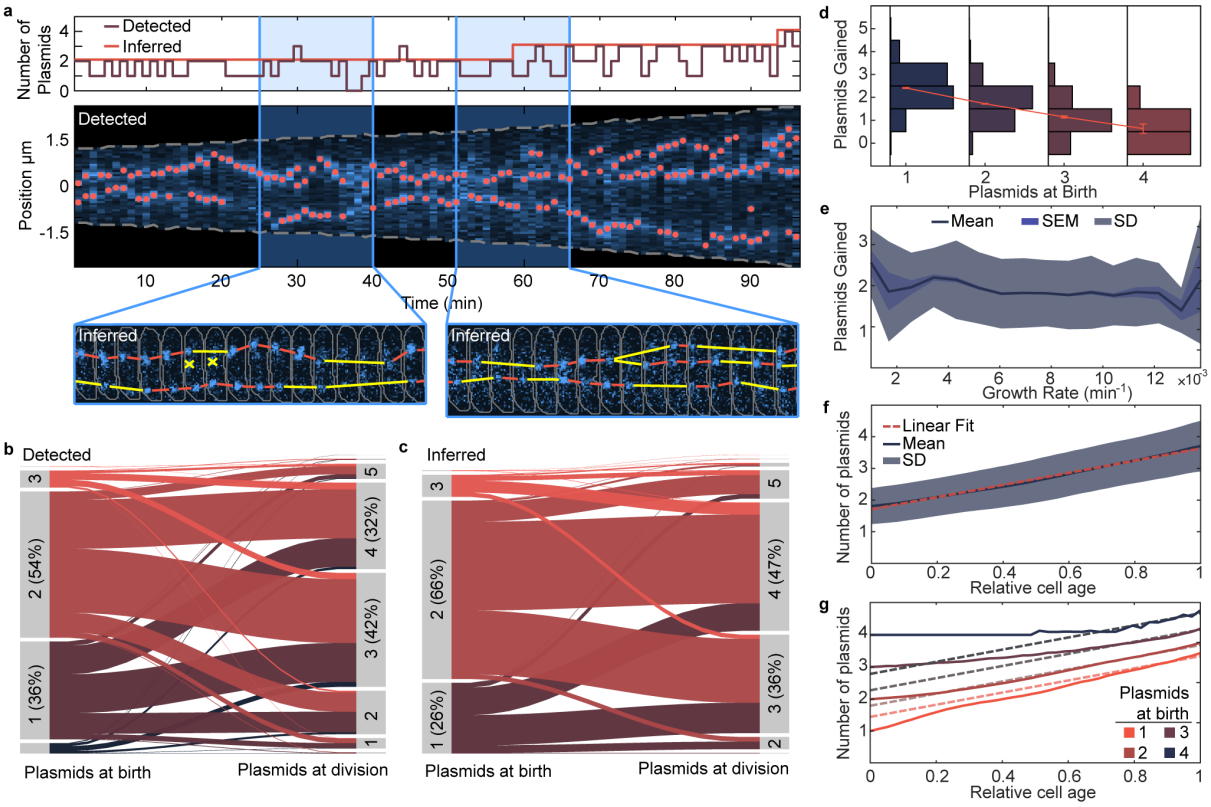

Fig. 2: Tracking F-plasmid. **a**, Example cell cycle of the *E. coli* strain DLT3125, hosting an F-plasmid encoding a mVenus-ParB fusion which binds and labels the plasmid. Top: detected and inferred number of plasmids during the cell cycle. Middle: Kymograph of the fluorescent signal of the example cell with detected plasmid locations (red dots). Gray dashed lines indicated cell boundaries. Bottom: Detailed view of the indicated regions. Red and yellow lines indicate the inferred tracking. Yellow lines indicate frames in which missing foci were inferred (false negatives). Red cross indicates foci identified as false positives. Gray lines indicate cell contours. **b**, Alluvial plot showing how the number of plasmids changes from birth to division in the raw data. Frame rate is 1 min^{-1}. **c**, Same as **b** but with the inferred data. **d**, Side-ways histograms of plasmids gained during a cell cycle plotted against the number of plasmids at birth. The redline depicts the mean number of plasmids gained for each group ± the standard error of the mean (SEM). **e,** plasmids gained during a cell cycle plotted against the growth rate ± SEM and standard deviation (SD). **f**, Mean number of plasmids ± SD plotted against relative cell age (0 is birth and 1 is division). The red dashed line is a linear fit. **g**, The mean number of plasmids grouped by the number of plasmids at birth plotted against relative cell age. The dashed lines are linear fits with the same slope as in **f** with the intercept chosen by fitting to the last portion of each group.

★Track allows us to probe the timing of plasmid replication within the cell cycle. We found that the mean number exhibits a surprising linear relationship with time (Fig. 2f). This is consistent with the net replication rate being constant in time i.e. plasmids are produced at the same rate irrespective of how many are present. However, the explicit nature of the sizer regulation could be seen by binning cell cycles according to the number of plasmids at birth (Fig. 2g). Cells born with fewer/more plasmid than average have a higher/lower net plasmid replication rate initially before returning to the population mean production rate towards the end of the cell cycle. This results in outliers converging rapidly to the mean. While negative feedback has long been known to underlie plasmid copy number

control[25–29], we believe that this is the first time it has been characterized at the level of the cell cycle. Furthermore, our result, that the effect of the regulation is to push the system back to a constant net replication rate, can now be used to test models of copy number control and motivate further study.

Next we used ⋆Track to analyse the replication of the origin region of the *Escherichia coli* chromosome. We used a strain in which the origin of replication (*ori*) is visualized through the P1 *parS*/ParB labeling system using an mTurquoise2-ParB fusion and imaged several thousand cell cycles using the same mother machine device as we used for studying F plasmid. As for that case, foci are not always visible or detected and spurious foci can confound interpretation (Fig. 3a, see Fig. S7b for further examples). This makes it challenging to determine the time of duplicated *ori* separation with certainty. Recent studies have used a replisome reporter to identify initiation events[33,34], which occur approximately 15 min before separation of *ori* foci[35]; however similar detection issues occur. These irregularities could be seen by comparing the number of *ori* on the first and last frames of the cell cycle (Fig. 3b). Significant cell populations have no detected foci at birth or only one focus at division. However, ⋆Track corrected almost all these inconsistencies (Fig. 3c).

Surprisingly, given the slow-growth conditions used (median doubling time of 140 min), we found that 14% of cells were born with two *ori* (this is a lower bound since a single focus could consist of two unsegregated *ori*). Plotting the number of *ori* against cell age revealed two periods of the cell cycle during which *ori* separation, and hence, replication initiation occurs (Fig. 3d and Fig. 3e). The majority population, cells born with a single *ori* focus, replicated their origin in the first third of the cell cycle. However in about 15% of these cells additional replication events occurred toward the end of the cell cycle producing pre-divisional cells with three or four *ori* foci (Fig. 3e,f). Cells born with two foci only initiated replication in this later period and never at the beginning of the cell cycle, though the majority of such cells do not initiate replication at all and so produce daughter cells with one *ori* focus each.

While analyzing replication initiation in terms of relative cell age is useful, the prevailing understanding is that chromosome replication initiates at an invariant cell volume per *ori*[33,34,36]. Consistent with this, we found that *ori* focus duplication in cells with a single *ori* occurred at a cell volume of about 0.8 μm^3, whereas in cells with two *ori* it occurred at twice this volume (Fig. S9). As such, there is really only one replication period defined by the volume per *ori*, but this period overlaps the division event and therefore we are able to observe two different values for the volume at initiation in the same population. This supports previous evidence for the invariant volume per *ori* hypothesis, which has been largely based on comparing populations under different growth conditions.

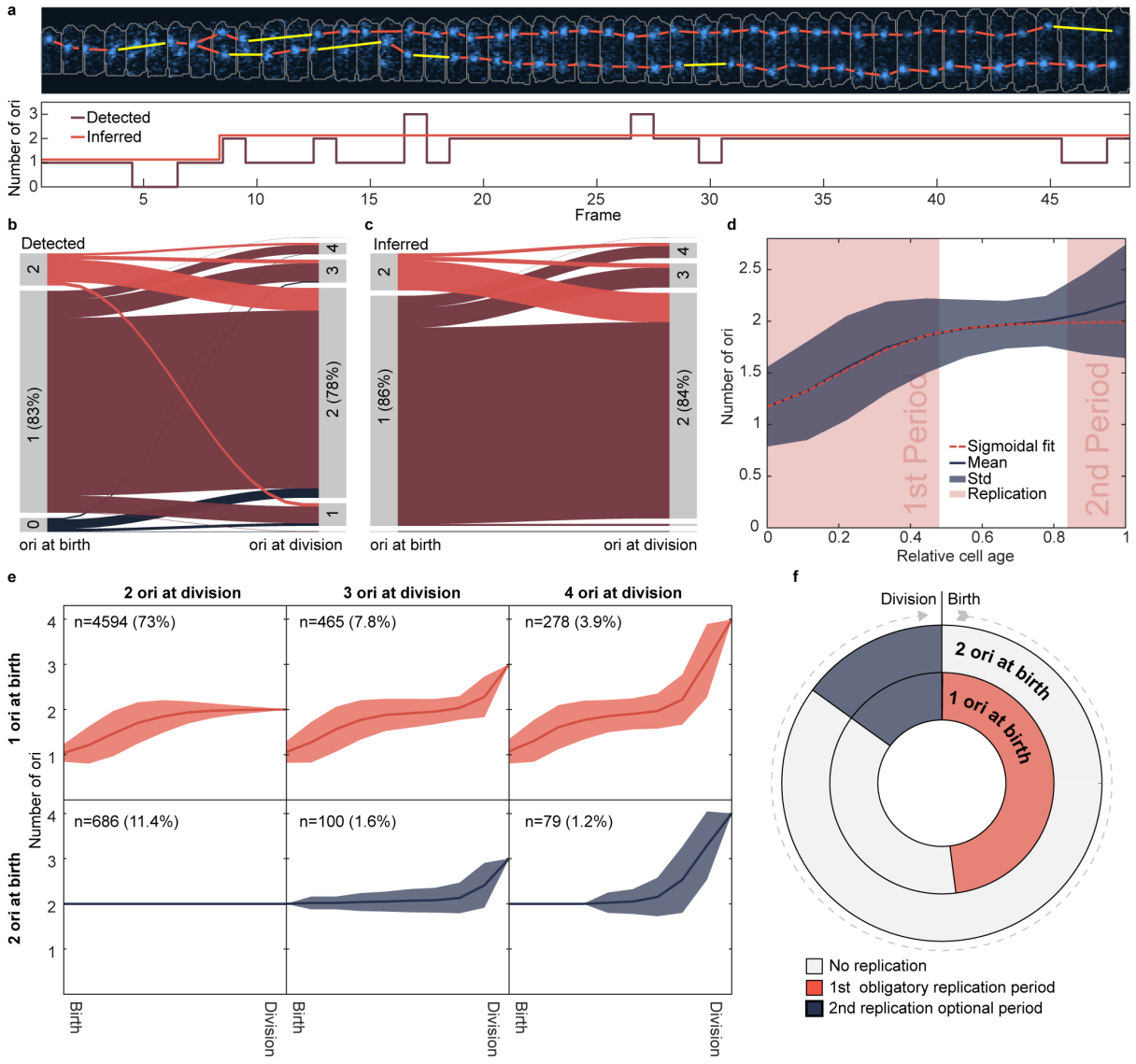

Fig. 3: Replication of *ori* occurs during two periods in the cell cycle. **a**, An example cell cycle of the *E. coli* strain IS130. It encodes mTourquise2-ParB which labels the origin of replication (*ori*). Top: Fluorescent signal with tracked *ori*. Bottom: Number of *ori*, both detected and inferred. Frame interval is 5 min. **b**, Alluvial plot showing how the number of *ori* changes from birth to division of the detected foci. **c**, Same as **b** but after ⋆Track was used. **d**, Mean number of *ori* ± standard deviation plotted against relative cell age (0 is birth and 1 is division). The two areas shaded in light red indicate periods during the cell cycle in which *ori* is replicated. The red dashed line is a sigmoidal fit from 0 to 0.85 (excluding the second period). See also Fig. S8. **e**, Data from **d** grouped by number of *ori* at birth (rows) and number of *ori* at division (columns). **f**, Cartoon of replication periods. The plots in **b-e** were created from 6265 cell cycles.

⋆Track is a powerful new tool for the tracking of plasmid and chromosomal DNA loci. It provides predictions for the timing of replication (initial foci splitting) events that should provide mechanistic insight across a range of systems, and not only for bacteria. ⋆Track can also be applied to any persistent particles that increase in copy number during the cell cycle and could therefore be used to study protein complexes that display this behavior[37–40]. The algorithm is deterministic (there are no stochastic or deep learning components) and therefore produces reproducible results with the user

having to specify only four estimatable parameters. ⋆Track is available at https://gitlab.gwdg.de/murray-group/StarTrack.

4.1.5.1. Limitations

⋆Track is designed for the tracking of small numbers of persistent particles. However the first part of the algorithm (the A* step) does not require persistence and can be used in isolation to identify trajectory segments of non-persistence particles. This can be achieved through the setting of a flag in the code. However, with higher particle numbers the A* step will become increasingly computationally expensive (take too long to run). This will depend on the provided parameters through the threshold cost for acceptable links c_{th} (see methods). Nevertheless, the method should also be useful for applications with intermediate particle numbers in which accuracy is paramount.

4.1.6. Author Contributions

R.K. and S.M.M. designed the research; R.K. developed, tested and applied the method; I.S. performed origin tracking experiments; R.K and S.M.M wrote the paper.

4.1.7. Declaration of Interests

The authors declare no competing interests.

4.1.8. Acknowledgements

We thank Frédéric Boccard for the strain RM3 and the plasmid pFHC2973, Jaan Mannik for strain RH3 and Jean-Yves Bouet for strain DLT1215 and the plasmid pJYB234.

4.1.9. References

1. Manzo, C. & Garcia-Parajo, M. F. A review of progress in single particle tracking: from methods to biophysical insights. *Rep. Prog. Phys.* **78**, 124601 (2015).
2. Shen, H. *et al.* Single Particle Tracking: From Theory to Biophysical Applications. *Chem. Rev.* **117**, 7331–7376 (2017).
3. Meijering, E., Dzyubachyk, O. & Smal, I. Chapter nine - Methods for Cell and Particle Tracking. in *Methods in Enzymology* (ed. conn, P. M.) vol. 504 183–200 (Academic Press, 2012).
4. Chenouard, N. *et al.* Objective comparison of particle tracking methods. *Nat. Methods* **11**, 281–289 (2014).
5. Tinevez, J.-Y. *et al.* TrackMate: An open and extensible platform for single-particle tracking. *Methods* **115**, 80–90 (2017).
6. Jaqaman, K. *et al.* Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods* **5**, 695–702 (2008).
7. Sage, D., Neumann, F. R., Hediger, F., Gasser, S. M. & Unser, M. Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics. *IEEE Trans. Image Process.* **14**, 1372–1383 (2005).
8. Chenouard, N., Bloch, I. & Olivo-Marin, J. Multiple Hypothesis Tracking for Cluttered Biological Image Sequences. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 2736–3750 (2013).

9. Oviedo-Bocanegra, L. M., Hinrichs, R., Rotter, D. A. O., Dersch, S. & Graumann, P. L. Single molecule/particle tracking analysis program SMTracker 2.0 reveals different dynamics of proteins within the RNA degradosome complex in *Bacillus subtilis*. *Nucleic Acids Res.* **49**, e112–e112 (2021).

10. Sbalzarini, I. F. & Koumoutsakos, P. Feature point tracking and trajectory analysis for video imaging in cell biology. *J. Struct. Biol.* **151**, 182–195 (2005).

11. Tal, S. & Paulsson, J. Evaluating quantitative methods for measuring plasmid copy numbers in single cells. *Plasmid* **67**, 167–173 (2012).

12. Sanchez, A. *et al.* Stochastic Self-Assembly of ParB Proteins Builds the Bacterial DNA Segregation Apparatus. *Cell Syst.* **1**, 163–173 (2015).

13. Bouet, J. Y., Rech, J., Egloff, S., Biek, D. P. & Lane, D. Probing plasmid partition with centromere-based incompatibility. *Mol. Microbiol.* **55**, 511–525 (2005).

14. Männik, J., Castillo, D. E., Yang, D., Siopsis, G. & Männik, J. The role of MatP, ZapA and ZapB in chromosomal organization and dynamics in *Escherichia coli*. *Nucleic Acids Res.* **44**, 1216–1226 (2016).

15. Mercier, R. *et al.* The MatP/matS Site-Specific System Organizes the Terminus Region of the E. coli Chromosome into a Macrodomain. *Cell* **135**, 475–485 (2008).

16. Li, Y., Youngren, B., Sergueev, K. & Austin, S. Segregation of the *Escherichia coli* chromosome terminus. *Mol. Microbiol.* **50**, 825–834 (2003).

17. Nielsen, H. J., Ottesen, J. R., Youngren, B., Austin, S. J. & Hansen, F. G. The *Escherichia coli* chromosome is organized with the left and right chromosome arms in separate cell halves: *E. coli* chromosome segregation. *Mol. Microbiol.* **62**, 331–338 (2006).

18. Wang, P. *et al.* Robust Growth of Escherichia coli. *Curr. Biol.* **20**, 1099–1103 (2010).

19. Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I. & Elf, J. Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proc. Natl. Acad. Sci.* **114**, 9170–9175 (2017).

20. Köhler, R., Kaganovitch, E. & Murray, S. M. High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS. *eLife* **11**, (2022).

21. Stylianidou, S., Brennan, C., Nissen, S. B., Kuwada, N. J. & Wiggins, P. A. SuperSegger: robust image segmentation, analysis and lineage tracking of bacterial cells: Robust segmentation and analysis of bacteria. *Mol. Microbiol.* **102**, 690–700 (2016).

22. Sliusarenko, O., Heinritz, J., Emonet, T. & Jacobs-Wagner, C. High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol. Microbiol.* **80**, 612–627 (2011).

23. Hart, P. E., Nilsson, N. J. & Raphael, B. A Formal Basis for the Heuristic Determination. *IEEE Trans. Syst. Sci. Cybern.* 8 (1968).

24. Ershov, D. *et al.* TrackMate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat. Methods* **19**, 829–832 (2022).

25. Pinto, U. M., Pappas, K. M. & Winans, S. C. The ABCs of plasmid replication and segregation. *Nat. Rev. Microbiol.* **10**, 755–765 (2012).

26. Chattoraj, D. K. Control of plasmid DNA replication by iterons: no longer paradoxical. *Mol. Microbiol.* **37**, 467–476 (2000).

27. Del Solar, G. & Espinosa, M. Plasmid copy number control: an ever-growing story. *Mol. Microbiol.* **37**, 492–500 (2000).

28. Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Q. Rev. Biophys.* **34**, 1–59 (2001).

29. Igor Konieczny, Katarzyna Bury, Aleksandra Wawrzycka, & Katarzyna Wegrzyn. Iteron Plasmids. *Microbiol. Spectr.* **2**, (2014).

30. Onogi, T., Miki, T. & Hiraga, S. Behavior of Sister Copies of Mini-F Plasmid after Synchronized Plasmid Replication in *Escherichia coli* Cells. *J. Bacteriol.* **184**, 3142–3145 (2002).

31. Walter, J. C. *et al.* Physical Modeling of a Sliding Clamp Mechanism for the Spreading of ParB at Short Genomic Distance from Bacterial Centromere Sites. *iScience* **23**, (2020).

32. Jun, S., Si, F., Pugatch, R. & Scott, M. Fundamental principles in bacterial physiology—history, recent progress, and the future with focus on cell size control: a review. *Rep. Prog. Phys.* **81**, 056601 (2018).

33. Wallden, M., Fange, D., Lundius, E. G., Baltekin, Ö. & Elf, J. The Synchronization of Replication and Division Cycles in Individual E. coli Cells. *Cell* **166**, 729–739 (2016).

34. Si, F. *et al.* Mechanistic Origin of Cell-Size Control and Homeostasis in Bacteria. *Curr. Biol.* **29**, 1760-1770.e7 (2019).

35. Reyes-Lamothe, R., Possoz, C., Danilova, O. & Sherratt, D. J. Independent Positioning and Action of Escherichia coli Replisomes in Live Cells. *Cell* **133**, 90–102 (2008).

36. Si, F. *et al.* Invariance of Initiation Mass and Predictability of Cell Size in Escherichia coli. *Curr. Biol.* **27**, 1278–1287 (2017).

37. Murray, S. M. & Sourjik, V. Self-organization and positioning of bacterial protein clusters. *Nat. Phys.* **13**, 1006–1013 (2017).

38. Hofmann, A., Mäkelä, J., Sherratt, D. J., Heermann, D. & Murray, S. M. Self-organised segregation of bacterial chromosomal origins. *eLife* **8**, e46564 (2019).

39. Richards, D. M., Hempel, A. M., Flärdh, K., Buttner, M. J. & Howard, M. Mechanistic Basis of Branch-Site Selection in Filamentous Bacteria. *PLOS Comput. Biol.* **8**, e1002423 (2012).

40. Roberts, M. A. J., Wadhams, G. H., Hadfield, K. A., Tickner, S. & Armitage, J. P. ParA-like protein uses nonspecific chromosomal DNA binding to partition protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 6698–703 (2012).

4.1.10. Supplementary Figures

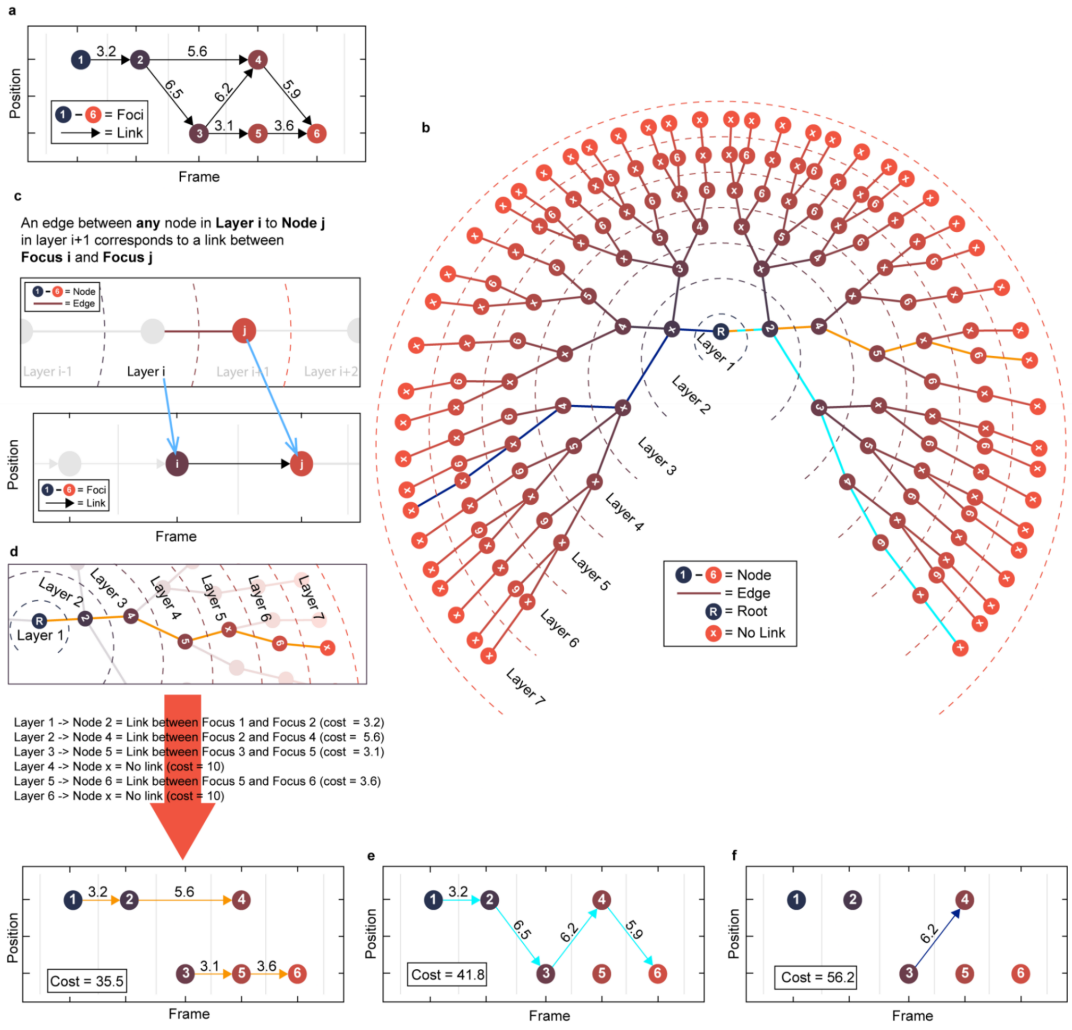

Fig. S1: Layered Graph. **a**, Series of 6 foci over 5 frames. ★Track first generates a set of acceptable links from every focus to any other focus on a subsequent frame. Acceptable links are those with a cost below a specified threshold, in this case 10 (indicated by arrows), which is obtained from the cost function and the maximum number of skipped frames. **b**, The layered graph generated from **a**. Each direct path from layer 1 to layer 7 corresponds to a unique tracking of the foci in **a**, such that each focus has at most 1 incoming and outgoing link. The graph contains all such trackings. An edge from any node in **layer** *i* to a numbered **node** *j* in the next layer corresponds to a link between focus *i* and focus *j*. An edge from any node in **layer** *i* to a **node** labelled 'x' in the next layer corresponds to focus *i* having no outgoing link. Note that it is the layer, not the node, from which the edge originates that determines the meaning of the edge. Each edge has a weight (not shown here) which is equal to the cost of linking the two foci (for edges into numbered nodes) or the cost of a focus not having an outgoing link (for edges into the 'x' node). In the latter case, the cost is equal to the threshold mentioned in **a** or if lower the cost of going to the end (see materials). The A*-algorithm generates this graph dynamically as it searches for the optimal tracking. **c**, Cartoon of how to translate an edge in the layered graph (top panel) into a link in the tracking (bottom panel). **d**, Example of how to interpret a path in the graph from **b** and how to translate it into trajectory segments. Top: Example path, highlighted in yellow from **b**. Bottom: Resulting trajectory segments with the cumulative cost of all edges of the path. The threshold cost (10) is associated to both foci 4 and 6 for not having an outgoing link. **e** and **f**, trajectories corresponding to two other paths from **b** as in **d.** The color of the links corresponds to the path in **b** with the same color.

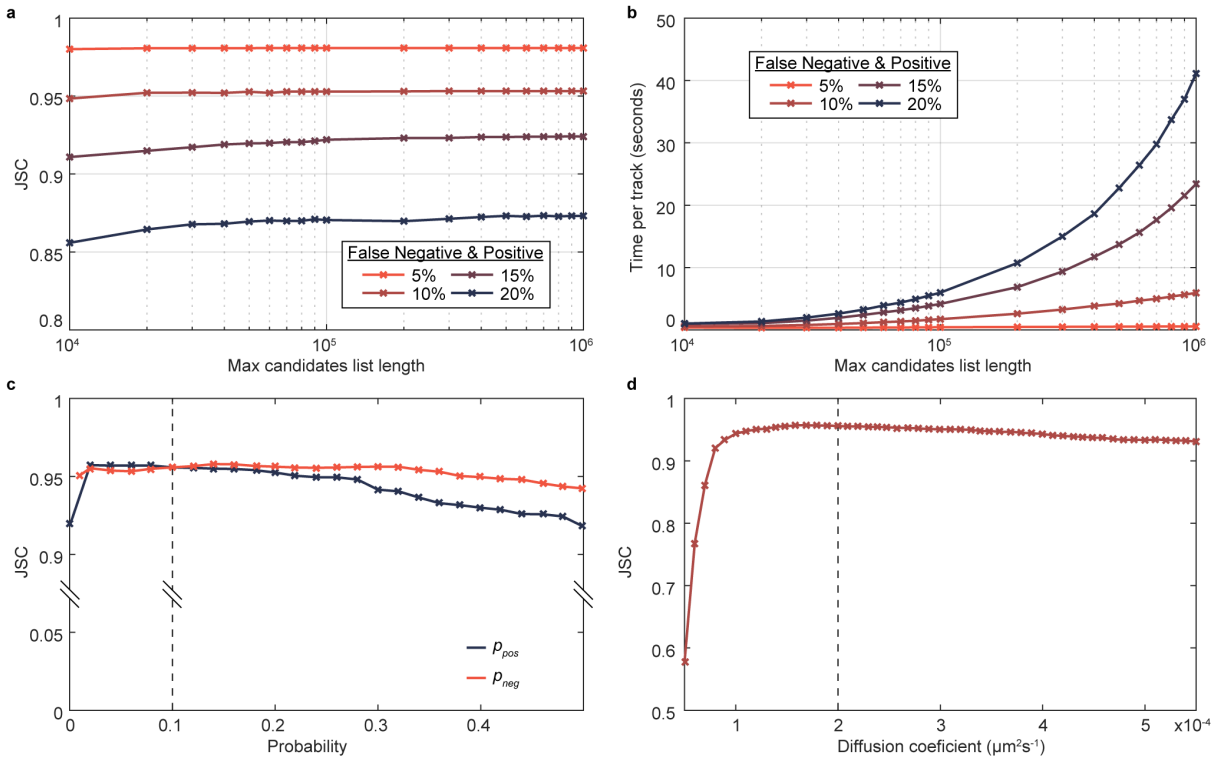

Fig. S2: Robustness and sensitivity of ★Track. **a**, The JSC (as in Fig. 1) of simulated trajectories with 5-20% false positives/negatives added/removed plotted against maximum length of the candidate list of the A* implementation. **b**, Same as **a** but for the runtime of the algorithm. Maintaining the candidate list beyond 10^6 entries does not improve accuracy while increasing runtimes. **c**, Robustness of chosen parameters. JSC of simulated trajectories with 10% false positives/negatives as a function of the parameters p_{pos} and p_{neg}. The horizontal dashed line represents the ground truth probability of the simulated data. Note that the first data point of p_{neg} is 0.01 and not 0. This is because the algorithm is unable to track data that includes false negatives when the user specifies that there are none ($p_{neg} = 0$). **d**, Same as **c** but for the diffusion coefficient.

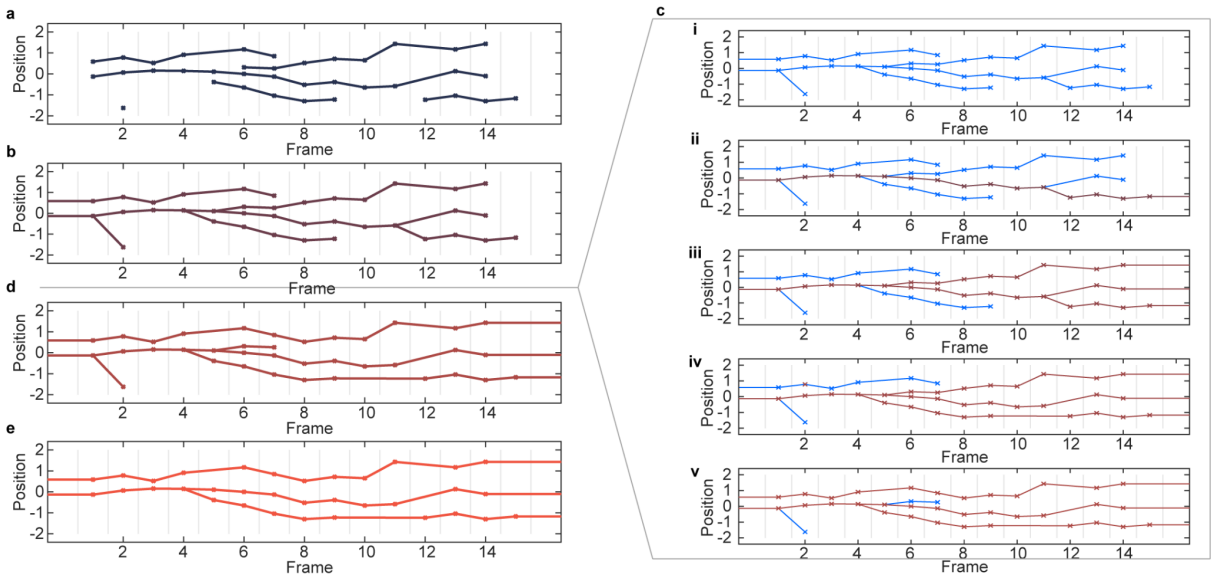

Fig. S3: Stitching A* output together. **a**, Example output of linked foci produced by A* tracking. **b**, The first focus of every trajectory segment is linked to an earlier focus (chosen by lowest cost). Foci on the first frame are linked to imaginary foci before the first frame. **c**, Detailed overview of the procedure between **b** and **d**. **i**, The goal is to have as many foci as possible connected to an imaginary frame before the beginning (frame 0) and to an imaginary frame after the end (in this case frame 16) through its links to other foci. Initially no focus is connected in this way and we colour the entire tracking blue. **ii**, For each loose end of the tracking, starting at the last loose end (here on frame 15), we attempt to integrate it into the tracking by linking it to the imaginary last frame or replacing an existing link such that the overall tracking improves (see methods for details of how the cost of the entire tracking is calculated). Here the last loose end is linked to the imaginary last frame of the tracking. This results in 16 foci being connected as required (coloured red). **iii**, Same as in **ii** for the two loose ends on frame 14. 10 additional foci are now connected as required. **iv**, The next earliest loose end (frame 9) is interwoven by replacing the link from frame 11 to 12 by a link from frame 9 to 12. This results in 5 additional foci being connected. **v**, The next loose end (frame 7) is interwoven into the tracking by replacing an existing link. This connects 6 more foci, however since the replaced link was not part of a splitting event this also disconnects 2 foci from the imaginary last frame. The remaining two loose ends cannot be interwoven into the tracking in such a way that the tracking is improved. **d**, The result of the interweaving procedure. **e**, Remaining unconnected foci are labelled as false positives and removed.

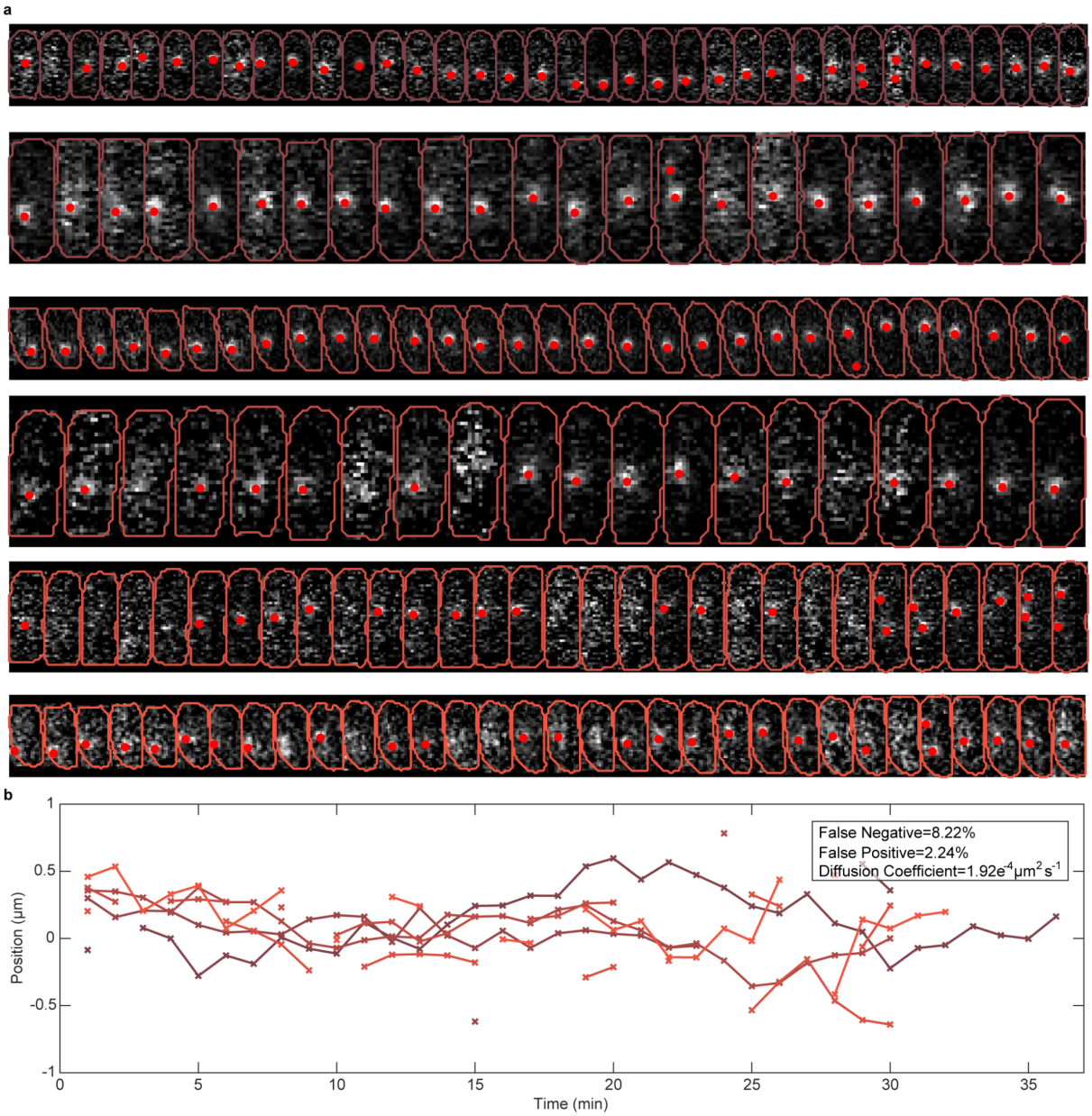

Fig. S4: Estimating parameters from data with a low density of foci. **a**, Example cell cycle fragments of the *E. coli* strain DLT3145 containing, except on some isolated frames, a single focus of ParB-mVenus (labeling F plasmid) . Cell contours are on top of the fluorescent signal and detected foci are highlighted by red dots. Frame rate is 1 min^{-1}. **b**, Foci trajectories from the above cycle fragments. The trajectories contain 9 false positives and 21 false negatives. This is easy to determine by looking at the number of foci on each frame. We found 604 analysable cycle fragments. They contained 361 false positives and 1323 false negatives in a total of 16103 foci. The resulting estimates are a false negative occurrence of 8.22% and a false positive occurrence of 2.24%. Further we estimate the diffusion coefficient of 1.92 x 10^{-4} $\mu m^2 s^{-1}$ by the variance of the displacement along the long axis between adjacent frames containing only one focus (the step-size distribution) divided by twice the time step.

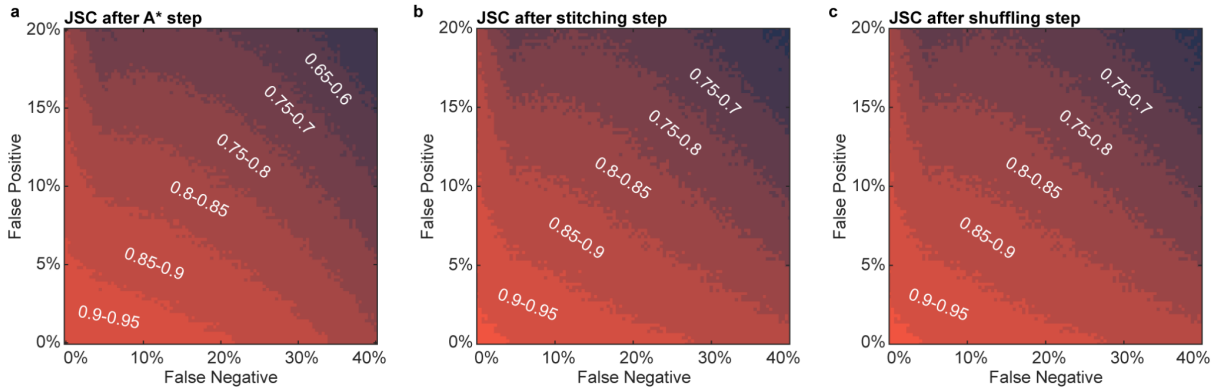

Fig. S5: A breakdown of the impact of each step of the algorithm. Shown is the JSC of links, which differs from the JSC or localisations presented in Figure 1. As the shuffling step only modifies existing links between foci, without altering the foci included in the tracking, the JSC defined in Figure 1 would remain unaltered. For this reason, we switched to the JSC of links to better capture the effect of the shuffling step on the accuracy of the tracking. **a**, JSC after the A* step. **b**, JSC after the stitching step. **c**, JSC after the shuffling step. The stitching step results in an improvement in the link accuracy especially at high false negative/positive rates. The effect is small under this measure since typically only a small fraction of links are missing or incorrect. Note that despite our switch to the JSC between links, there is minimal disparity between **b** and **c**. This can be attributed to the infrequent need for shuffling to enhance tracking accuracy.

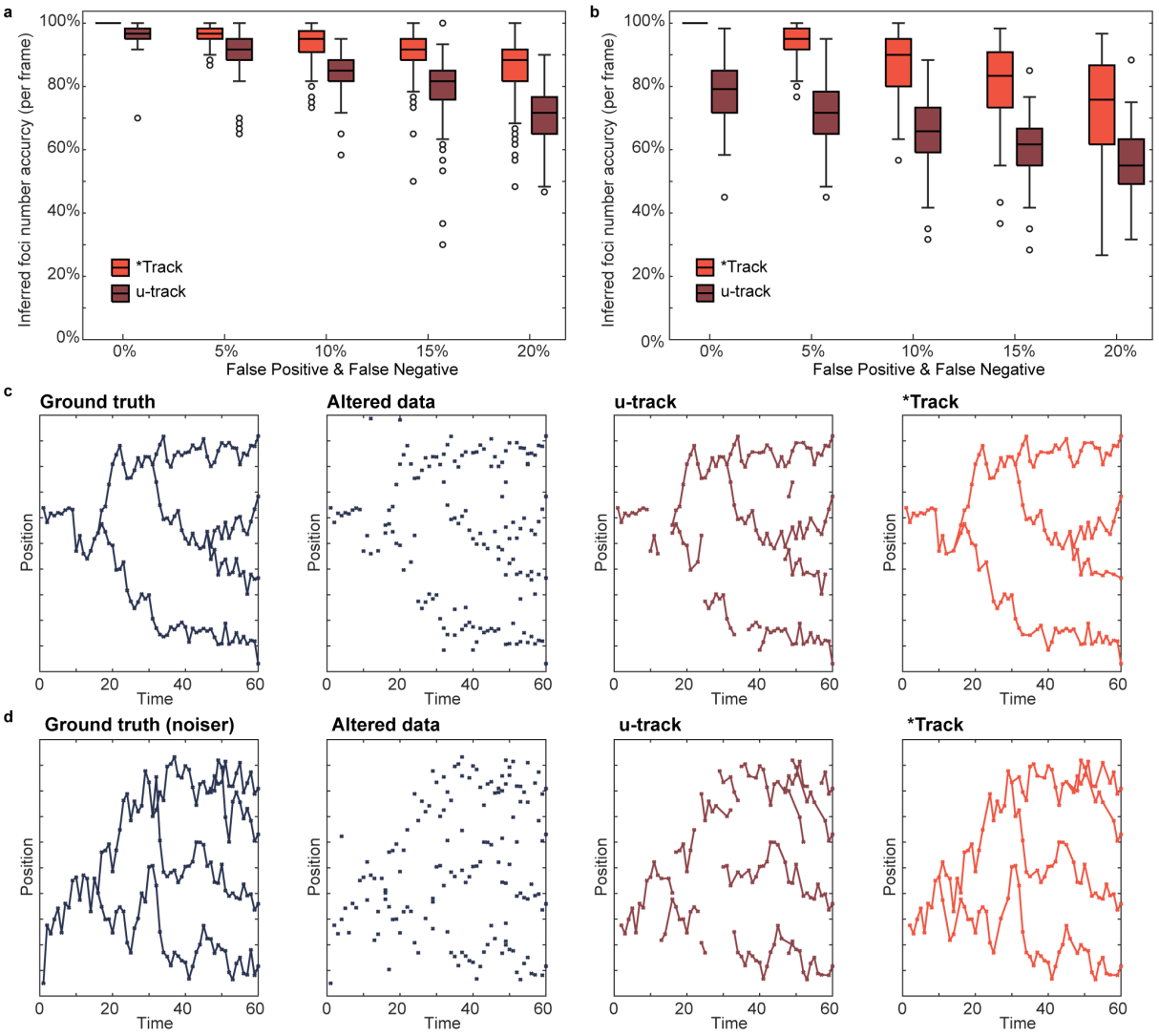

Fig. S6: A comparison of ★Track and u-track. Due to the differences in the two methods (u-track can produce disconnected segments of the trajectory whereas ★Track cannot), we use the percentage of frames in which the number of inferred foci is equal to the ground truth as a measure of comparison of the methods. The parameters for u-track were chosen to give the best results and are found in Table S1 and in the code repository. **a**, Box charts comparing the accuracy of ★Track to u-track. The ground truth was altered similarly to Figure 1c before being processed by the tracking algorithms. Each box chart was produced from 100 simulations. Each box chart displays the following information: the median, the lower and upper quartiles, any outliers (defined as values more than 1.5 times the interquartile range away from the quartiles), and the minimum and maximum values that are not outliers. **b**, as in **a** but for more mobile foci as in Figure 1d. **c** Example from **a** with 10% false positives & false negatives. **d**, Example from **b** with 10% false positives & false negatives.

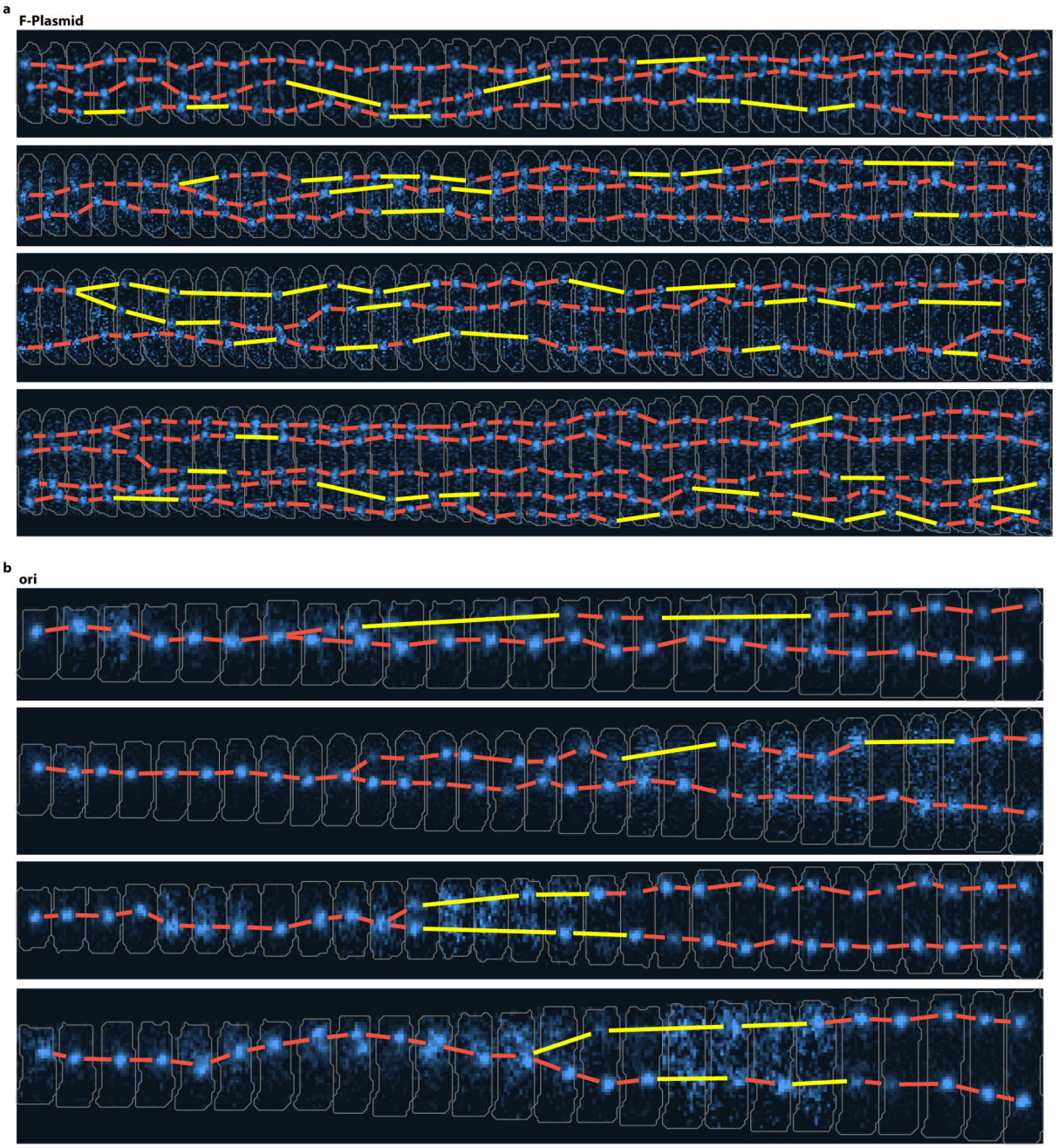

Fig. S7: Examples of ⋆Tracked foci for ParB-mVenus labelled F-Plasmid and mTurqouise2-ParB labelled ori. **a**, cell cycle fragments (40 frames) from *E. coli* strain DLT3145 with tracked ParB-mVenus foci. Frame rate is 1 min^{-1}. **b**, Cell cycles from *E. coli* strain IS130 with tracked mTurqouise2-ParB foci. Frame interval is 5 min. Yellow lines indicate links with frame skips.

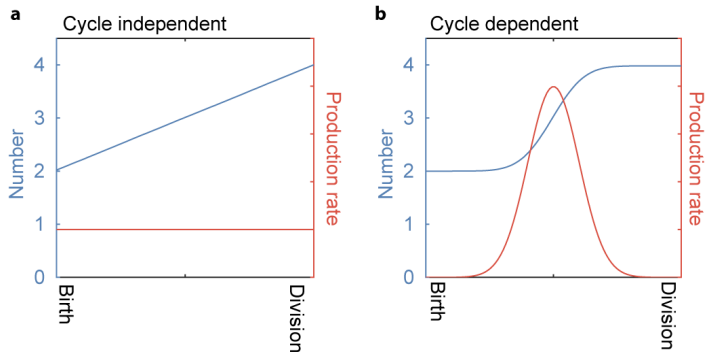

Fig. S8: Cycle independent and cycle dependent replication. **a**, Cycle independent replication. At each point in time there is the same probability of replication (orange line). This causes the number (blue line) to increase linear **b**, Cycle dependent replication. There is a period during the middle of the cycle where replication occurs (orange line). This creates a sigmoidal change in the number.

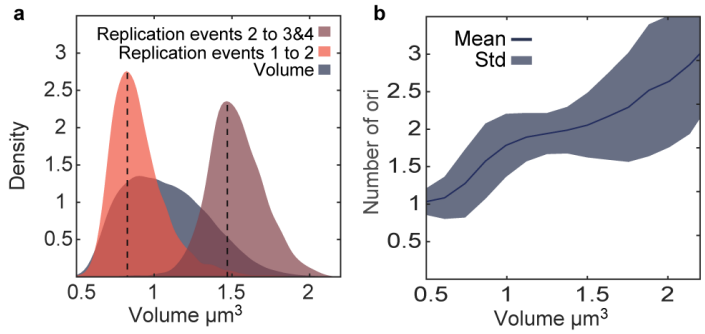

Fig. S9: Second replication during the late cell cycle supports that volume per ori triggers replication. **a**, The distribution of replication events from 1 to 2 ori (orange) and from 2 to 3 or 4 ori (brown) plotted against cell volume. Black dashed lines indicate the peaks of the distribution at roughly 0.8 µm^3 and 1.5 µm^3. For reference the distribution of all measured volumes is plotted in gray. **b**, Mean number of ori ± standard deviation plotted against volume.

Table S1 Parameters used with u-track

Parameter	Value
timeWindow	6
minTrackLen	2
linearMotion	0
minSearchRadius	6
maxSearchRadius	6
brownStdMult	3
useLocalDensity	1
brownScaling	[0.05, 0.01]
lenForClassify	5
linStdMult	3
linScaling	[1, 0.01]
maxAngleVV	30
gapPenalty	5
gapExcludeMS	1
strategyBD	0.001

4.1.11. Supplementary Videos

Supplementary Video 1: Video of the generation of the layered graph displayed in Fig. S1b. The links corresponding to the current path are shown in the bottom-middle.

Supplementary Video 2: Video of the A*-like path finding process. Only the colored part of the graph is generated. The links corresponding to the current best candidate path are shown in the bottom-middle. Additionally the partial cost of the edges of the candidate path, the heuristic estimate of completing this path and the total cost estimate are shown.

5. Paper III

Despite extensive research on ParA's role in plasmid partitioning, its precise function remains only partly understood. A critical aspect of this process involves interactions between ParA and other components, such as ParB and the nucleoid, which are crucial for efficient partitioning. Numerous studies have delved into this subject, contributing to a deeper understanding. However, a key element of ParA seems to have been overlooked: the intracellular localization patterns of ParA and their impact on its overall function still require clarification.

Current computational models struggle to generate the ParA localization pattern when plasmids are stable positioned within the cell. These models predict that ParA is distributed fairly uniformly throughout the nucleoid. This discrepancy between simulations and experimental observations calls for further investigation into the mechanisms governing ParA dynamics. In this manuscript, we propose a molecular mechanism, termed ParA-ParA recruitment, where nucleoid-bound ParA dimers recruit cytosolic ParA dimers to the nucleoid in their immediate vicinity. Utilizing high-throughput single-cell data, we demonstrate that incorporating ParA-ParA recruitment into our simulations adequately reproduces ParA dynamics observed *in vivo*.

My contribution to this paper includes all aspects of this paper.

This paper has not yet been submitted, it is in the final stages of preparation and is expected to be submitted soon.

5.1. Putting the Par back into ParABS: A Novel Mode of Plasmid Partitioning Driven by ParA Oscillations

Robin Köhler, Seán M. Murray*

Max Planck Institute for Terrestrial Microbiology and LOEWE Centre for Synthetic Microbiology (SYNMIKRO), Karl-von-Frisch Straße 14, 35043 Marburg, Germany

*Email: sean.murray@synmikro.mpi-marburg.mpg.de

5.1.1. Abstract

The ParABS system is crucial for the faithful segregation and inheritance of many bacterial chromosomes and low-copy number plasmids. Despite extensive research, the spatio-temporal dynamics of ParA and its interplay with the other components of the system remain elusive. In this study, on the F-plasmid, we utilize high-throughput imaging, quantitative data analysis, and computational modeling to explore the *in vivo* dynamics of ParA, and ParB, and their interactions with the nucleoid. In accordance with previous literature, our study shows that ParA undergoes collective migration between cell halves multiple times per cell cycle, resulting in oscillations over time. Our findings reveal that a constricted nucleoid plays an essential role in these migrations; following its formation, migration events are typically set in motion by a plasmid crossing the mid-cell. Quantitatively, we show that ParB follows ParA and displaces it due to ParB-induced hydrolysis. We postulate that this hydrolysis, along with the relocalization of a significant amount of ParB such as a ParB-coated plasmid crossing mid-cell, triggers the migration of ParA. By incorporating ParA-ParA recruitment (nucleoid-bound ParA recruits cytosolic ParA) and a model of a nucleoid into our simulations, we successfully reproduce the ParA oscillations observed *in vivo*. We further show that without ParA-ParA recruitment, plasmids are not equally partitioned between the two lobes of a constricted nucleoid, which can lead to plasmid loss. ParA-ParA recruitment effectively resolves this issue. Overall, this study provides a comprehensive mechanistic understanding of ParA dynamics, shedding light on the underlying principles governing its oscillatory behaviour.

5.1.2. Introduction

In bacterial cell division, the precise segregation and positioning of genetic material is essential to produce viable offspring. This process is crucial for both chromosomal inheritance and the spatial distribution of low copy number plasmids like the F-plasmid. With an average of 3 F-plasmids per cell

(Köhler et al., 2022), random positioning cannot ensure stability and an active partitioning system is necessary. The most prevalent partitioning system is the ParABS system (Gerdes et al., 2000), also found on the F-plasmid, composed of three components: (i) *parS*, a centromere-like binding site; (ii) ParA, a Walker-type ATPase; (iii) ParB, a protein that binds to *parS*. Both ParA and ParB dimerizes and the latter binds to *parS* and slides along DNA for several kilobases, enabling the buildup of a substantial amount of ParB on the plasmid, forming a condensed region called the partitioning complex (Osorio-Valeriano et al., 2019; Taylor et al., 2021). Furthermore, ParB-coated plasmids move towards the direction of the greatest gradient of ParA and stimulate the ATPase activity of ParA. ParA, in its dimeric ATP-bound state, binds DNA non-specifically, coating the nucleoid of the cell (Ebersbach & Gerdes, 2001; G. E. Lim et al., 2005). Upon ATP hydrolysis, ParA detaches from the DNA and requires a prolonged cytosolic conformational transition before re-binding to DNA (Vecchiarelli et al., 2010). This minimalistic system distributes plasmids equally along the cell's long axis, ensuring equal inheritance to daughter cells during division.

Despite significant advancements in understanding the ParABS system, some aspects remain elusive. Among them, the dynamics of ParA stand out as a particularly intriguing subject for investigation, which this work aims to address.

Three noteworthy features of ParA include (1) its oscillatory behaviour within the cell, with the majority of ParA alternatingly localized between cell halves (Ah-Seng et al., 2013; Ebersbach & Gerdes, 2001, 2004), (2) its ability to diffuse on the nucleoid, and (3) its ability to form filaments *in vitro* as observed through electron microscopy (Ebersbach et al., 2006; G. E. Lim et al., 2005). This demonstrates that ParA can form structures larger than dimers under the right conditions and likely possesses multiple domains for interacting with other ParA. These observations initially led to a proposed model for plasmid partitioning and positioning, in which ParA forms helical filaments within the cell. The constant polymerization and depolymerization of these filaments act as a motor to move the plasmids. However, the formation of *in vivo* filaments within the cell has never been conclusively proven, causing this model to fall out of favor for diffusion ratchet models.

The next generation of models, diffusion ratchet models, propose that ParA uniformly binds to the nucleoid, with nucleoid-bound ParA and plasmid-bound ParB forming tethers (H. C. Lim et al., 2014; Surovtsev et al., 2016). This arrangement allows plasmids to harness chromosomal fluctuations and move towards the highest incoming ParA flux, a hypothesis supported by various models (Ietswaart et al., 2014; Köhler et al., 2022; Sugawara & Kaneko, 2011). ParA diffusion on the nucleoid results in positional flux differences based on the plasmid's position. For example, if the plasmid is positioned near the nucleoid's edge, a smaller quantity of nucleoid-bound ParA will diffuse towards the plasmid from that direction compared to the other directions. This allows each plasmid to "sense" the nucleoid's edges and find its target position. Furthermore, if the average ParA diffusion distance is

greater than or equal to half the length of the nucleoid, flux balance (the flux is the same from opposite sides) is only achieved when all plasmids are equally positioned along the nucleoid's long axis. However, one caveat with these models is their inability to generate the oscillatory pattern of ParA in the presence of multiple plasmids. ParA in these models is uniformly distributed around the plasmids with peaks or sinks of ParA at the plasmid based on the model's parameters. Despite the inability to produce oscillations, the models prove to be sufficient to ensure faithful segregation of plasmids.

This highlights that there is a discrepancy between our understanding of the system and the findings obtained through experimental research. Examining closely related proteins of ParA may provide insights into the how and why of ParA oscillations.

MinD of the MinCDE system is a prime candidate. MinD, another Walker-type ATPase, also dimerizes (Ramm et al., 2019) and works in conjunction with MinC and MinE to localize the FtsZ ring. FtsZ plays a crucial role in establishing the division site of the cell, through the assembly of a functional septum. If the MinCDE system is absent or impaired, cells will produce anucleate and mini cells (de Boer et al., 1989). The accumulation of FtsZ is inhibited by MinC, which passively follows the movements of MinD and MinE (Huang et al., 2003). The cell membrane serves as a binding site for MinD, and its ATPase activity is stimulated by MinE, leading to the dissociation of MinD from the membrane into the cytosol. Upon entering the cytosol, MinD exchanges ADP for ATP and rebinds to the membrane region with the lowest concentration of MinE, resulting in periodic spatial relocalization of both proteins between the cell poles. This leads, as a time average, to an intracellular gradient of MinE, MinD, and consequently MinC, which peaks at the poles and reaches its lowest point at mid-cell. The FtsZ ring assembles at the mid-cell, marking the future division site, due to the minimal presence of MinC, which hinders its accumulation. An important insight gained from modelling the system is the need for MinD-MinD recruitment, in which membrane-bound MinD helps cytosolic MinD to bind in close proximity (Halatek & Frey, 2012; Huang et al., 2003). Without this recruitment, no oscillations can be sustained. The similarities between MinD and ParA suggest that ParA may have a recruitment function, akin to what is assumed for MinD, in order to create the oscillations.

In this paper, we investigate the oscillatory nature of ParA by using a combination of high-throughput imaging and quantitative data analysis. We analyzed thousands of cell cycles to characterize ParA dynamics in detail, which led us to the discovery that nucleoid constriction formation triggers stable and prolonged ParA oscillations. We propose a mechanism, ParA-ParA recruitment, that can explain ParA oscillations while being consistent with our *in vivo* observations. By adding ParA-ParA recruitment and an accurate representation of the nucleoid to our previously developed model, we demonstrate that these oscillations are essential for preventing plasmid loss in cells with an early

nucleoid constriction. Ultimately, our findings provide a comprehensive understanding of ParA oscillations and their functional significance within the ParABS system.

5.1.3. Results

5.1.3.1. The Impact of Plasmid Movement on ParA Oscillatory Behaviour

To study the oscillations of ParA within the cell, we employed a high-throughput microfluidic approach using a custom-fabricated "mother machine" device in conjunction with a segmentation, tracking, and foci detection pipeline (as shown in Figure S1). To capture all components of the ParABS system and their relationship with the nucleoid, we utilized a triple-labeled *E. coli* strain (ParA-mVenus, ParB-mTurquoise2, and HU-mCherry) (Guilhas et al., 2020).

First, we evaluated the distribution of ParA within living cells (Figure 1A depicts a representative cell cycle). At the onset of the cell cycles, ParA is dispersed evenly. Later (at ~30min), ParA collectively migrates to one half of the cell (red shaded region). It stays in this asymmetric state (as defined in Figure S2A) for 20 minutes until another mass migration of ParA (at ~50min) takes place. The ParA remains in this asymmetric state until cell division. This exemplary cell cycle highlights a key characteristic of ParA oscillations: They do not follow a smooth wave function. Rather, they exhibit a square function-like pattern (Figure 1A & Figure S2A), in which all ParA migrate, in a short migration period, into one cell half and remain there stably until the next migration.

Figure 1BC shows the absolute ParA asymmetry as a function of cell length. The data reveal that the asymmetry in the ParA signal increased with cell length and saturated at approximately 6 µm, indicating the presence of a size limiting factor beyond which the asymmetry does not increase further. Further, the average length between migration events was on average 21.5 minutes (Figure 1D), with a standard deviation of 17.1 minutes. Figure S2D shows the average time between migration events plotted against time since birth, which exhibited a saturation point at around 100 minutes. Similarly, Figure S2E depicts a similar pattern when plotting this time against cell length, with an increase in asymmetry up to 6 µm followed by a plateau. These observations highlight that both asymmetry and the duration between migration events increased as the cell cycle progresses.

On average, each cell underwent two migration events between birth and division, resulting in the cell being in an asymmetric state for an average of 38.8% of the time (Figure S2C). The ParA signal in newborn cells was highly heterogeneous due to the fact that the asymmetry increases as the cell cycle progresses, resulting in the majority of cells being in an asymmetric state at division (Figure 1C). On average, upon division, one daughter cell inherited 73% of the ParA signal while the other only inherited 27% (Figure S2F).

Examination of the trajectories of ParB-coated plasmids revealed a link between plasmid movement and ParA migration (Figure 1E). In the majority of cases, the migration of ParA was initiated once a plasmid crossed mid-cell from the ParA-depleted side to the ParA-rich side. Furthermore, the migration of ParA followed a particular scheme. Simultaneously, the concentration of ParA decreases uniformly in one half and increases uniformly in the other (Figure S2G). There was no clear evidence of a retracting gradient as described by previous literature. By aligning the asymmetry of ParA relative to the moment a plasmid crossed mid-cell, the significance of plasmid movement becomes evident (Figure 1F). The asymmetry of ParA was the highest prior to the crossing event and six minutes later the sign of the asymmetry changed, indicating that ParA migrated away from the cell half that the plasmid had just crossed into. We observed the same behaviour on a single cellular level as 65% of ParA migration events are preceded by plasmid crossing events (Figure S3).

In summary, we confirmed that ParA exhibits a highly asymmetric distribution within the cell and is inherited unevenly between daughter cells. The asymmetry and the time between migration events both increase with cell length and time since birth. Moreover, we discovered that the majority of migration events are initiated by plasmids crossing the mid-cell.

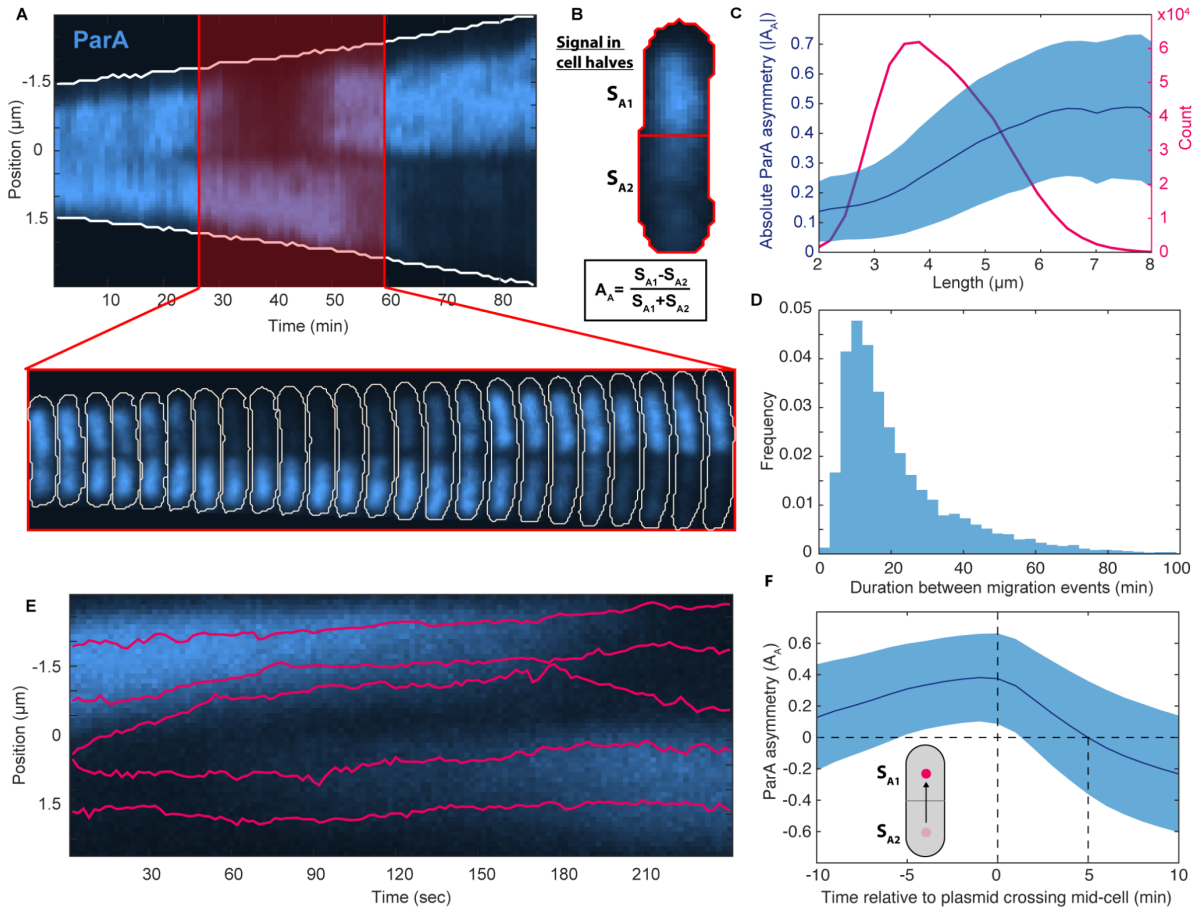

Figure 1. Characterizing ParA migration events and their dependence on plasmid movements. **(A)** Top: Intensity profile of ParA-mVenus along the long axis of one cell (1 minute time intervals). Bottom: Single cell overview of the red highlighted

region. **(B)** The cell is split in the middle into two halves. The signal of ParA-mVenus is measured in both halves (S_{A1}, S_{A2}) to calculate how asymmetric ParA is distributed: $A_A = \frac{S_{A1} - S_{A2}}{S_{A1} + S_{A2}}$. **(C)** Blue: Absolute ParA asymmetry of cells plotted against length (mean ± SD). Red: Number of data points at each length. **(D)** Histogram of the time between migration events (see Figure S2A for the definition of migration event). **(E)** Intensity profile of ParA-mVenus along the long axis of one cell (2 second time intervals). Red lines are plasmid trajectories. **(F)** The change in ParA asymmetry relative to a plasmid crossing the middle of the cell (mean ± SD). The cells were aligned such that S_{A1} corresponds to the signal of ParA in the cell half that the plasmid crossed into. The panels **(C)**, **(D)** and **(F)** were created from a dataset containing 5357 cell cycles.

5.1.3.2. Odd vs Even: How Nucleoid Constriction Alters ParA Asymmetry

In this section, we focused on the interplay between ParA, ParB, and the nucleoid (Figure 2A) and found that the asymmetry of ParA increases after the formation of the nucleoid constriction and the nucleoid has a bilobed shape. The nucleoid constriction forms before the completion of replication and segregation of the newly replicated sister nucleoids (see Figure S4 and materials and methods for nucleoid constriction detection).

Before the constriction, ParA asymmetry remained constant; however, it increased following its formation (Figure 2B), suggesting a strong connection between the migration events of ParA and the nucleoid constriction. Additionally, the nucleoid constriction introduced a difference between cells with even and odd numbers of plasmids (Figure 2C). Before constriction formation, no differences in asymmetry were observed between cells with odd and even numbers of plasmids (Figure 2C, inset), but after constriction formation, cells with odd numbers of plasmids exhibited greater ParA asymmetry based on cell length.

This raises the question: why does the nucleoid constriction distinctly affect cells with even and odd numbers of plasmids? An analysis of the spatial distribution of plasmids in cells containing three plasmids offers an explanation. Before constriction formation, plasmids were regularly positioned with one peak at the middle and two peaks at the 1/6 and 5/6 positions (Figure S5A, Top). This distribution changed after constriction formation, resulting in the middle peak splitting into two (Figure S5A, bottom). The observed shift in plasmid localization can be attributed to a decreased amount of DNA at the center of the cell, which is visible as the nucleoid constriction. This reduced amount of DNA led to a decrease in the amount of nucleoid-bound ParA at the cell center, prompting any plasmid situated there to relocate, typically to either side of the constriction where the concentration of ParA is higher. Consequently, after the constriction forms, the middle peak disappeared and two new peaks (left and right to the constriction) emerged. This occurred exclusively in cells with an odd number of plasmids, as regularly positioning an even number of plasmids will never result in one plasmid being situated at the center of the cell (Figure S5B). Another interesting occurrence in cells with an odd number of plasmids was the frequent traversal of the central plasmid from one lobe of the nucleoid to the other (Figure S5C). This movement is driven by the majority of

ParA being positioned opposite to the central plasmid and therefore attracting it to cross the nucleoid constriction. The crossing of the plasmid triggers a migration event, leaving the plasmid once more on the ParA-depleted side, and the cycle repeats. As a result, the central plasmid oscillates between cell halves, which is a common occurrence in our experimental data.

Greater imbalance between plasmids in cell halves resulted in higher asymmetry, as evidenced by comparing cells with the same number of plasmids but varying degrees of plasmid imbalance (Figure 2D). In all cases, as imbalance increased, so did asymmetry. Interestingly, regular positioning continued to persist after nucleoid constriction formation. Plasmids were no longer regularly positioned along the entire cell but were distributed regularly within each lobe of the nucleoid. This is clearly visible in cells with three (1:2) and five (2:3) plasmids (Figure 2D), which exhibited different numbers and locations of peaks in both cell halves, all of which were regularly positioned within their half of the cell.

In summary, there is a strong connection between ParA asymmetry and nucleoid constriction, with the latter not only increasing asymmetry but also affecting the localization of plasmids. Further, despite changes in localization, plasmids maintain regular positioning within their respective lobe of the nucleoid following the formation of a constriction.

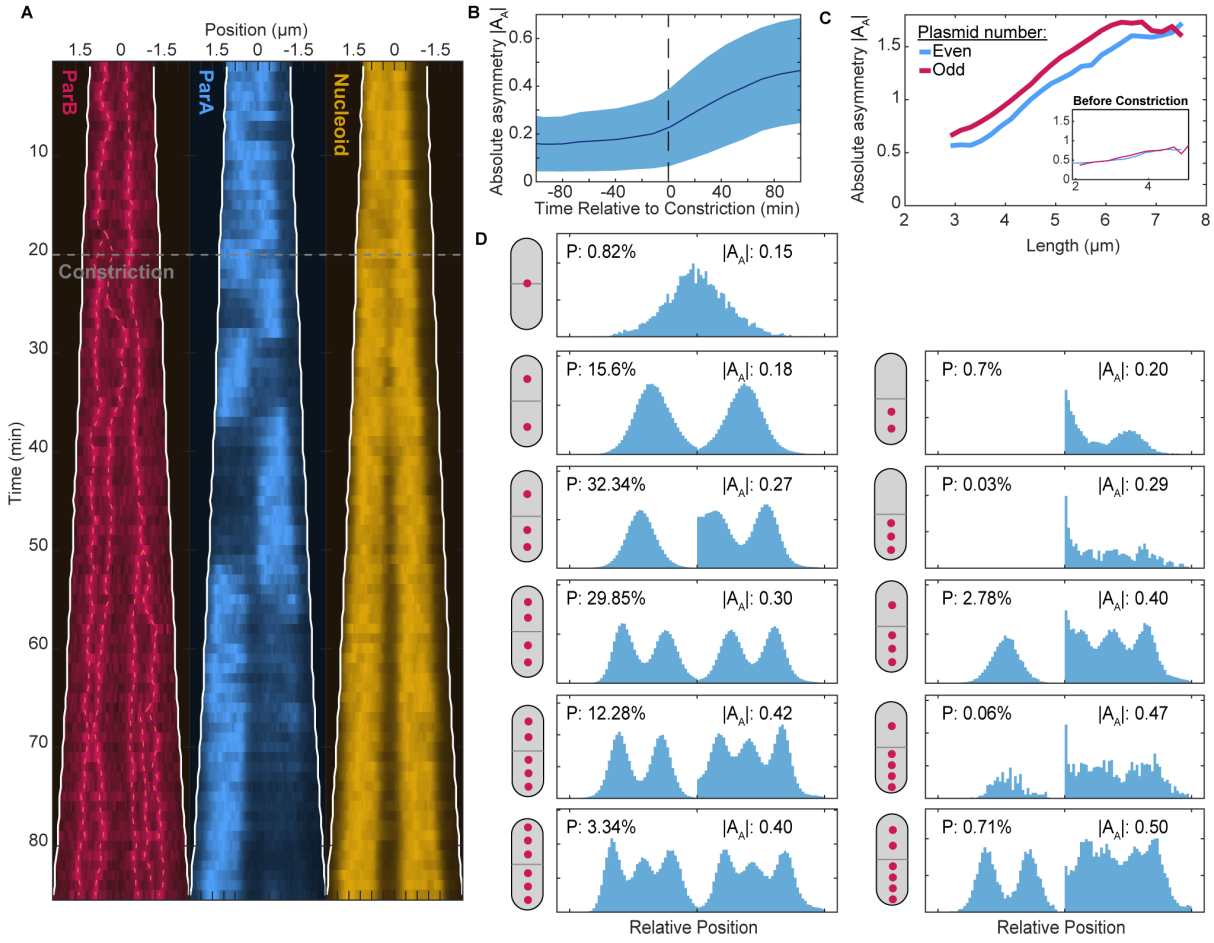

Figure 2. Nucleoid constriction is necessary for stable ParA oscillations. **(A)** Kymograph of Nucleoid (HU-mCherry), ParA-mVenus, ParB-mTurquoise2 of a cell. Red dashed lines indicate tracked foci. The gray dashed line denotes the first frame in which the nucleoid constriction was detected. **(B)** Average ParA asymmetry (as in Figure 1D) relative to nucleoid constriction (mean ± std). **(C)** The absolute asymmetry of ParA for cells with even and odd numbers of plasmids after the formation of the nucleoid constriction. Inset: The same but prior to the constriction. **(D)** Histograms of the distribution of plasmids for different configurations of plasmids as indicated by the cartoon on the left of each plot. The values presented within the plot correspond to the population size (P) and the average absolute ParA asymmetry ($|A_A|$) of the corresponding configuration. The panels **(B)**, **(C)** and **(D)** were created from a dataset containing 5357 cell cycles.

5.1.3.3. The ParA-ParB Dance: A Potential Trigger for ParA Migration Events

What triggers a migration of ParA from one cell half to the other? As previously mentioned, 53% of migration events were preceded by a plasmid crossing mid-cell (Figure S3). ParB stimulates the hydrolysis of ParA and when a plasmid crosses the mid-cell, a significant amount of plasmid-bound ParB is relocated from one side of the cell to the other. This changes the amount of ParA being hydrolyzed in each half of the cell such that the concentration of ParA is lower in the cell half with more ParB.

To explore the possibility that ParB displaces its own attractor, ParA, and therefore triggers migration events, we examined the relationship between the asymmetry of ParA and ParB. By plotting the time series of the asymmetry values of ParA and ParB for one cell cycle, we can visualize how they affect each other. If the relocation of a significant amount of ParB between cell halves triggers the migration of ParA away from the side the ParB was relocated to and if ParB is attracted to ParA, we anticipate a counterclockwise spiral (illustrated as a cartoon in Figure 3 A,B). Indeed, our analysis of real cell cycles revealed this pattern, after nucleoid constriction formation (Figure 3 C,D). By calculating the angular velocity of the spiral post-constriction, we can determine its spin. A positive angular velocity indicates a counterclockwise spin, while a negative value denotes a clockwise spin. The majority of cell cycles exhibited a positive angular velocity and, by extension, the counterclockwise spiral pattern (Figure 3 E), explicitly demonstrating that ParB displaces ParA while being attracted to it, consistent with previous literature.

In conclusion, we propose that the relocation of ParB, triggered by plasmids crossing the mid-cell in a cell with a constricted nucleoid, results in varying amounts of ParA being hydrolyzed between the nucleoid lobes. This, in turn, leads to the migration of ParA towards the side with less ParB. In the absence of a plasmid crossing event, fluctuations in the amounts of ParB bound to the plasmid could also drive the migration of ParA.

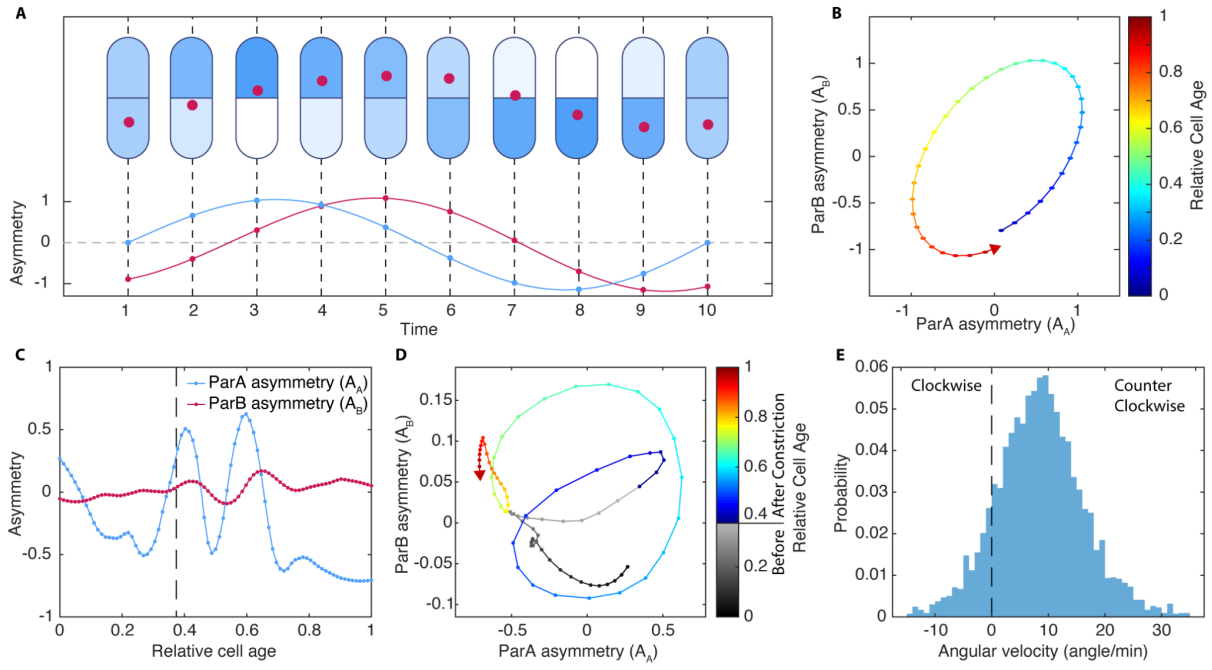

Figure 3. The dance of ParA and ParB. **(A)** Cartoon representation of the asymmetry between ParA (blue) and ParB (red) in a bacterial cell. Top: A cell with one ParB-coated plasmid actively chasing ParA. Bottom: the time series of the asymmetry of both ParA (A_A) and ParB (A_B). **(B)** Asymmetry plot: Trajectory of the interplay between ParA and ParB asymmetry. The color shows relative cell age ranging from 0 (birth) to 1 (division). **(C)** Asymmetry of ParA and ParB of a real cell cycle. The dashed line indicates the time point when a nucleoid constriction was first detected. Both curves were smoothed using a

robust locally weighted scatterplot smoothing algorithm with a smoothing parameter of 0.2 to transform the curves into more wave-like forms. **(D)** The asymmetry plot of the same cell cycle as in **(C)**. The color shows relative cell age, while gray indicates data points prior to nucleoid constriction detection. **(E)** A histogram of the angular velocity of 5357 cell cycles, after the nucleoid constriction was detected, of the trajectories as shown in **(D)**.

5.1.3.4. Putting the Par back into ParABS

We hypothesize that the system exhibits two distinct modes of plasmid positioning: (1) Prior to the formation of the nucleoid constriction, the plasmids are regularly positioned across the nucleoid to ensure equal distribution of plasmids among the sister nucleoids following replication; (2) Post constriction formation, regular positioning is no longer a priority. Each daughter cell will inherit one nucleoid and the focus shifts towards even partitioning of plasmids, first between the lobes of the constricted nucleoid and afterward between sister nucleoids. The increased frequency of ParA migration events after the formation of the nucleoid constriction supports this kind of switch from regular positioning to a partitioning mechanism (although regular positioning is still maintained on each nucleoid separately). To test this hypothesis, we introduced an accurate representation of the nucleoid into our previously developed model of the ParABS system. Our simulations demonstrate that, in the absence of additional mechanisms and with only regular positioning, plasmids are unable to traverse between the lobes of the nucleoid once the constriction is severe enough (Figure S6A). Therefore, we believe that the migration of ParA is needed for partitioning plasmids between lobes and, once replicated, between sister nucleoids. This process ensures even partitioning and, as a result, promotes equal inheritance of plasmids to the daughter cells.

But how is ParA able to exclusively coat one lobe of the nucleoid or, if fully replicated, only one of the sister nucleoids? Currently, there exists no model (including ours, Figure S6B,C) capable of reproducing this behaviour of ParA, implying that some fundamental aspect of the ParABS system is still missing. A potential mechanism for achieving this asymmetry between sister nucleoids is ParA-ParA recruitment, similar to how recruitment is required for the MinCDE system to oscillate. ParA-ParA recruitment describes the process of nucleoid-bound ParA interacting with cytosolic ParA, promoting the cytosolic ParA to bind in close proximity to the nucleoid-bound ParA (Figure 4 A). To achieve this, the interaction between the ParA may assist the cytosolic ParA to attain its DNA-binding competent state or, if already in this state, may increase its DNA-binding affinity. Independent of the exact mechanism, ParA-ParA recruitment creates a positive feedback loop in which less ParA is kicked off the nucleoid with fewer ParB-coated plasmids due to ParB-induced hydrolysis, resulting in a higher concentration of ParA on the same nucleoid compared to the other nucleoid. This higher concentration, in turn, attracts more ParA through recruitment, further amplifying the asymmetry (Figure 4 B). This loop continues and allows the majority of ParA to be located on one nucleoid until a plasmid, attracted by the significantly higher ParA concentration, traverses the gap between sister nucleoids. To summarize, the nucleoid with fewer plasmids contains the majority of nucleoid-bound

ParA, which attracts plasmids from the other nucleoid. The addition of this mechanism allows the system to ensure an even distribution of plasmids between nucleoids, since the migration of ParA will continue until the number of plasmids is balanced between nucleoids.

Incorporating ParA-ParA recruitment into our improved model allows it to produce ParA localization patterns almost identical to the ones observed *in vivo* (Figure 4C,D,E). In the absence of nucleoid constriction, ParA is evenly distributed across the nucleoid (Figure 4C). As nucleoid constriction forms, ParA migrates in a chaotic fashion from one side of the cell to the other (Figure 4D). Once a substantial constriction has formed, migration events become more stable, leading to prolonged asymmetric states of ParA (Figure 4E). Similar to our *in vivo* observations, migration events are caused by plasmids crossing mid-cell.

Here, we propose a model with two distinct modes of plasmid positioning: (1) regular positioning prior to nucleoid constriction; (2) partitioning post-constriction. ParA-ParA recruitment is a crucial mechanism for achieving these two different modes, as it allows ParA to migrate between cell halves post-constriction. Incorporating an accurate representation of the nucleoid and ParA-ParA recruitment in our previously developed model led to ParA localization patterns identical to those observed *in vivo*.

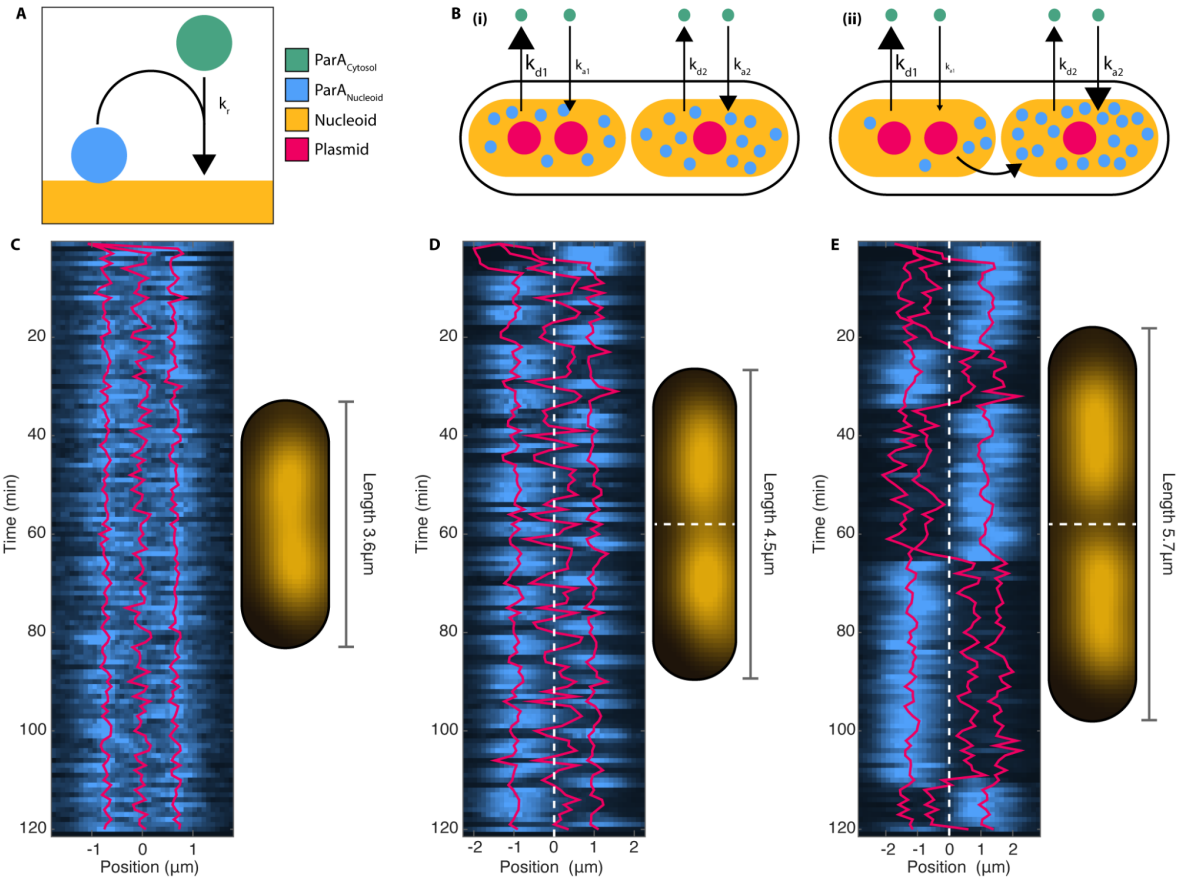

Figure 4. ParA recruitment is sufficient to produce ParA migration. **(A)** A cartoon of nucleoid-bound ParA recruiting cytosolic ParA. **(B)** A cartoon of plasmid-induced ParA migration in a cell with two nucleoids. **(i)** Plasmids contain ParB which stimulates ParA hydrolysis, therefore increasing the rate of ParA dissociation from the nucleoid. The nucleoid with more plasmids has a higher dissociation rate (k_{d1}) than the other nucleoid (k_{d2}) and therefore contains less nucleoid-bound ParA. This difference in ParA is further increased by ParA recruitment. Since the other nucleoid has more ParA, its association rate (k_{a2}) of cytosolic ParA is greater than the rate of the other nucleoid (k_{a1}). **(ii)** This causes ParA to migrate, further increasing the difference in association rate between the nucleoids. Once the majority of ParA is bound to the nucleoid with fewer plasmids, the plasmid on the other nucleoid, the central plasmid, is attracted by the ParA and starts moving towards the other nucleoid. **(C)** Simulations using our previous model with the inclusion of ParA recruitment and accurate depiction of the nucleoid. Left, kymograph of ParA dynamics. Right, nucleoid profile. **(D)** Same as in **(C)** but with a nucleoid that is biloped. **(E)** Same as in **(C)** but with two separated sister nucleoids.

5.1.3.5. The Role of ParA Oscillations in the ParABS System

We postulate that ParA oscillations are needed to deal with plasmid replication after the nucleoid constriction is formed. There are different growth conditions leading to different timings of nucleoid constriction formation (Govers et al., 2023). In fast growth conditions, the cell is born with two fully replicated nucleoids due to overlapping replication cycles. Without any exchange of plasmids between nucleoids in these kinds of growth conditions, plasmid replication can result in a runaway reaction, leading to one nucleoid accumulating a disproportionate amount of plasmids. Each plasmid has the same probability to replicate, which means that replication events are more likely to occur on the nucleoid with more plasmids, leading to an even higher chance of subsequent replications on the same nucleoid. As mentioned above, depending on growth conditions, the nucleoid may either separate late during the cell cycle, in which case regular positioning alone would suffice to ensure stability, or it may already be replicated and separated at birth. In the latter scenario, plasmid loss will occur (Figure 5A) without a mechanism to transfer plasmids between nucleoids and prevent the problem of self-reinforcing asymmetry by replication. Migration of ParA provides a solution to this issue by moving plasmids to the less occupied nucleoid and thereby causing an even partitioning of plasmids.

To assess the impact of nucleoid constriction timing on plasmid loss, we performed simulations across 10 generations under two conditions: with ParA-ParA recruitment and without it (Figure 5B). The simulations encompassed six different nucleoid constriction timings and employed the following simplifications: the nucleoid instantly constricts and separates at the specified cell age, all cells have the same width and length (1.0 μm x 5.8 μm), and all cells have the same cycle duration of the population mean (113.7 min). Our results show the crucial importance of constriction timing in the system, as evidenced by the significantly higher loss rate of plasmids in the non-recruiting ParA condition compared to the recruiting ParA condition when the nucleoid constriction forms early (<0.2) in the cell cycle. However, when constriction occurs late in the cell cycle, the differences in loss rate between the two conditions become negligible (>0.2). ParA migration mitigates plasmid loss events, ensuring stable and accurate segregation of plasmids during cell division. Furthermore, the ParA-ParA

recruitment reduces the variation of plasmid number within the populations (Figure 5B). It is only when the constriction forms very late in the cell cycle (>0.8) that the standard deviation of the number of plasmids between the two conditions is equal. This is attributed to the inability of the system to partition plasmids between nucleoids, introducing higher asymmetry when it comes to distributing plasmids to the offspring.

These findings support our hypothesis that an additional mechanism, such as ParA-ParA recruitment, is required to prevent plasmid loss in cells with early nucleoid constriction formation and keep the distribution of the number of plasmids within the population tight. We conclude that, to ensure plasmid stability across various growth conditions, an active translocation mechanism between cell halves is needed post nucleoid constriction.

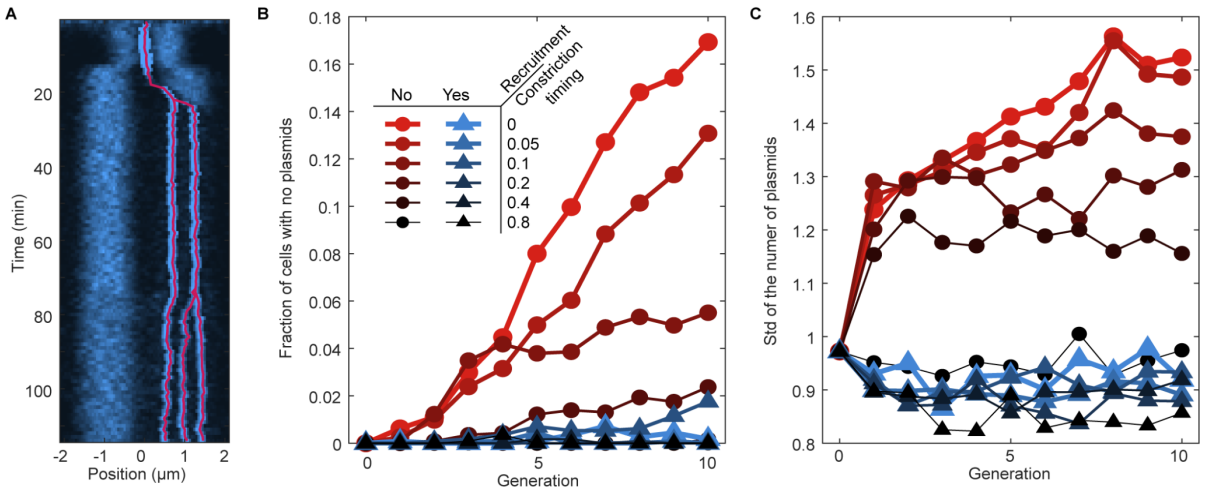

Figure 5. Simulations show that ParA-ParA recruitment significantly reduces plasmid loss rate. **(A)** Example of a simulated cell without ParA-ParA recruitment, producing offspring which lost the plasmid. **(B)** Simulations of the ParABS system, with (red circles) and without (blue triangles) ParA-ParA recruitment, are conducted across 6 distinct nucleoid formation timings (0 birth, 1 division) spanning 10 generations. Each data point represents 1000 simulated cells. At the end of every generation, the final localizations of plasmids of each cycle are used to determine the plasmid inheritance distribution to the next generation. From this distribution, the initial plasmid numbers for 1000 new cell cycles in the subsequent generation are selected. The frequency and timing of plasmid replication are randomly selected from our experimental data. The initial distribution of plasmids at generation 0 is equal to the distribution of plasmids at birth from our experimental data. The fraction of cells without plasmids is plotted for each generation to visualize the results. **(C)** Same as **(B)** but instead of the fraction of cells with no plasmid, the standard deviations of the distribution of plasmid inheritance at the end of a generation are shown.

5.1.4. Discussion

Over the last several decades, extensive research on the ParABS system has shed light on numerous aspects of its function and role in plasmid segregation. Several breakthroughs have been made regarding ParB, such as the discovery that it possesses a CTP-binding domain and that DNA-bound ParB can recruit cytosolic ParB to bind in close proximity when bound to DNA (Tišma et al., 2022).

Despite these advances, many aspects of the ParABS system, particularly the dynamics of ParA, remain elusive. To date, no theory exists with a comprehensive explanation for ParA oscillations without certain limitations or shortcomings. Previous attempts have posited that a plasmid follows a retracting gradient, which in turn causes ParA oscillations (Ah-Seng et al., 2013; Ietswaart et al., 2014; Szardenings et al., 2011). This notion, however, contradicts our *in vivo* observations, where in 47% of the cases (Figure S3) ParA migrates between cell halves without any accompanying plasmid movement. Moreover, the functional significance of ParA oscillations remains unclear, and their necessity in the cellular context is unknown. The objective of this study was to address these existing gaps in our comprehension of ParA.

In this study, we quantitatively examined and characterized the localization patterns of ParA, discovering that ParA oscillations begin after the nucleoid constricts and that they correlate with plasmids crossing mid cell. We propose that ParA-ParA recruitment is necessary for these oscillations and that this recruitment, and consequently the ParA oscillations, can prevent plasmid loss. We support these claims by incorporating ParA-ParA recruitment and an accurate representation of the nucleoid into our previously developed model. We demonstrate that these features are sufficient to reproduce the spatial and temporal behaviour of ParA observed *in vivo*.

We find that ParA-ParA recruitment is a highly plausible mechanism, as evidenced by analogous examples such as MinD and, more recently, ParB. These instances showcase similar behaviors where a cytosolic protein is recruited by a membrane-bound or DNA-bound protein. Moreover, *in vitro* studies have revealed that ParA can form filaments, indicating its potential for self-interaction beyond mere dimer formation under suitable conditions. These observations might be an artifact of ParA-ParA recruitment. Nevertheless, the hypothesis that ParA possesses such an ability seems very likely, considering the evidence from this study, previous studies and examples of other proteins displaying the similar behaviours. Future studies can aim to validate our hypothesis by attempting to identify a specific domain, if it exists, that is necessary for ParA-ParA recruitment. Discovering such a domain could have significant implications not only for plasmids but also for chromosomal ParABS systems. By further investigating the ParA-ParA interaction and its underlying molecular mechanisms, we can deepen our understanding of the ParABS system and its role in the segregation of various different cargos, ultimately expanding the current knowledge of bacterial plasmid segregation and chromosomal partitioning.

Nonetheless, it should be noted that there exist alternative mechanisms to ParA-ParA recruitment, also capable of producing the observed ParA localization patterns. For example, one alternative involves nucleoid-bound ParA interacting with other nucleoid-bound ParA to stabilize and reduce its basal hydrolysis rate. In cells with two nucleoids, this enables ParA to remain longer on the nucleoid with a higher concentration of ParA, creating a bistable system reminiscent of the *in vivo* observations. Yet,

this mechanism seems unlikely, given the brief migration period during which the majority of ParA shifts from one cell half to the other.

Previous studies modelling the ParABS system have assumed that their implementation is sufficient to ensure faithful segregation (Hu et al., 2017a; Surovtsev et al., 2016). However, these studies failed to recognize the importance of an accurate representation of the nucleoid. In all instances, the nucleoid was simply represented as a box with uniform DNA-density. Under such conditions, regular or approximate positioning is sufficient to ensure faithful segregation, as corroborated by our simulations where the nucleoid constriction forms at the end of the cell cycle. It is only through the incorporation of an early constricted nucleoid that it becomes evident that the existing mechanisms are not adequate for reliable segregation.

An intriguing aspect, not analyzed in this study, is the asymmetric inheritance of ParA. We frequently observed that ParA was solely inherited by one daughter cell, leaving the other without any ParA. This phenomenon is quite curious, given that ParA is a critical component of the system and is required for accurate positioning. However, the cell half without ParA most often contains the majority of plasmids. Consequently, even if a daughter cell does not inherit any ParA, it possesses the bulk of the plasmids which can be used to produce new ParA. In line with this, we noticed cells that lost all ParA but managed to recover within one or two generations, eventually containing the population average of ParA again.

In conclusion, this study has taken important steps toward unraveling the complexities of the ParABS system, particularly in the context of ParA dynamics. We propose that ParA-ParA recruitment is sufficient to reproduce our *in vivo* observations and provide a rationale for the biological relevance of ParA oscillations. Our findings have the potential to significantly impact our understanding of plasmid segregation and, by extension, chromosomal partitioning done by ParABS systems in bacteria. While our work has addressed some of the gaps in our knowledge, it has also paved the way for future investigations to further validate and refine our proposed mechanism.

5.1.5. Material and Methods

Strains and growth condition, Microfluidics, Microscopy and Image Processing are similar to our previous paper (Köhler et al., 2022) except three slight differences: (1) The strain DLT3125 was replaced by strain DLT31XX which has two additional fluorophores; (2) During microscopy in addition to the YFP-signal the CFP- and RFP-signal were also captured; (3) The end of the model section incorporates an explanation of how the nucleoid and ParA-ParA recruitment were implemented.

5.1.5.1. Strains and growth condition

F plasmid experiments use strain DLT3059 with containing the plasmid pJYB249 (Sanchez et al., 2015), a derivative of the *E. coli* K-12 strain DLT1215 (Bouet et al., 2005) containing the mini-F plasmid derivative pJYB234. This plasmid carries a functional ParB-mTurquoise2 and ParA-mVenus fusion. Overnight cultures were grown at 37 °C in LB-Media containing 10 μg/ml thymine + 10 μg/ml chloramphenicol.

5.1.5.2. Microfluidics

Like the original mother machine (P. Wang et al., 2010), our design consists of a main channel through which nutrient media flows and narrow growth-channels in which cells are trapped. However, we follow (Baltekin et al., 2017) and include (i) a small opening at the end of each growth channel (ii) a waste channel connected to that opening to allow a continuous flow of nutrients through the growth channels (iii) an inverted growth-channel that is used to remove the background from fluorescence and phase contrast. We used a silicon wafer with this design to create the mother machine. We poured a polydimethylsiloxane (PDMS) mixture composed of a ratio of 1:7 (curing agent:base) over the wafer and let it rest at low pressure in a degasser for ~30 min to remove air bubbles inside. The PDMS was then baked at 80 °C overnight (~16 h). The cured PDMS was peeled off the wafer. Before imaging, the chip is bonded to a glass slide using a plasma generator (30 s at 75 W) and subsequently baked for a further 30 min at 80 °C, while the microscope is prepared.

5.1.5.3. Microscopy

We used a Nikon Ti microscope with a 100 x/1.45 oil objective and a Hamamatsu Photonics camera for all imaging. For imaging cells of strain DLT3125 we used a mother machine. Overnight cultures were inoculated into fresh media (M9+0.5% glycerol + 0.2% casamino acids + 0.04 mg/mL thymine + 0.2 mg/mL leucine + 10 μg/mL chloramphenicol) for 4 hours at 30 °C before imaging. Cells were loaded into the chip through the main channel and the chip was placed into a preheated microscope at 30 °C. The cells were constantly supplied with fresh media by pumping 2 μL/min of M9+0.5% glycerol + 0.2% casamino + 0.04 mg/mL thymine + 0.2 mg/mL leucine through the microfluidic chip. Cells were grown for 2 hr inside the microscope before imaging. Cells were imaged at 1 minute intervals for approximately 72 hr. Phase contrast, RFP-signal, CFP-signal and YFP-signal were captured. Imaging was repeated independently with similar results.

5.1.5.4. Image processing

Our image processing pipeline for mother-machine experiments consists of three parts: (I) preprocessing, (II) segmentation and foci finding, and (III) cell and foci tracking. While Parts I and III

use custom Matlab scripts, Part II is based on SuperSegger (Stylianidou et al., 2016) (Stylianidou et al., 2016), a Matlab-based package for segmenting and tracking bacteria within microcolonies (original code is available at https://github.com/wiggins-lab/SuperSegger), that we modified to better handle high-throughput data. SuperSegger employs pre-trained neural networks to segment cells by identifying their boundaries. It comes with a pre-trained model for E. coli which worked very well with our data. Therefore there was no need to train our own neural network. SuperSegger is capable of tracking cells however the tracking did not work properly with mother-machine images and so we developed our own method. Nevertheless, acknowledging that one of the main components of our pipeline, the segmentation, uses SuperSegger we refer to the entire pipeline as MotherSegger (code is available at https://gitlab.gwdg.de/murray-group/MotherSegger/-/tree/PaperParABS).

In Part I, each frame of an acquired image stack is aligned (the offset between frames in x and y is removed). Afterwards the image stack is rotated so the growth channels are vertical. A mask of the mother machine layout is fitted to the phase contrast, using cross-correlation, to identify where the growth channels are located. Each growth channel is extracted from the image stack and the flipped inverted channel is subtracted to remove the background from both the fluorescence signal and phase contrast. The images are then segmented and fluorescent foci are identified using Supersegger.

In Part III, both foci and cells are tracked. Since cells cannot change their order inside the growth channel, they can be tracked by matching similar cell length between frames (starting from the bottom of the growth channels). Once individual cell cycles are identified, the foci positions found by Supersegger are re-specified relative to the bounding box of the cell (the smallest rectangular image containing the cell mask) on each frame. Since cells are vertical in the channels without any significant tilting, the bounding box is aligned with the cell axes. Within each cell cycle, foci are tracked between frames by finding the closest focus on the next frame inside the same cell cycle. The effect of growth on foci position was neglectable since cells grew on average much less than one pixel per frame at the 1 min frame rate and 100 min doubling time used here. Finally, half the cell length was subtracted from the foci positions along the long cell axis (vertical direction) so that 0 corresponds to the middle of the cell. The sign of the positions was also adjusted so that negative positions refer to the old-pole proximal side of the cell.

To filter out potential segmentation errors, cell cycles that do not have exactly 1 parent and 2 daughters are excluded from analysis along with their immediate relatives (with the exception of those who are pushed out of the growth channel). For the analysis of foci trajectories, we considered only trajectories coming from at least 12 consecutive frames with the same number of foci. For pB171, we used (unmodified) SuperSegger to process images of cells growing on agarose pads.

Foci were tracked using ⋆Track (https://gitlab.gwdg.de/murray-group/StarTrack).

5.1.5.5. Model

Our model is an extension of the previous DNA-relay model (Surovtsev et al., 2016a) that incorporates diffusion on the nucleoid (hopping) and basal hydrolysis of ParA-ATP and uses analytic expressions for the fluctuations rather than a second order approximation. Like the DNA relay it is a 2D off-lattice stochastic model and updates positions in discrete time steps dt. The implementation was written in C++ (code is available at https://gitlab.gwdg.de/murray-group/hopping_and_relay/-/tree/PaperParABS). It consists of the following components. ParA associates to the DNA non-specifically in its ATP-dependent dimer state with the rate k_a. Once associated, ParA (i.e. ParA-ATP dimers) moves in two distinct ways: (i) Diffusive motion on the nucleoid with the diffusion coefficient D_h. This is an effective description of the movement of dimers due to transient unbinding events that allow them to 'hop' between DNA-strands. We do not consider the alternative scenario in which dimers transfer between DNA strands when the latter come into contact. In this scenario the effective diffusion coefficient would depend on the parameters describing the DNA fluctuations (D_A and $\sigma_{x,y}$). (ii) Between hopping events, each bound ParA dimer experiences the elastic fluctuations of the DNA strand it is bound to. This is implemented as elastic (spring-like) fluctuations around its initial position. Dimers dissociate from the nucleoid due to either basal ATP hydrolyse at a rate k_d or due to hydrolysis stimulated by ParB on the plasmid. The latter is modelled as a ParB-coated disc and ParB-ParA tethers form whenever the disk comes in contact with a ParA dimer. ParB-stimulated hydrolysis then breaks these tethers at a rate k_h, returning ParA to the cytosolic pool. The plasmid experiences the elastic force of every tethered ParA and moves according the its intrinsic diffusion coefficient D_p and the resultant force of all tethers. An overview of this scheme is shown in Figure 3A.

As in the DNA relay model we have made some simplifications that we next make explicit. First, we only modelled three states of ParA: 'nucleoid associated' and 'cytosolic' and 'tethered'. Second, cytosolic ParA are assumed to be well mixed. This is justified based on the slow conformation changes needed to return it to a state competent for DNA-binding (Vecchiarelli et al., 2010). Third, no individual ParB molecules were modelled, rather the plasmid is treated as a disk coated with enough ParB that each nucleoid-bound ParA that makes contact with the plasmid instantaneously finds a ParB partner, therefore removing the need to model individual ParB. This is justified by the substantially higher local concentration of ParB compared to ParA at the plasmid.

The nucleoid is modelled as a rectangle with the dimensions $L \times W$. The positions of ParA and the plasmid(s), are updated every time step dt as follows. Between hopping events, each nucleoid associated ParA dimer fluctuates about a home position x_h. The new position $x(t + dt)$ of each dimer is given by $x(t + dt) = x_h + \delta x$, where δx is drawn with probability $p(\delta x, dt \mid x(t) - x_h)$ where x(t) is its original position (see section 'Over-damped spring') and the normalised spring constant (f/k_BT above) along each dimension is $1/\sigma_{x,y}^2$ and the diffusion coefficient D_A. During hopping events $x(t)$ and x_h are both offset by a value drawn from a Gaussian distribution with $\mu = 0$ and $\sigma = \sqrt{2D_h dt}$ for both dimensions. The displacement of the plasmid is determined similar to each ParA dimer but according to the resultant force acting on it. This

resultant force vector has an effective spring constant equal to the spring constant of a single tether times the number of tethers and acts towards an equilibrium position $x_p(t) + \sum_{tethers}(x_h - x(t))/n$, where $x_p(t)$ is the plasmid position and the sum is over all (n) tethers. We ignore the effects of Torque. The intrinsic diffusion coefficient of the plasmid is D_p. If the plasmid has no tethers attached then it moves by normal diffusion, with displacements drawn from a Gaussian distribution with $\mu = 0$ and $\sigma = \sqrt{2D_p dt}$. The x and y components of all positions are updated independently and all simulations in this paper were run until the system reached equilibrium before acquiring data used for analysis.

ParA-ParA recruitment is implemented as such: Each nucleoid bound ParA has the option to recruit, with the rate k_r, a cytosolic ParA based on the amount of cytosolic ParA.

To implement an accurate depiction of the nucleoid into our simulations we used the nucleoid profile from a cell from our experiment. We performed some smoothing to the nucleoid, cut off the upper and lower quantile and normalized output such that the lowest value is equal to 0 and the highest value is equal to 1. This representation is then used as a field that influences the diffusion of ParA on the nucleoid. It uses the first order approximation similar to (Branka & Heyes, n.d.)in combination with the effector factor F_n to bias the movement of the ParA to go up the gradient of the field.

5.1.6. References

Ah-Seng, Y., Rech, J., Lane, D., & Bouet, J.-Y. (2013). Defining the Role of ATP Hydrolysis in Mitotic Segregation of Bacterial Plasmids. *PLoS Genetics*, *9*(12), 12. https://doi.org/10.1371/journal.pgen.1003956

Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I., & Elf, J. (2017). Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proceedings of the National Academy of Sciences*, *114*(34), 9170–9175. https://doi.org/10.1073/pnas.1708558114

Branka, A. C., & Heyes, D. M. (n.d.). *Algorithms for Brownian dynamics simulation*.

de Boer, P. A. J., Crossley, R. E., & Rothfield, L. I. (1989). A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli. *Cell*, *56*(4), 641–649. https://doi.org/10.1016/0092-8674(89)90586-2

Ebersbach, G., & Gerdes, K. (2001). The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(26), 26. https://doi.org/10.1073/pnas.261569598

Ebersbach, G., & Gerdes, K. (2004). Bacterial mitosis: Partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Molecular Microbiology*, *52*(2), 2. https://doi.org/10.1111/j.1365-2958.2004.04002.x

Ebersbach, G., Ringgaard, S., Moller-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. *Molecular Microbiology*, *61*(6), 6. https://doi.org/10.1111/j.1365-2958.2006.05322.x

Gerdes, K., Moller-Jensen, J., & Jensen, R. B. (2000). Plasmid and chromosome partitioning: Surprises from phylogeny. *Molecular Microbiology*, *37*(3), 3. https://doi.org/10.1046/j.1365-2958.2000.01975.x

Govers, S., Campos, M., Tyagi, B., Laloux, G., & Jacobs-Wagner, C. (2023). Apparent simplicity and emergent robustness in bacterial cell cycle control. *BioRxiv*. https://doi.org/10.1101/2023.01.16.524295

Guilhas, B., Walter, J.-C., Rech, J., David, G., Walliser, N. O., Palmeri, J., Mathieu-Demaziere, C., Parmeggiani, A., Bouet,

J.-Y., Le Gall, A., & Nollmann, M. (2020). ATP-Driven Separation of Liquid Phase Condensates in Bacteria. *Molecular Cell*, *79*(2), 2. https://doi.org/10.1016/j.molcel.2020.06.034

Halatek, J., & Frey, E. (2012). Highly Canalized MinD Transfer and MinE Sequestration Explain the Origin of Robust MinCDE-Protein Dynamics. *Cell Reports*, *1*(6), 741–752. https://doi.org/10.1016/j.celrep.2012.04.005

Hu, L., Vecchiarelli, A. G., Mizuuchi, K., Neuman, K. C., & Liu, J. (2017). Brownian Ratchet Mechanism for Faithful Segregation of Low-Copy-Number Plasmids. *Biophysical Journal*, *112*(7), 7. https://doi.org/10.1016/j.bpj.2017.02.039

Huang, K. C., Meir, Y., & Wingreen, N. S. (2003). Dynamic structures in *Escherichia coli*: Spontaneous formation of MinE rings and MinD polar zones. *Proceedings of the National Academy of Sciences*, *100*(22), 12724–12728. https://doi.org/10.1073/pnas.2135445100

Ietswaart, R., Szardenings, F., Gerdes, K., & Howard, M. (2014). Competing ParA structures space bacterial plasmids equally over the nucleoid. *PLoS Computational Biology*, *10*(12), 12. https://doi.org/10.1371/journal.pcbi.1004009

Köhler, R., Kaganovitch, E., & Murray, S. M. (2022). High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS. *ELife*, *11*, e78743. https://doi.org/10.7554/eLife.78743

Köhler, R., & Murray, S. M. (2022). *High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS* [Preprint]. Biophysics. https://doi.org/10.1101/2022.03.29.486196

Lim, G. E., Derman, A. I., & Pogliano, J. (2005). Bacterial DNA segregation by dynamic SopA polymers. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(49), 49. https://doi.org/10.1073/pnas.0507222102

Lim, H. C., Surovtsev, I. V., Beltran, B. G., Huang, F., Bewersdorf, J., & Jacobs-Wagner, C. (2014). Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *ELife*, *3*(3), 3. https://doi.org/10.7554/eLife.02758

Osorio-Valeriano, M., Altegoer, F., Steinchen, W., Urban, S., Liu, Y., Bange, G., & Thanbichler, M. (2019). ParB-type DNA Segregation Proteins Are CTP-Dependent Molecular Switches. *Cell*, *179*(7), 1512-1524.e15. https://doi.org/10.1016/j.cell.2019.11.015

ParB proteins can bypass DNA-bound roadblocks via dimer-dimer recruitment. (2022). *SCIENCE ADVANCES*.

Ramm, B., Heermann, T., & Schwille, P. (2019). The E. coli MinCDE system in the regulation of protein patterns and gradients. *Cellular and Molecular Life Sciences*, *76*(21), 4245–4273. https://doi.org/10.1007/s00018-019-03218-x

Stylianidou, S., Brennan, C., Nissen, S. B., Kuwada, N. J., & Wiggins, P. A. (2016). SuperSegger: Robust image segmentation, analysis and lineage tracking of bacterial cells: Robust segmentation and analysis of bacteria. *Molecular Microbiology*, *102*(4), 690–700. https://doi.org/10.1111/mmi.13486

Sugawara, T., & Kaneko, K. (2011). Chemophoresis as a driving force for intracellular organization: Theory and application to plasmid partitioning. *Biophysics*, *7*, 77–88. https://doi.org/10.2142/biophysics.7.77

Surovtsev, I. V., Campos, M., & Jacobs-wagner, C. (2016). DNA-relay mechanism is sufficient to explain ParA-dependent intracellular transport and patterning of single and multiple cargos. *Proceedings of the National Academy of Sciences*, 1–9. https://doi.org/10.1073/pnas.1616118113

Szardenings, F., Guymer, D., & Gerdes, K. (2011). ParA ATPases can move and position DNA and subcellular structures. *Current Opinion in Microbiology*, *14*(6), 6. https://doi.org/10.1016/J.MIB.2011.09.008

Taylor, J. A., Seol, Y., Budhathoki, J., Neuman, K. C., & Mizuuchi, K. (2021). CTP and parS coordinate ParB partition complex dynamics and ParA-ATPase activation for ParABS-mediated DNA partitioning. *ELife*, *10*, e65651. https://doi.org/10.7554/eLife.65651

Vecchiarelli, A. G., Han, Y.-W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertümpfel, C., Funnell, B. E., & Mizuuchi, K. (2010). ATP control of dynamic P1 ParA-DNA interactions: A key role for the nucleoid in plasmid partition. *Molecular Microbiology*, *78*(1), 1. https://doi.org/10.1111/j.1365-2958.2010.07314.x

Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., & Jun, S. (2010). Robust Growth of Escherichia coli. *Current Biology*, *20*(12), 1099–1103. https://doi.org/10.1016/j.cub.2010.04.045

5.1.7. Supplementary Figures

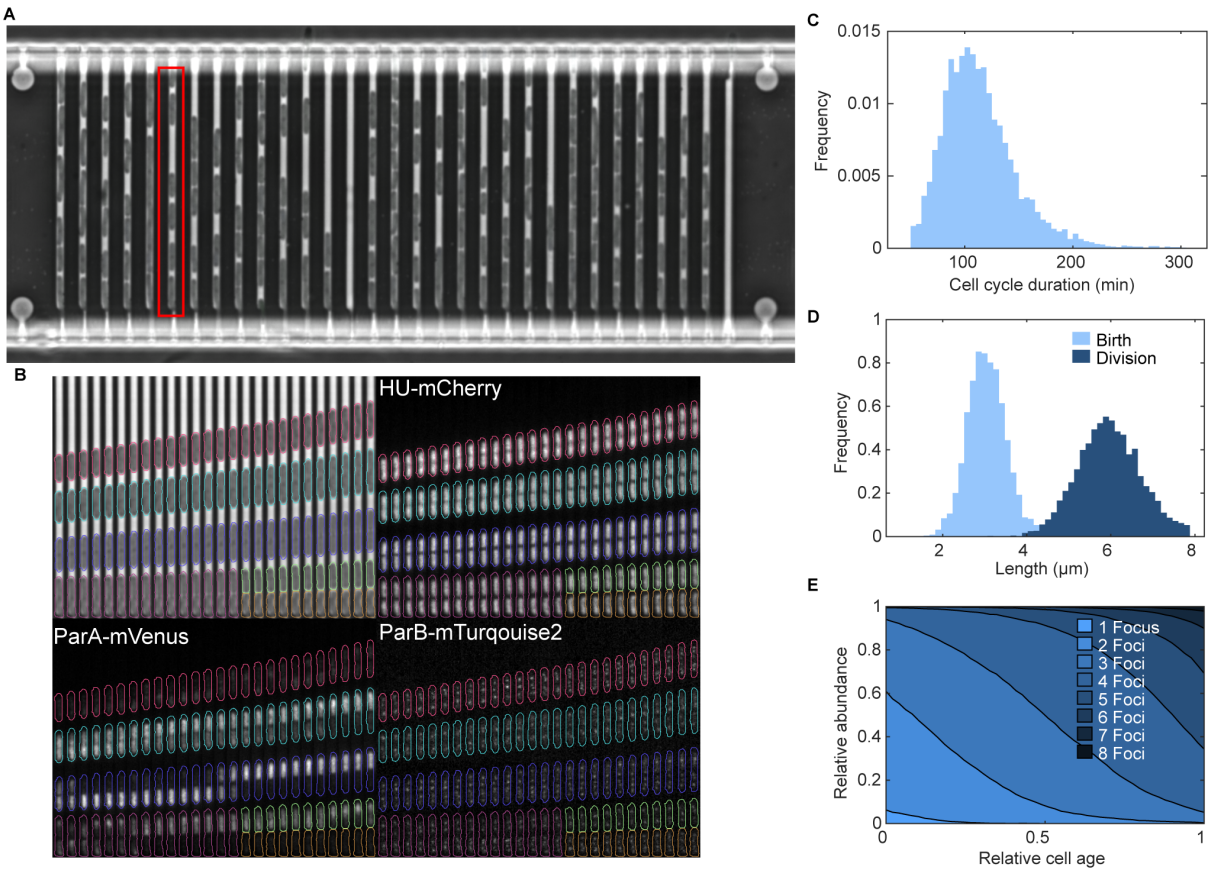

Figure S1. Overview, mother machine experiment. **(A)** Growth channels of one field of view. **(B)** Kymograph of one growth channel of the mother machine device (The one highlighted in red in panel **(A)**). Phase contrast and 3 fluorescent channels are depicted with overlaid cell contours (same color = same cell cycle). **(C)** Histogram of cell cycle duration of all acquired cell cycles (113.7 ± 34.8 min, mean ± SD). **(D)** Histogram of birth and division size of all acquired cell cycles. **(E)** Relative abundance of ParB foci plotted against relative cell age. The panels **(C)**, **(D)** and **(E)** were created from a dataset containing 5357 cell cycles.

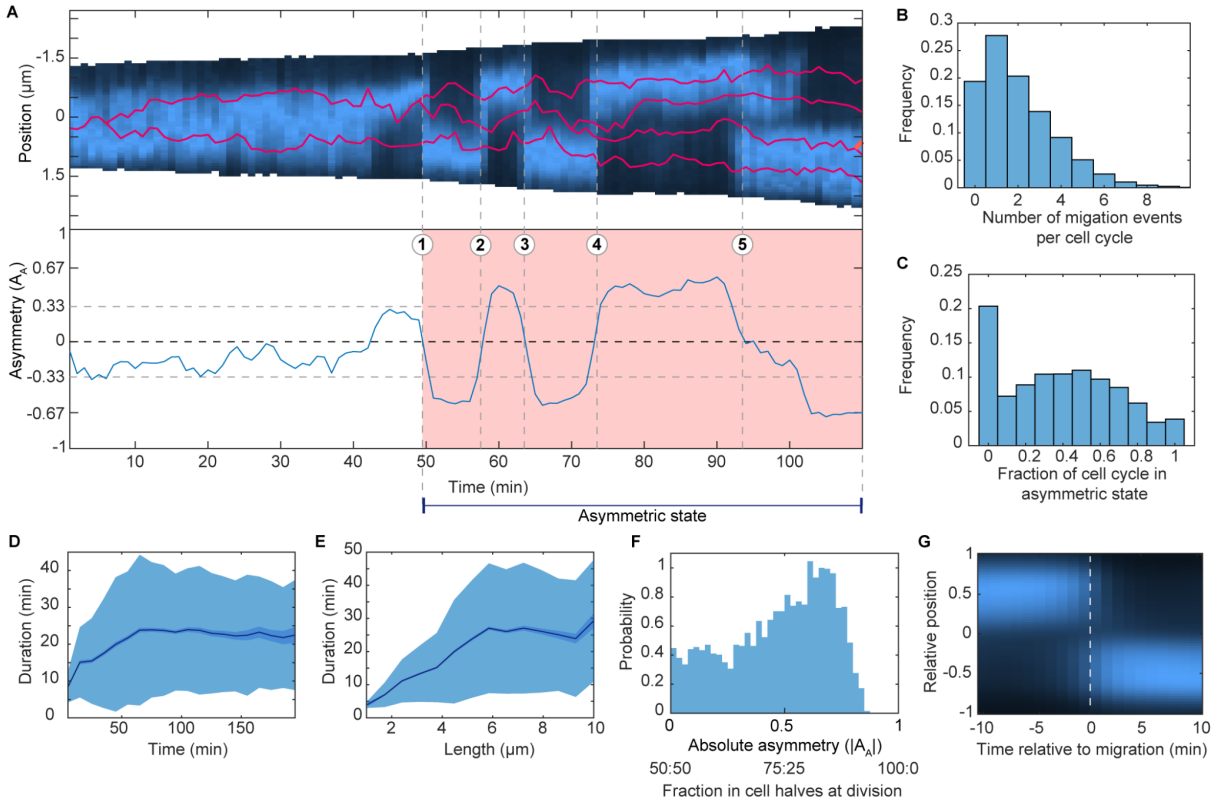

Figure S2. Properties of ParA migration patterns. **(A)** Example of the annotation of migration events. Top: Kymograph of ParA-mVenus signal. Red lines indicate plasmid trajectories obtained from tracking foci of ParB-mTurquoise2. Bottom: Asymmetry of ParA. When the asymmetry exceeds the threshold indicated by the gray dashed lines (2/3 of the signal localized in one cell half), the area between the zero-crossings is marked as ParA being in an asymmetric state (red shaded region). These zero-crossing points (1-5) are labeled as migration events of ParA. **(B)** Distribution of the number of migration events per cell cycle. **(C)** Fraction of the cell cycle in which ParA is in an asymmetric state. **(D)** The duration between migration events plotted against time since birth. **(E)** The duration between migration events plotted against cell length. **(F)** The absolute asymmetry of ParA on the last frame from each cell cycle or the fraction of ParA in both cell halves directly before division. **(G)** Average kymograph of the ParA signal relative to ParA migration events. The ParA signal was aligned such that the majority of ParA is localized in the cell half with a relative position greater than 0. The panels **(B)** to **(F)** were created from a dataset containing 5357 cell cycles.

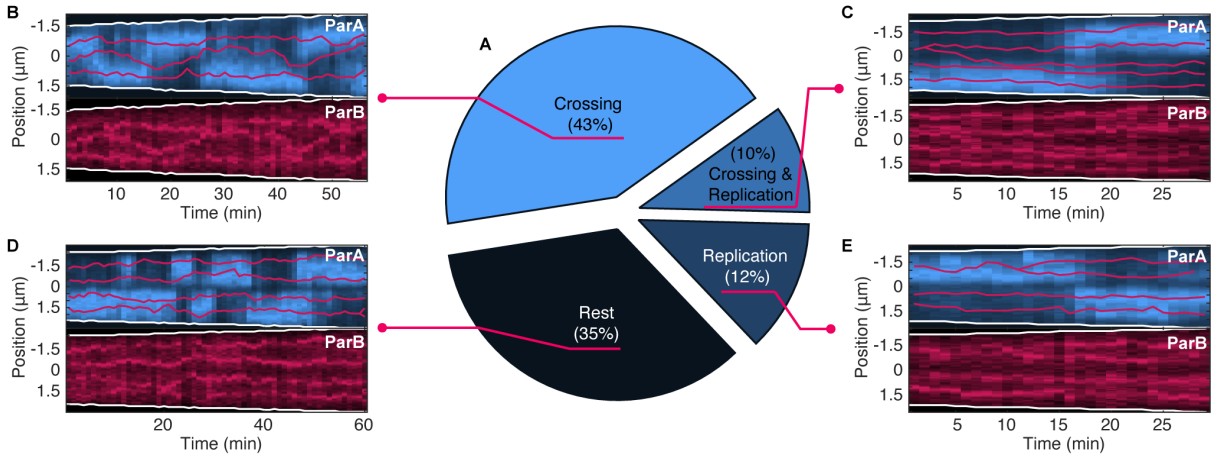

Figure S3. What precedes a migration event. **(A)** Pie chart of what cellular event precedes (10 min or less) a ParA migration event. Crossing: A plasmid crossing mid cell. Replication: A plasmid is replicated. Rest: Neither crossing nor replication. **(B)** A fragment of a cell cycle in which a crossing event precedes a migration event. **(C)** Same as **(B)** but for crossing and replication. **(D)** Same as **(B)** but for neither replication nor crossing. **(E)** Same as **(B)** but for replication. Panel **(A)** was created from a dataset containing 5357 cell cycles.

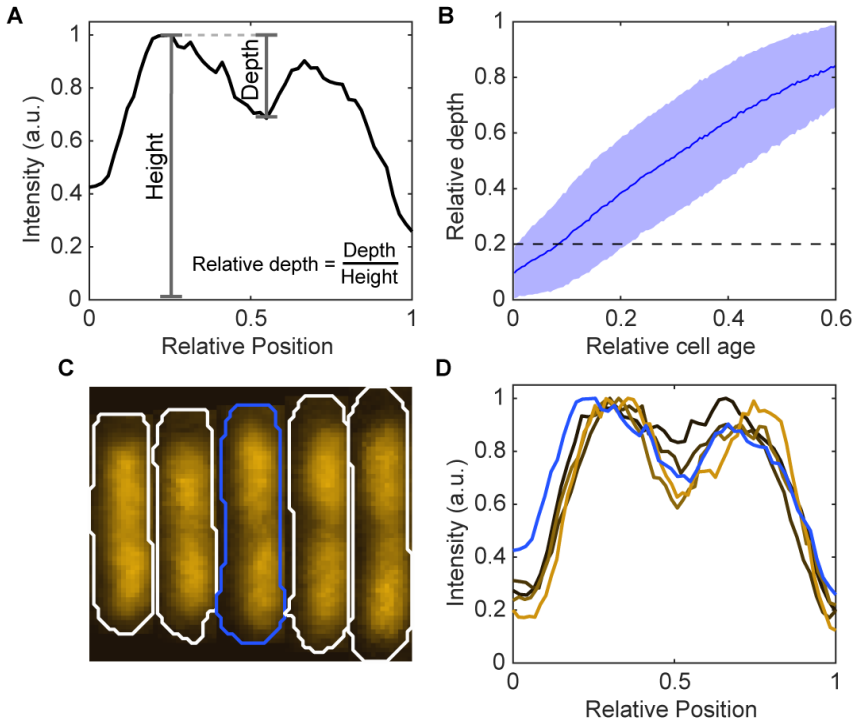

Figure S4. Nucleoid constriction detection. **(A)** Example of intensity profile of HU-mCherry from one cell. Highlighted are the height (max intensity) and the depth (distance between height and valley) and how to calculate the relative depth. **(B)** Average relative depth plotted against relative cell age; shade region is the standard deviation. The dashed line indicates the threshold depth for nucleoid constriction detection. This threshold is equal to the 95% percentile of all relative depths on the first frame of each cell cycle (n=5357). **(C)** Example of nucleoid constriction formation. The cell outlined in blue indicates the initial instance where a nucleoid constriction was detected, and this feature persists in all subsequent cells. **(D)** Line profiles of the 5 cells shown **(C)**. The blue line corresponds to the cell with the blue outline.

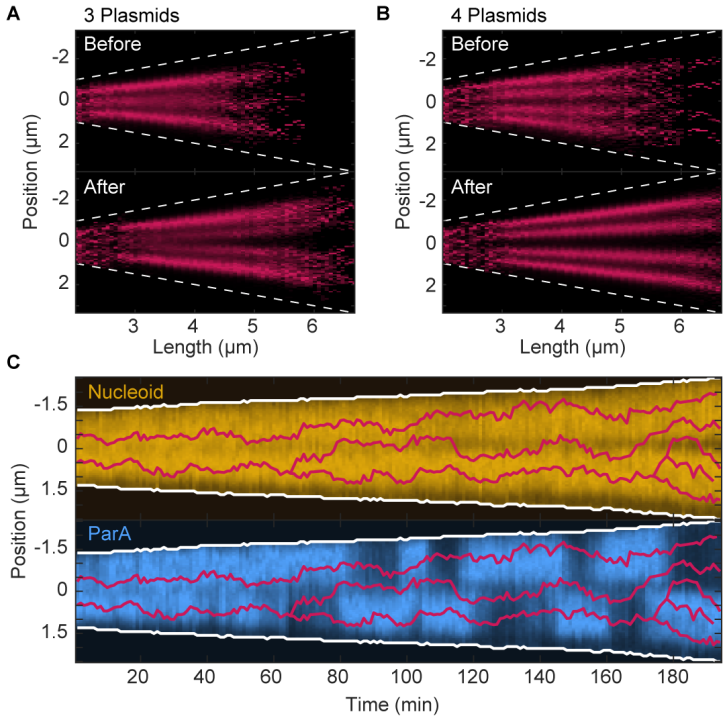

Figure S5. Divergent plasmid localization between cells with odd and even numbers of plasmids. **(A)** Top: Kymograph with the distribution of plasmids at different cell lengths containing three plasmids before the nucleoid is constricted. Bottom: Same as top but after the nucleoid is constricted. **(B)** Same as **(A)** but for cells with 4 plasmids. **(C)** Kymograph of the nucleoid (HU-mCherry) and ParA-mVenus of a cell cycle. Red lines indicate plasmid trajectories.

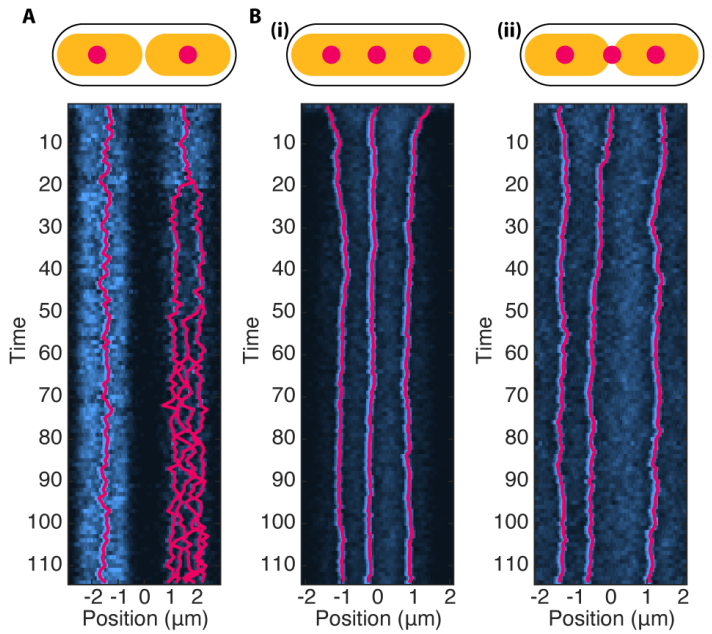

Figure S6. ParA-ParA recruitment is required for ParA migration events. **(A)** Top: Cartoon of the configuration of plasmids (red) and nucleoids (yellow). Bottom: Kymograph of a simulated cell with the same configuration as in Top. After four replication events, all but one plasmid are located over one nucleoid. **(B)** Cells with different nucleoid shapes do not produce migration events: **(i)** A simulated cell with a single nucleoid and three plasmids. **(ii)** Same as **(i)** but with two separated nucleoids.

6. Synopsis

The ParABS system is the most common partitioning system for low copy number plasmids and has been the subject of extensive research over the years. Despite these efforts, there are still many aspects of the system that remain poorly understood. In particular, the lack of quantitative data on a single cellular level is a great obstacle to furthering our comprehension of the ParABS system and validating current models. This underscores the need for more quantitative and rigorous investigations into the ParABS system to understand the mechanisms required for faithful plasmid segregation. In the first part of this section, the contributions of this thesis addressing this issue are summarized. Additionally, we will examine the impact of physical modeling and how modelling can serve as a guide for experimental research. Furthermore, we will delve into a few of the important questions that have been addressed, as well as those that persist regarding the ParABS system. Lastly, we will also address areas where the papers present in this thesis may benefit from additional research and refinement.

The primary contribution of this thesis is the identification of the dynamics that govern the movement of plasmids partitioned by the ParABS system. This was made possible by tagging the plasmid with mVenus-ParB and monitoring its localization during the cell cycle at one-minute intervals over thousands of cell cycles. A significant improvement and distinction from previous studies is our utilization of a microfluidic device in conjunction with our in-house developed image analysis pipeline, enabling the analysis of experiments conducted over multiple days. This way, we were able to observe signatures within the dynamics of plasmid movement, which revealed that the plasmid behaves like a particle attached to a spring, resulting in precise and directed positioning. For the F-plasmid, regular positioning was unequivocally confirmed, rather than oscillations, which were partially observed in a different system. Leveraging these experimental insights, we developed a stochastic model, based on the DNA-relay model, which was able to accurately capture our *in vivo* observations. This supports the mechanisms proposed by the model. Moreover, we demonstrated that regular positioning is not the only mode of positioning employed by the ParABS system, as the pB171-plasmid was observed to undergo pole-to-pole oscillations. Remarkably, our model is also capable of explaining this observation, further supporting its validity. We determined that the primary distinction between the two regimes is due to the combined effects of basal hydrolysis and ParA hopping. These processes establish a characteristic length scale, λ, which describes the distance ParA diffuses on the nucleoid before undergoing basal hydrolysis and being released into the cytosol. This length sets the plasmid's “sensing radius”, which determines its ability to detect nucleoid boundaries and other plasmids. If the specified length scale is shorter than half the cell's length, achieving regular positioning with a single plasmid is not feasible.

It is noteworthy that our simulations revealed the lowest energy consumption in the oscillatory regime, thus providing an advantageous incentive to operate as closely to it as possible without jeopardizing the stability of the plasmid. This might explain the position of both F- and pB171-plasmids within the phase space so close to the transition between oscillations and regular positioning. In the case of pB171, the plasmid-bearing strain carries a delta *ΔpcnB* mutation, which decreases the average number of plasmids per cell and therefore reduces the plasmid concentration, potentially pushing the system more towards the oscillatory regime. This is due to the fact that as the plasmid number and vice versa plasmid concentration decreases, ParA must travel greater distances to ensure regular positioning of the plasmids. An investigation that modifies the hydrolysis rate of ParA to alter the average distance it travels on the nucleoid could prove to be very insightful.

Using our ⋆Track algorithm, we were able to analyze plasmid movement in greater detail and address long-standing questions regarding the oscillations of ParA. By employing the same high-throughput procedures as before, we characterized ParA oscillations and revealed the interplay between ParA, ParB, and the nucleoid. We discovered that a constricted nucleoid and the movement of ParB-coated plasmids triggered ParA migration events between cell halves, resulting in ParA oscillations over time. Current models are unable to produce these oscillations. However, an extension of our model, which includes a nucleoid constriction and ParA-ParA recruitment, was able to accurately reproduce ParA oscillations. Simulations revealed their importance in preventing plasmid loss in cells in which the nucleoid constricts early in the cell cycle. Overall, the study provides a solid explanation for the dynamics of ParA oscillations and contributes to a deeper understanding of bacterial plasmid segregation and chromosomal partitioning.

All in all, the combination of microfluidic, high-throughput experiments and physical modelling has proven to be an exceptionally powerful approach for investigating complex biological systems, such as the ParABS system. Utilizing a microfluidic device, such as the mother machine, allows for a highly controlled environment in which cells grow in steady-state conditions for extended periods of time. It enables the extraction of thousands of cell cycles, allowing for in-depth investigation of scientific questions, even within small subpopulations, often yielding valuable insights, like the first paper's observation of large cells with only one plasmid displaying low amplitude oscillations. This high-throughput data also serves as an ideal resource for evaluating models that aim to describe the underlying mechanisms of a system, as it offers the means to directly compare the output of the model and *in vivo* observations. By examining the agreement between the model and the observations, we can assess the model's predictive accuracy and deduce the crucial mechanisms governing the system. Furthermore, if the model fails to reproduce the observations, it indicates that a fundamental aspect of the system is missing. This was the case for the ParA oscillations, which we were able to reproduce only after adding new components to the model. Ultimately, many of the significant findings in this

thesis were made possible due to the synergistic combination of rigorous and quantitative observations with physical modelling.

The second publication about ⋆Track did not focus on the ParABS system and instead on the development of a method enabling precise tracking of persistent and replicating particles, such as DNA loci. It performs better than established algorithms in this specific niche and proves to be immensely beneficial when examining DNA replication systems in both chromosomes and plasmids, thanks to its high accuracy in predicting replication events and inferring the number of loci in each frame, as demonstrated in the publication. Consequently, it allows for the analysis of data with great amounts of noise even when tracking multiple objects, a task unattainable with naive tracking solutions. The third paper immensely benefited from this algorithm as all trajectories shown were produced using ⋆Track.

⋆Track, MotherSegger and our ParABS model ("hopping and relay") are all open source and publicly available.

6.1. Open Questions

Numerous questions remain unanswered regarding the ParABS system. One aspect is the impact of ParB-induced hydrolysis of ParA. Researchers showed how ParA oscillations are significantly faster in cells carrying a mutated ParB with lower efficiency to stimulate the ATPase activity of ParA (Ah-Seng et al., 2013). This behaviour is counterintuitive and does not agree with the predictions of our model. In our model, if we reduce the rate of ParB-induced hydrolysis, the oscillations become slower, directly opposite to the *in vivo* observations. This implies that there are still many unknowns that remain to be uncovered.

Similar to ParA, ParB presents a number of challenging questions that have yet to be fully answered, leaving a significant gap in our knowledge regarding the mechanisms underlying the formation of the partition complex. While ParB sliding offered a new perspective on how a large amount of ParB can be loaded onto DNA with only a few *parS* sites, it remains unclear how ParB forms the partition complex. One possible hypothesis for its formation is through ParB-ParB bridges, where two DNA-bound ParB dimers form a bridge, bringing together distant DNA strands. The lifetime and frequency of these bridges, as observed in simulations, play a crucial role in the formation and condensation of the partition complex into a spatially confined blob, as seen *in vivo*.

6.2. Limitations

Despite the significant findings of the publications presented in this thesis, it is important to acknowledge some limitations and weaknesses of these studies.

The first publication primarily concentrated on cells with 1 or 2 plasmids, without delving into a comprehensive analysis of cells containing a higher number of plasmids. This is due to the difficulty in accurately identifying the number and location of foci once multiple plasmids (>2) are present. Due to this limitation, our study focused only on the analysis of a subpopulation of cells, which may have caused us to overlook phenomena that occur only at high plasmid numbers, which is unlikely but not impossible. This will not be an obstacle for future studies, as we have developed ★Track, a tool that facilitates precise tracking of cells with multiple foci and enables inference of missing or merged foci. Another weakness of this publication is related to the pB171-plasmid, which employed the FROS system instead of a ParB fusion, as done in the F-plasmid, to label the plasmids. In the FROS system, TetR-mCherry binds to a TetO-array consisting of 120 repeats, which is located on the plasmid. Excessive TetR binding to the TetO-array creates an impassable obstacle for the replication machinery, ultimately leading to plasmid loss. Hence, it is possible that the observed behaviour of the pB171-plasmid does not accurately reflect its native behaviour, and instead may have been altered by the binding of TetR-mCherry. For example, the average number of plasmids per cell might have decreased before imaging due to the FROS system.

The second publication introduces the ★Track algorithm, which has certain limitations. Its primary constraint is the number of objects it can effectively track. As the density of particles increases, the A* step becomes progressively more computationally demanding, potentially rendering ★Track unusable when applied outside its intended scope. Another limitation is its specificity, as it was designed specifically for tracking persistent and replicating objects. While tracking transient objects is possible by omitting the algorithm's later steps, this approach sacrifices ★Track's ability to predict splitting events. Overall, the algorithm caters to a specific niche in which it excels but applying it to data outside this scope may result in some drawbacks.

The third paper does present certain limitations. The most significant limitation is that the proposed ParA-ParA recruitment mechanism remains unproven. Although it explains the observed *in vivo* observations, is consistent with filament formation *in vitro*, and provides a notable advantage in preventing plasmid loss in specific situations, it still remains a hypothesis. Furthermore, we have not been able to reproduce one particular *in vivo* observation regarding ParB-hydrolysis mutants, which exhibit less stimulation of ParA's ATPase activity than the wild type ParB. These mutants result in faster ParA oscillations, contrary to our simulation findings. This discrepancy suggests that there may be an additional, as-yet-unidentified mechanism at play.

6.3. Future work

Our next step in investigating the ParABS system is to test our hypothesis about ParA-ParA recruitment. A potential strategy to confirm its existence is by identifying the domain responsible for

this process. This can be achieved by examining both existing and novel ParA mutations to determine if any eliminate ParA oscillations. Informed predictions can be made about which domain within the approximately 400-amino-acid protein is crucial for recruitment. A recent paper suggests that ParA binds to the promoter region of the par operon cooperatively (Boudsocq et al., 2021), proposing that a winged-helix-domain in the center of ParA is necessary for such cooperativity. This serves as a promising starting point for further investigation. If this approach proves unsuccessful, we plan to construct a library of ParA mutants and systematically test them until the desired phenotype, if it exists, is found.

A flow cytometer could be used to swiftly screen mutants by observing the ParA distribution within the population of a specific mutant. If oscillations cease, ParA is inherited more evenly, leading to a narrower, less dispersed distribution of ParA within the population. This should be detectable using flow cytometry, enabling the rapid screening of various mutations. However, a limitation of this method is that other mutations, besides disrupting recruitment, can also halt ParA oscillations. For instance, the interaction between ParA and ParB is essential, as is the ability of ParA to bind DNA non-specifically. Distinguishing between these different non-oscillating ParA phenotypes may prove challenging. To address this issue, examining plasmid stability could be a valuable approach. We hypothesize that the non-recruiting ParA phenotype will result in a minor increase in the loss rate of the plasmid, while other mutations will have a more pronounced effect.

Identifying the precise domain implicated in ParA-ParA recruitment could have broader implications beyond merely clarifying the mechanism of ParA oscillations. These findings could also benefit the study of chromosomal ParABS systems and their role in DNA segregation. In the end, addressing this question would constitute a substantial contribution to the scientific community, enhancing our knowledge of bacterial genetics and the fundamental principles governing spatial organization within cells.

6.4. Final Remarks

This thesis offers a comprehensive examination of the plasmid partitioning system known as ParABS. It begins with a historical overview of the system, emphasizing key insights gained over decades of outstanding scientific research. From its discovery in the 1980s to the present day, many researchers have contributed invaluable work investigating the system, which serves as the foundation of this thesis. Following the introduction, three papers are presented, two directly about the ParABS system and another focusing on an algorithm developed to enable a more detailed analysis of the system. We demonstrated that different ParABS systems exhibit various behaviours while sharing common underlying mechanisms, as evidenced by cutting-edge microfluidic experiments, image analysis, and physical modelling. We identified a crucial length scale, the average distance ParA diffuses on the

nucleoid, which controls the behaviour of a ParABS system, in our simulations. Additionally, we proposed a hypothesis explaining why the system operates close to an oscillating instability – to minimize ATP consumption. The last paper presents further data offering a novel perspective on the system. We showed that the behaviour of the system varies depending on whether the nucleoid is constricted or fully replicated & separated, and switches from a positioning system on the nucleoid to a partitioning system between nucleoids. All in all, this thesis provides valuable insights into the ParABS system and hopefully contributes to future studies in the ongoing effort to gain a deeper understanding of our world.

7. Acknowledgements

Shout out to my supervisor Seán M. Murray for his mentorship and enduring my imprecise nature during the last 3 years! His unwavering support has enabled me to enhance my scientific skills and fully develop my potential over the course of the last three years.

I would also like to express my gratitude to the Max Planck Institute for Terrestrial Microbiology for offering the opportunity and essential infrastructure required to carry out my research. Their invaluable support has been instrumental in the progress and success of my work.

Lastly, I would like to extend my heartfelt appreciation to all members of my TAC-committee and my coworkers for their engaging discussions, as well as the scientific guidance and emotional support they have generously provided throughout my PhD, except Giovanni :P

8. References

Abeles, A. L., Friedman, S. A., & Austin, S. J. (1985). Partition of Unit-copy Miniplasmids to Daughter Cells. *Journal of Molecular Biology*, *185*, 261–272.

Ah-Seng, Y., Rech, J., Lane, D., & Bouet, J.-Y. (2013). Defining the Role of ATP Hydrolysis in Mitotic Segregation of Bacterial Plasmids. *PLoS Genetics*, *9*(12), 12. https://doi.org/10.1371/journal.pgen.1003956

Austin, S., Hart, F., Abeles, A., & Sternberg, N. (1982). Genetic and physical map of a P1 miniplasmid. *Journal of Bacteriology*, *152*(1), 63–71. https://doi.org/10.1128/jb.152.1.63-71.1982

Baltekin, Ö., Boucharin, A., Tano, E., Andersson, D. I., & Elf, J. (2017). Antibiotic susceptibility testing in less than 30 min using direct single-cell imaging. *Proceedings of the National Academy of Sciences*, *114*(34), 9170–9175. https://doi.org/10.1073/pnas.1708558114

Boudsocq, F., Salhi, M., Barbe, S., & Bouet, J.-Y. (2021). Three ParA Dimers Cooperatively Assemble on Type Ia Partition Promoters. *Genes*, *12*(9), 1345. https://doi.org/10.3390/genes12091345

Branka, A. C., & Heyes, D. M. (n.d.). *Algorithms for Brownian dynamics simulation*.

Breier, A. M., & Grossman, A. D. (2007). Whole-genome analysis of the chromosome partitioning and sporulation protein Spo0J (ParB) reveals spreading and origin-distal sites on the Bacillus subtilis chromosome. *Molecular Microbiology*, *64*(3), 703–718. https://doi.org/10.1111/j.1365-2958.2007.05690.x

Campbell, C. S., & Mullins, R. D. (2007). In vivo visualization of type II plasmid segregation: Bacterial actin filaments pushing plasmids. *The Journal of Cell Biology*, *179*(5), 5. https://doi.org/10.1083/jcb.200708206

Castaing, J.-P., Bouet, J.-Y., & Lane, D. (2008). F plasmid partition depends on interaction of SopA with non-specific DNA. *Molecular Microbiology*. https://doi.org/10.1111/j.1365-2958.2008.06465.x

Chenouard, N., Smal, I., de Chaumont, F., Maška, M., Sbalzarini, I. F., Gong, Y., Cardinale, J., Carthel, C., Coraluppi, S., Winter, M., Cohen, A. R., Godinez, W. J., Rohr, K., Kalaidzidis, Y., Liang, L., Duncan, J., Shen, H., Xu, Y., Magnusson, K. E. G., … Meijering, E. (2014). Objective comparison of particle tracking methods. *Nature Methods*, *11*(3), 281–289. https://doi.org/10.1038/nmeth.2808

Davis, M. A., & Austin, S. J. (1988). Recognition of the P1 plasmid centromere analog involves binding of the ParB protein and is modified by a specific host factor. *The EMBO Journal*, *7*(6), 1881–1888. https://doi.org/10.1002/j.1460-2075.1988.tb03021.x

Davis, M. A., Martin, K. A., & Austin, S. J. (1992). Biochemical activities of the ParA partition protein of the P1 plasmid. *Molecular Microbiology*, *6*(9), 1141–1147. https://doi.org/10.1111/j.1365-2958.1992.tb01552.x

de Boer, P. A. J., Crossley, R. E., & Rothfield, L. I. (1989). A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in E. coli. *Cell*, *56*(4), 641–649. https://doi.org/10.1016/0092-8674(89)90586-2

Dunham, T. D., Xu, W., Funnell, B. E., & Schumacher, M. A. (2009). Structural basis for ADP-mediated transcriptional regulation by P1 and P7 ParA. *The EMBO Journal*, *28*(12), 1792–1802. https://doi.org/10.1038/emboj.2009.120

Ebersbach, G., & Gerdes, K. (2001). The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(26), 26. https://doi.org/10.1073/pnas.261569598

Ebersbach, G., & Gerdes, K. (2004). Bacterial mitosis: Partitioning protein ParA oscillates in spiral-shaped structures and positions plasmids at mid-cell. *Molecular Microbiology*, *52*(2), 2. https://doi.org/10.1111/j.1365-2958.2004.04002.x

Ebersbach, G., Ringgaard, S., Moller-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. *Molecular Microbiology*, *61*(6), 6. https://doi.org/10.1111/j.1365-2958.2006.05322.x

Erdmann, N., Petroff, T., & Funnell, B. E. (1999). Intracellular localization of P1 ParB protein depends on ParA and *parS*. *Proceedings of the National Academy of Sciences*, *96*(26), 14905–14910. https://doi.org/10.1073/pnas.96.26.14905

Funnell, B. E. (1988). Participation of Escherichia coli integration host factor in the P1 plasmid partition system. *Proceedings of the National Academy of Sciences*, *85*(18), 6657–6661. https://doi.org/10.1073/pnas.85.18.6657

Funnell, B. E. (1991). The P1 plasmid partition complex at parS. The influence of Escherichia coli integration host factor and of substrate topology. *Journal of Biological Chemistry*, *266*(22), 14328–14337. https://doi.org/10.1016/S0021-9258(18)98688-6

Gerdes, K., Moller-Jensen, J., & Jensen, R. B. (2000). Plasmid and chromosome partitioning: Surprises from phylogeny. *Molecular Microbiology*, *37*(3), 3. https://doi.org/10.1046/j.1365-2958.2000.01975.x

Gordon, G. S., Sitnikov, D., Webb, C. D., Teleman, A., Straight, A., Losick, R., Murray, A. W., & Wright, A. (1997). Chromosome and low copy plasmid segregation in E-coli: Visual evidence for distinct mechanisms. *Cell*, *90*(6), 6.

Govers, S., Campos, M., Tyagi, B., Laloux, G., & Jacobs-Wagner, C. (2023). Apparent simplicity and emergent robustness in bacterial cell cycle control. *BioRxiv*. https://doi.org/10.1101/2023.01.16.524295

Guilhas, B., Walter, J.-C., Rech, J., David, G., Walliser, N. O., Palmeri, J., Mathieu-Demaziere, C., Parmeggiani, A., Bouet, J.-Y., Le Gall, A., & Nollmann, M. (2020). ATP-Driven Separation of Liquid Phase Condensates in Bacteria. *Molecular Cell*, *79*(2), 2. https://doi.org/10.1016/j.molcel.2020.06.034

Halatek, J., & Frey, E. (2012). Highly Canalized MinD Transfer and MinE Sequestration Explain the Origin of Robust MinCDE-Protein Dynamics. *Cell Reports*, *1*(6), 741–752. https://doi.org/10.1016/j.celrep.2012.04.005

Hart, P. E., Nilsson, N. J., & Raphael, B. (1968). A Formal Basis for the Heuristic Determination. *IEEE TRANSACTIONS ON SYSTEMS SCIENCE AND CYBERNETICS*, 8.

Hatano, T., Yamaichi, Y., & Niki, H. (2007). Oscillating focus of SopA associated with filamentous structure guides partitioning of F plasmid. *Molecular Microbiology*, *64*(5), 5. https://doi.org/10.1111/j.1365-2958.2007.05728.x

Hayes, F., Radnedge, L., Davis, M. A., & Austin, S. J. (1994). The homologous operons for P1 and P7 plasmid partition are autoregulated from dissimilar operator sites. *Molecular Microbiology*, *11*(2), 249–260. https://doi.org/10.1111/j.1365-2958.1994.tb00305.x

Hu, L., Vecchiarelli, A. G., Mizuuchi, K., Neuman, K. C., & Liu, J. (2015). Directed and persistent movement arises from mechanochemistry of the ParA/ParB system. *Proceedings of the National Academy of Sciences*, *112*(51), 51. https://doi.org/10.1073/pnas.1505147112

Hu, L., Vecchiarelli, A. G., Mizuuchi, K., Neuman, K. C., & Liu, J. (2017a). Brownian Ratchet Mechanism for Faithful Segregation of Low-Copy-Number Plasmids. *Biophysical Journal*, *112*(7), 7. https://doi.org/10.1016/j.bpj.2017.02.039

Hu, L., Vecchiarelli, A. G., Mizuuchi, K., Neuman, K. C., & Liu, J. (2017b). Brownian ratchet mechanisms of ParA-mediated partitioning. *Plasmid*, *92*, 12–16. https://doi.org/10.1016/j.plasmid.2017.05.002

Huang, K. C., Meir, Y., & Wingreen, N. S. (2003). Dynamic structures in *Escherichia coli*: Spontaneous formation of MinE rings and MinD polar zones. *Proceedings of the National Academy of Sciences*, *100*(22), 12724–12728. https://doi.org/10.1073/pnas.2135445100

Hwang, L. C., Vecchiarelli, A. G., Han, Y.-W., Mizuuchi, M., Harada, Y., Funnell, B. E., & Mizuuchi, K. (2013). ParA-mediated plasmid partition driven by protein pattern self-organization. *The EMBO Journal*, *32*(9), 1238–1249. https://doi.org/10.1038/emboj.2013.34

Ietswaart, R., Szardenings, F., Gerdes, K., & Howard, M. (2014). Competing ParA structures space bacterial plasmids equally over the nucleoid. *PLoS Computational Biology*, *10*(12), 12. https://doi.org/10.1371/journal.pcbi.1004009

Jones, C. W., & Armitage, J. P. (2017). Essential Role of the Cytoplasmic Chemoreceptor TlpT in the *De Novo* Formation of Chemosensory Complexes in Rhodobacter sphaeroides. *Journal of Bacteriology*, *199*(19). https://doi.org/10.1128/JB.00366-17

Jun, S., & Mulder, B. (2006). Entropy-driven spatial organization of highly confined polymers: Lessons for the bacterial chromosome. *Proceedings of the National Academy of Sciences*, *103*(33), 12388–12393. https://doi.org/10.1073/pnas.0605305103

Köhler, R., Kaganovitch, E., & Murray, S. M. (2022). High-throughput imaging and quantitative analysis uncovers the nature of plasmid positioning by ParABS. *ELife*, *11*, e78743. https://doi.org/10.7554/eLife.78743

Le Gall, A., Cattoni, D. I., Guilhas, B., Mathieu-Demazière, C., Oudjedi, L., Fiche, J.-B., Rech, J., Abrahamsson, S., Murray, H., Bouet, J.-Y., & Nollmann, M. (2016). Bacterial partition complexes segregate within the volume of the nucleoid. *Nature Communications*, *7*, 12107. https://doi.org/10.1038/ncomms12107

Leonard, T. A., Butler, P. J., & Löwe, J. (2005). Bacterial chromosome segregation: Structure and DNA binding of the Soj dimer ? a conserved biological switch. *The EMBO Journal*, *24*(2), 270–282. https://doi.org/10.1038/sj.emboj.7600530

Lim, G. E., Derman, A. I., & Pogliano, J. (2005). Bacterial DNA segregation by dynamic SopA polymers. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(49), 49. https://doi.org/10.1073/pnas.0507222102

Lim, H. C., Surovtsev, I. V., Beltran, B. G., Huang, F., Bewersdorf, J., & Jacobs-Wagner, C. (2014). Evidence for a DNA-relay mechanism in ParABS-mediated chromosome segregation. *ELife*, *3*(3), 3. https://doi.org/10.7554/eLife.02758

Lin-Chao, S., Chen, W.-T., & Wong, T.-T. (1992). High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. *Molecular Microbiology*, *6*(22), 3385–3393. https://doi.org/10.1111/j.1365-2958.1992.tb02206.x

Lindow, J. C., Britton, R. A., & Grossman, A. D. (2002). Structural Maintenance of Chromosomes Protein of *Bacillus subtilis* Affects Supercoiling In Vivo. *Journal of Bacteriology*, *184*(19), 5317–5322. https://doi.org/10.1128/JB.184.19.5317-5322.2002

Livny, J., Yamaichi, Y., & Waldor, M. K. (2007). Distribution of centromere-like parS sites in bacteria: Insights from comparative genomics. *Journal of Bacteriology*, *189*(23), 23. https://doi.org/10.1128/JB.01239-07

MacCready, J. S., Hakim, P., Young, E. J., Hu, L., Liu, J., Osteryoung, K. W., Vecchiarelli, A. G., & Ducat, D. C. (2018). Protein gradients on the nucleoid position the carbon-fixing organelles of cyanobacteria. *ELife*, *7*, e39723. https://doi.org/10.7554/eLife.39723

MacCready, J. S., & Vecchiarelli, A. G. (2021). Positioning the Model Bacterial Organelle, the Carboxysome. *MBio*, *12*(3), e02519-19. https://doi.org/10.1128/mBio.02519-19

Manzo, C., & Garcia-Parajo, M. F. (2015). A review of progress in single particle tracking: From methods to biophysical insights. *Reports on Progress in Physics*, *78*(12), 124601. https://doi.org/10.1088/0034-4885/78/12/124601

Martin, K. A., Friedman, S. A., & Austin, S. J. (1987). Partition site of the P1 plasmid. *Proceedings of the National Academy of Sciences*, *84*(23), 8544–8547. https://doi.org/10.1073/pnas.84.23.8544

Meijering, E., Dzyubachyk, O., & Smal, I. (2012). Methods for Cell and Particle Tracking. In *Methods in Enzymology* (Vol. 504, pp. 183–200). Elsevier. https://doi.org/10.1016/B978-0-12-391857-4.00009-4

Mishra, D., Jakhmola, A., & Srinivasan, R. (2022). A role for the last C-terminal helix of the F plasmid segregating protein SopA in nucleoid binding and plasmid maintenance. *Plasmid*, *119–120*, 102617. https://doi.org/10.1016/j.plasmid.2022.102617

Mohl, D. A., Easter, J., & Gober, J. W. (2008). The chromosome partitioning protein, ParB, is required for cytokinesis in Caulobacter crescentus: ParB is required for cell division in Caulobacter. *Molecular Microbiology*, *42*(3), 741–755. https://doi.org/10.1046/j.1365-2958.2001.02643.x

Mori, H., Kondo, A., Ohshima, A., Ogura, T., & Hiraga, S. (1986). Structure and function of the F plasmid genes essential for partitioning. *Journal of Molecular Biology*, *192*(1), 1–15. https://doi.org/10.1016/0022-2836(86)90459-6

Motallebi-Veshareh, M., Rouch, D. A., & Thomas, C. M. (1990). A family of ATPases involved in active partitioning of diverse bacterial plasmids. *Molecular Microbiology*, *4*(9), 1455–1463. https://doi.org/10.1111/j.1365-2958.1990.tb02056.x

Ogura, T., & Hiraga, S. (1983). Partition mechanism of F plasmid: Two plasmid gene-encoded products and a cis-acting region are involved in partition. *Cell*, *32*(2), 2. https://doi.org/10.1016/0092-8674(83)90454-3

Osorio-Valeriano, M., Altegoer, F., Steinchen, W., Urban, S., Liu, Y., Bange, G., & Thanbichler, M. (2019). ParB-type DNA Segregation Proteins Are CTP-Dependent Molecular Switches. *Cell*, *179*(7), 1512-1524.e15. https://doi.org/10.1016/j.cell.2019.11.015

Ramm, B., Heermann, T., & Schwille, P. (2019). The E. coli MinCDE system in the regulation of protein patterns and gradients. *Cellular and Molecular Life Sciences*, *76*(21), 4245–4273. https://doi.org/10.1007/s00018-019-03218-x

Roberts, M. A. J., Wadhams, G. H., Hadfield, K. A., Tickner, S., & Armitage, J. P. (2012). ParA-like protein uses nonspecific chromosomal DNA binding to partition protein complexes. *Proceedings of the National Academy of Sciences*, *109*(17), 6698–6703. https://doi.org/10.1073/pnas.1114000109

Shen, H., Tauzin, L. J., Baiyasi, R., Wang, W., Moringo, N., Shuang, B., & Landes, C. F. (2017). Single Particle Tracking: From Theory to Biophysical Applications. *Chemical Reviews*, *117*(11), 7331–7376. https://doi.org/10.1021/acs.chemrev.6b00815

Soh, Y.-M., Davidson, I. F., Zamuner, S., Basquin, J., Bock, F. P., Taschner, M., Veening, J.-W., De Los Rios, P., Peters, J.-M., & Gruber, S. (2019). Self-organization of *parS* centromeres by the ParB CTP hydrolase. *Science*, *366*(6469), 1129–1133. https://doi.org/10.1126/science.aay3965

Stylianidou, S., Brennan, C., Nissen, S. B., Kuwada, N. J., & Wiggins, P. A. (2016). SuperSegger: Robust image segmentation, analysis and lineage tracking of bacterial cells: Robust segmentation and analysis of bacteria. *Molecular Microbiology*, *102*(4), 690–700. https://doi.org/10.1111/mmi.13486

Sugawara, T., & Kaneko, K. (2011). Chemophoresis as a driving force for intracellular organization: Theory and application to plasmid partitioning. *Biophysics*, *7*, 77–88. https://doi.org/10.2142/biophysics.7.77

Surovtsev, I. V., Campos, M., & Jacobs-wagner, C. (2016). DNA-relay mechanism is sufficient to explain ParA-dependent intracellular transport and patterning of single and multiple cargos. *Proceedings of the*

National Academy of Sciences, 1–9. https://doi.org/10.1073/pnas.1616118113

Szardenings, F., Guymer, D., & Gerdes, K. (2011). ParA ATPases can move and position DNA and subcellular structures. *Current Opinion in Microbiology*, *14*(6), 6. https://doi.org/10.1016/J.MIB.2011.09.008

Taylor, J. A., Seol, Y., Budhathoki, J., Neuman, K. C., & Mizuuchi, K. (2021). CTP and parS coordinate ParB partition complex dynamics and ParA-ATPase activation for ParABS-mediated DNA partitioning. *ELife*, *10*, e65651. https://doi.org/10.7554/eLife.65651

Thiem, S., & Sourjik, V. (2008). Stochastic assembly of chemoreceptor clusters in Escherichia coli. *Molecular Microbiology*, *68*(5), 1228–1236. https://doi.org/10.1111/j.1365-2958.2008.06227.x

Thompson, S. R., Wadhams, G. H., & Armitage, J. P. (2006). The positioning of cytoplasmic protein clusters in bacteria. *Proceedings of the National Academy of Sciences*, *103*(21), 8209–8214. https://doi.org/10.1073/pnas.0600919103

Tišma, M., Panoukidou, M., Antar, H., & Dekker, C. (2022). ParB proteins can bypass DNA-bound roadblocks via dimer-dimer recruitment. *SCIENCE ADVANCES*.

Tran, N. T., Laub, M. T., & Le, T. B. K. (2017). SMC Progressively Aligns Chromosomal Arms in Caulobacter crescentus but Is Antagonized by Convergent Transcription. *Cell Reports*, *20*(9), 2057–2071. https://doi.org/10.1016/j.celrep.2017.08.026

Turmo, A., Gonzalez-Esquer, C. R., & Kerfeld, C. A. (2017). Carboxysomes: Metabolic modules for CO2 fixation. *FEMS Microbiology Letters*, *364*(18). https://doi.org/10.1093/femsle/fnx176

Vecchiarelli, A. G., Han, Y.-W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertümpfel, C., Funnell, B. E., & Mizuuchi, K. (2010). ATP control of dynamic P1 ParA-DNA interactions: A key role for the nucleoid in plasmid partition. *Molecular Microbiology*, *78*(1), 1. https://doi.org/10.1111/j.1365-2958.2010.07314.x

Vecchiarelli, A. G., Neuman, K. C., & Mizuuchi, K. (2014). A propagating ATPase gradient drives transport of surface-confined cellular cargo. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(13), 13. https://doi.org/10.1073/pnas.1401025111

Viollier, P. H., Thanbichler, M., McGrath, P. T., West, L., Meewan, M., McAdams, H. H., & Shapiro, L. (2004). Rapid and sequential movement of individual chromosomal loci to specific subcellular locations during bacterial DNA replication. *Proceedings of the National Academy of Sciences*, *101*(25), 9257–9262. https://doi.org/10.1073/pnas.0402606101

Wang, P., Robert, L., Pelletier, J., Dang, W. L., Taddei, F., Wright, A., & Jun, S. (2010). Robust Growth of Escherichia coli. *Current Biology*, *20*(12), 1099–1103. https://doi.org/10.1016/j.cub.2010.04.045

Wang, X., Brandão, H. B., Le, T. B. K., Laub, M. T., & Rudner, D. Z. (2017). *Bacillus subtilis* SMC complexes juxtapose chromosome arms as they travel from origin to terminus. *Science*, *355*(6324), 524–527. https://doi.org/10.1126/science.aai8982

Wang, X., Hughes, A. C., Brandão, H. B., Walker, B., Lierz, C., Cochran, J. C., Oakley, M. G., Kruse, A. C., & Rudner, D. Z. (2018). In Vivo Evidence for ATPase-Dependent DNA Translocation by the Bacillus subtilis SMC Condensin Complex. *Molecular Cell*, *71*(5), 841-847.e5. https://doi.org/10.1016/j.molcel.2018.07.006

Watanabe, E., Wachi, M., Yamasaki, M., & Nagai, K. (1992). ATPase activity of SopA, a protein essential for active partitioning of F plasmid. *Molecular and General Genetics MGG*, *234*(3), 346–352. https://doi.org/10.1007/BF00538693

Yang, W., & Briegel, A. (2020). Diversity of Bacterial Chemosensory Arrays. *Trends in Microbiology*, *28*(1), 68–80. https://doi.org/10.1016/j.tim.2019.08.002