UPPSALA
UNIVERSITET

UPTEC X 18 017
Examensarbete 30 hp
Juni 2018

Mapping of Chromosome Dynamics over the Bacterial Cell Cycle

Konrad Gras

UPPSALA
UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
http://www.teknat.uu.se/student

Abstract

Mapping of Chromosome Dynamics over the Bacterial Cell Cycle

Konrad Gras

The replication of DNA, its compaction and segregation in dividing cells are challenges which all organisms are faced with. Ensuring that these processes occur without error is essential for the survival of the organism. While the major mechanisms governing chromosome replication and segregation have been elucidated in eukaryotic organisms, analogous processes and their details in prokaryotic organisms have been more challenging to analyse. In order to understand the processes behind the localisation of the chromosome during cell division, this project has aimed at analysing the dynamics of 13 fluorescently labelled loci over the cell cycle of Escherichia coli. The results of this project can be used for further analysis of the chromosome in a large-scale study where more loci are analysed to map the dynamics of the whole chromosome.

The fluorescent labelling was achieved by introducing the parS sequence at the chosen sites with lambda-Red recombination and expression of ParB fused to the fluorescent protein mCherry. The sequence was successfully introduced at eight different positions in eight separate strains. The introduced parS/ParB system was confirmed to result in fluorescent foci by fluorescence microscopy imaging of the strains on agarose pads. Three of these strains were analysed in a microfluidic PDMS chip platform with fluorescence microscopy. Microfluidic systems provide an advantage of capturing large amounts of cells and making it possible to analyse them continuously in the same conditions. Combining these systems with bright-field, phase contrast and fluorescence imaging, the growth rates of the cells and dynamics of the fluorescent foci were successfully analysed over several hours.

Handledare: Johan Elf
Ämnesgranskare: Disa Larsson Hammarlöf
Examinator: Jan Andersson
ISSN: 1401-2138, UPTEC X 18 017

Sammanfattning

Mikroorganismer finns överallt. De omger oss, interagerar med oss och har därmed en betydande roll i människans liv. Bland den enorma variation av mikroorganismer som har utforskats finns bakterierna. Antalet bakterieceller som vi bär på är större än antalet mänskliga celler som vi består av. De är involverade i en mängd processer som sker konstant i vår omgivning. Bakterierna som man har kunnat identifiera utgör en definitiv minoritet av de som finns där ute, bland annat på grund av hur utmanade det har varit att återskapa förhållanden som de lever i. Deras vikt kan inte överskattas. Människans vilja att förstå hur liv fungerar har lett till en enorm utforskning av hur bakterier fungerar som organismer, vilka processer de använder sig av och hur dessa kan ha utvecklats. Denna forskning har lett till en rad olika tillämpningar där dessa organismer används. De har utnyttjats enormt som modeller för att förstå livsviktiga processer. Genom att försöka upptäcka de mekanismer som är viktiga för att liv ska kunna finnas har den nya förståelsen drivit utvecklingen av nya metoder som har gynnat mänskligheten. Många frågor återstår dock och forskningen fortskrider för att hitta svaren på dessa.

En av dessa frågor har inte fått ett tydligt svar trots att många har tacklat den i mer än 50 år. Detta handlar om utmaningen som bakterierna står inför när de delas från en till två celler. Fokus ligger på vad som händer med bakteriens DNA under celldelningsprocessen. DNA finns packat som ett nystan i bakterien och innehåller all information som är nödvändig för att bakterierna ska kunna fungera korrekt och kunna överleva. Delningen av bakterier måste alltså innefatta kopiering av allt DNA och lokalisering av det till vardera bakteriecell, en process där detaljerna ännu inte har blivit tydliga.

En gång trodde många att bakterier var som säckar av oordnade molekyler, med DNA som ett nystan i bakteriens inre, en uppfattning som sedan flera år har förändrats. Forskare har förstått att DNA har en tydlig organisation som upprätthålls i cellerna. Denna har även visats ha en påverkan på specifika funktioner i organismerna. Ett antal olika modeller för fördelningen av DNA mellan cellerna och viktiga komponenter i processen har föreslagits, med olika grad av stödjande bevis, men de essentiella faktorerna har varit utmanande att identifiera.

Detta arbete syftar till att uppnå bättre förståelse av vad som händer med DNA hos bakterier när de genomgår celldelning. Genom att visualisera de olika delarna av denna livsviktiga molekyl och observera hur dess organisation förändras i realtid har det blivit möjligt att få en insikt i hur denna process går till i större detalj. Det kan svara på frågan om hur fördelningen av DNA sker mellan bakterieceller samt vilka mekanismer som kan vara involverade. Tidigare föreslagna modeller kan stödjas av resultaten från detta arbete, men nya upptäckter kan även resultera från detta projekt.

Table of contents

Abbreviations

bp	base pairs
CMOS	complementary metal-oxide-semiconductor
CRISPR	clustered regularly interspaced palindromic repeats
DNA	deoxyribonucleic acid
EMCCD	electron multiplying charge-coupled device
FISH	fluorescence *in situ* hybridization
IPTG	isopropyl β-D-1-thiogalactopyranoside
kb	kilobase
LB	lysogeny broth
NAP	nucleoid associated protein
PCR	polymerase chain reaction
PDMS	polydimethylsiloxane
SMC	structural maintenance of chromosome

1 Introduction

All organisms are met with the challenge of compacting their DNA inside cells where it occupies only a small fraction of the total volume of the cell. Before each cell division this structure has to be replicated in order for both of the sister cells to contain the same genetic information. These are essential processes important for the survival of the organism, both with regards to eukaryotes and prokaryotes. While many advances have been made in getting insight into the processes that govern cell division in eukaryotic organisms, they still remain unclear with regards to the prokaryotic cell cycle. It is not known if the chromosome is organized in the same way at specific loci during different stages of the cell cycle or if this process is stochastic. The effect of this organisation on gene expression is also not clear. The bacteria that make up a large part of the prokaryotes are abundant and interact with a wide variety of different organisms, including humans. Many of them, such as the well characterised bacteria *Escherichia coli* are used in various applications and understanding how they function is thus valuable. It helps understanding the processes essential to life. It also highlights how evolution has achieved functioning replication and cell division systems in eukaryotes and prokaryotes, both with regards to their similarities and differences. Achieving a better understanding of how the organisation of the bacterial chromosome changes during cell division is the aim of this project.

Understanding how the bacterial chromosome is organized and segregated during cell division has been a challenging task for several decades. In contrast to the replication and organisation of the eukaryotic chromosome elucidating the same properties of the prokaryotic chromosome has been difficult. This is partially because of the low accuracy of existing methods and the difficulty to identify essential genes for chromosome segregation (Reyes-Lamonthe *et al*. 2008). In cases where the bacterial replication rates exceed the division rates, the analysis can also become even more daunting due to overlapping replication cycles (Wallden *et al*. 2016). The observation and visualisation of loci on the bacterial chromosome has also been halted by the lack of methods which could achieve a sufficiently high resolution. This obstacle has been tackled during the last decades where methods which use fluorescent molecules combined with super-resolution microscopy have been developed. These allow for visualisation of single molecules, such as those inside cells (Reyes-Lamonthe *et al*. 2008).

Among the methods that have been applied in loci visualisation is fluorescence *in situ* hybridization (FISH). FISH uses DNA probes which hybridize to the target sequence on the chromosome. The sequences of the probes are complementary to the locus and are also labelled with a fluorophore. The probes are applied to fixed cells which have been permeabilized chemically in order to detect the binding between the probe and target sequence with fluorescence microscopy (Reyes-Lamonthe *et al.* 2008).

Another alternative for visualisation of different components in the cell is the use of fluorescent proteins. The gene for the fluorescent protein is often introduced at either end of the coding sequence for the target protein, resulting in expression of a fusion protein from the native promoter for the coding sequence. In contrast to FISH, where the cells have to be fixed and permeabilized the fluorescent proteins can be used to image live cells. The fluorescent protein fusions have been used to visualize different loci by expressing fluorescent proteins fused with various repressor proteins. The Lac and Tet repressors are both examples of proteins that have been used for this approach. By introducing the operator sequences for these repressors at specific positions on the chromosome, combined with expressing the repressor fused to a fluorescent protein the interaction can be observed with fluorescence microscopy (Reyes-Lamonthe *et al.* 2008). A similar approach uses proteins involved in chromosome segregation in certain organisms (Funnel 2016), the ParB protein and the *parS* sequence. The *parS* is similar to the operator sites in that ParB binds specifically to it. It has also been shown that ParB proteins form complexes at the *parS* site (Funnel 2016). This can be taken advantage of by expressing ParB with a fluorescent protein fused to it, resulting in a large number of fluorescent proteins at the *parS* site and thus, a strong fluorescent signal that can be observed with fluorescence microscopy (Reyes-Lamonthe *et al.* 2008).

With regards to the different parts and components of the chromosome itself some of them should be described. Contrary to the eukaryotic cells, the chromosome is not enclosed by a nucleus, but is instead folded inside the bacteria in a structure known as the nucleoid (Reyes-Lamonthe *et al.* 2012). It has been established that the nucleoid is a discrete structure and has a well-defined shape, where specific domains that make up the structure have been suggested (Reyes-Lamonthe *et al.* 2008, Kleckner *et al.* 2014, Lioy *et al.* 2018). The nucleoid has also been shown to be radially confined within the cell (Fisher *et al.* 2013). It comprises all of the genomic DNA in the cell as well as a number of proteins.

The origin and terminus are regions on the chromosome that are involved in replication initiation and termination, respectively. They have been shown to be present around the mid-cell position, with a left and right chromosome arm on each side of the cell (Nielsen *et al.* 2006). However, *E. coli* cells with origin and terminus at the poles of the cells have also been observed (Bates & Kleckner 2005). This has led to two different nucleoid configuration types being proposed as well as the possibility that the cells can switch between the two (Kleckner *et al.* 2015). While most bacteria have circular chromosomes, such as *E. coli*, some of them have linear variants like many eukaryotes do (Reyes-Lamonthe *et al*. 2012). Replication initiation occurs at the origin and is mediated by various proteins, with DnaA being a key component in the initiation process. DnaA is an ATPase involved in the separation of the double-stranded DNA, allowing for assembly of the replisome, a protein complex that mediates replication elongation along the DNA. The terminus region is comprised of a number of termination sites spanning approximately 400 kb to which the Tus protein is bound. Once the replisomes reach the terminus, the elongation is terminated in a process that is not fully understood (Reyes-Lamonthe *et al*. 2008).

The many efforts to get a better insight into the processes that govern cell division have resulted in a number of different models for chromosome segregation (Lemon & Grossmann 2001, Bates & Kleckner 2005, Nielsen *et al.* 2008). Various reviews have emerged that try to unify these chromosome segregation models and the recent advances in determination of which proteins could be important for the segregation process (Reyes-Lamonthe *et al.* 2008, Reyes-Lamonthe *et al.* 2012, Kleckner *et al.* 2014). The accepted view describes the nucleoid as a distinct structure which has specific domains. Replication is initiated at the origin around the mid-cell position, where the replisome is assembled and proceeds bidirectionally. The bacterial chromosome has a left and right arm termed replichores and each arm is replicated separately by a replisome (Nielsen *et al.* 2006). Replication termination occurs at the terminus region, which is also located at mid-cell. Following replication certain parts of the sister chromatids experience a short period of cohesion before they segregate. Various conclusions have been drawn regarding sister chromatid cohesion based on the performed studies (Bates & Kleckner 2005, Nielsen *et al.* 2006, Reyes-Lamonthe *et al.* 2008). Several studies have suggested that while a brief cohesion period is observed the sister chromatids are segregated sequentially. The newly replicated origins move to the 1/4 and 3/4 positions of the dividing cell, resulting in localisation close to the poles of the respective sister cells. Similarly, the termini become positioned near the mid-cell which becomes a new pole in the sister cells (Bates & Kleckner 2005). However, as described earlier the configuration of origin and terminus being present at the poles or at mid-cell in non-replicating cells have both been observed. A switching mechanism between the two has been proposed, but at replication initiation these components are always found at mid-cell in *E. coli* (Kleckner *et al.* 2015).

In order to analyse the dynamics of the different parts of the chromosome the aforementioned *parS*/ParB system has been used in this project. The *parS* sequence was introduced at different positions on the chromosome combined with expression of fluorescently labelled ParB from a plasmid in order to achieve distinct fluorescent signals resulting from the interaction between these components. With regards to the scale of the project, it would be desirable to achieve a high resolution analysis, making it possible to map the dynamics of the whole chromosome. This would however require a large number of *parS* sequences being introduced at different positions on the chromosome. The time constraints of the project make this large scale analysis difficult to achieve. It would be possible by using a methodology similar to the one presented by Lawson *et al.* (2017). This would involve using a CRISPR/Cas9 system combined with a pool of sgRNA sequences targeting a site on the chromosome where *parS* would be introduced with homologous recombination. CRISPR/Cas9 uses RNA sequences to target specific sites and cleave them enzymatically. The system and its other variants have been derived from a bacterial defence mechanism and have been used extensively for genetic modifications (Zerbini *et al.* 2017). Thus, its use would lead to cleavage of the chromosomal site at which *parS* has not been successfully integrated, resulting in survival of only the *E. coli* cells which have successfully integrated the sequence. However, this project has focused on performing a smaller scale analysis using λ-Red recombination for the introduction of the *parS* sequence at 13 sites in order to visualize

these loci with fluorescence microscopy and determine their dynamics during the cell cycle. Some of these sites have previously been used for chromosomal integration (Stouf *et al.* 2013, Zerbini *et al.* 2017, Lawson *et al*. 2017).

This project thus aims to elucidate the mechanisms of chromosome segregation in the bacteria *E. coli* over its cell cycle in order to achieve a better understanding of how the DNA in the mother cell is replicated and transferred to the daughter cells. While this is a small scale analysis with regards to the number of sites on the chromosome that are analysed, this approach can be used to determine if the described methodology could be viable in order to be able to perform the large scale analysis in the future. The results from the analysis performed in this project as well as the subsequent large-scale analysis could shed light on one of the essential processes for *E. coli* and the results could be representative for a large number of different organisms.

2 Background

2.1 Bacterial chromosome segregation models

The analysis of bacterial chromosome segregation has resulted in several different models which, in some cases, do not support each other. The differences between them show how challenging the study of this process has been. One of the earliest attempts at elucidating the mechanisms behind chromosome segregation suggested binding of the chromosome to the inner membrane of the bacteria and that replication elongation would result in growth of the chromosome along the length of the bacteria (Reyes-Lamonthe *et al.* 2008). Another model that has been widely discussed is the presence of a mechanism similar to the one found in eukaryotes, more specifically with regards to the transfer of the replicated chromosomes via their centromeres. The models that are based on these mitotic-like mechanisms have suggested the *migS* site in *E. coli* to function as a centromere-like sequence (Reyes-Lamonthe *et al.* 2008). Similar theories have been proposed for a *parS*/ParB system in *Caulobacter cresentus* and the protein Spo0J with several origin-proximal binding sites in *Bacillus subtilis* (Toro & Shapiro, 2010). However, in the case of *E. coli* the *migS* site has been shown to be non-essential for the process. While a mitotic spindle-like machinery has not been identified in *E. coli* the structural maintenance of chromosome (SMC) proteins MukB, MukE and MukF have been implicated in having an important role in chromosome segregation through binding to the origin in these bacteria. The mechanism behind their function is however not known (Wang *et al.* 2013, Badrinarayanan *et al.* 2015).

An often cited segregation model describes the chromosome segregation being mediated by a so called replication factory (Lemon & Grossman 2001, Reyes-Lamonthe *et al.* 2008). The first formulations of this model were based on results seen in *B. subtilis* and it was speculated that a similar mechanism would be present in other bacteria, such as *E. coli* (Lemon &

Grossman 1998). The model describes the presence of a replication factory formed by replisomes at the mid-cell position. During replication the DNA would be tethered through this factory, while it would remain fixed. It has been suggested that the replisomes release the replicated DNA towards either end of the replicating cell, mediating its segregation into the sister cells. The most recent variant of the model combined this with capturing of the replicated DNA at the cell poles via a suggested membrane-associated complex (Lemon & Grossman 2001). Several studies have since shown that the replisomes are mobile in *E. coli*, moving along the left and right chromosome arm respectively in the cell (Bates & Kleckner 2005, Nielsen *et al.* 2006, Reyes-Lamonthe *et al.* 2008). However, it should be noted that while the replisomes have been shown to be mobile, certain studies have presented data where they are not confined to the respective cell halves during replication (Wallden *et al.* 2016).

When analysing the positioning of certain loci on the bacterial chromosome during cell division, a delay between finished replication of a locus and the segregation of its sister locus has been observed. This has been described as a cohesion mechanism between the sister chromatids (Bates & Kleckner 2005). The mechanism was proposed following the observation that after certain loci had been replicated, the sisters did not separate until several minutes later. The same experiments also showed that the segregation of sister chromatids appeared to occur around the same time as splitting of the nucleoid, proposing that segregation occurs simultaneously for the whole chromosome (Bates & Kleckner 2005). While a brief delay between replication and segregation has been observed, some argue that it is not part of a dedicated mechanism. It has been suggested that the cohesion period could give time for other processes to occur, such as homologous recombination. Several studies also show that segregation occurs sequentially following replication rather than simultaneous segregation of the whole chromosome (Nielsen *et al.* 2006, Reyes-Lamonthe *et al.* 2008). The cohesion could result from entanglement between the sister chromatids. Segregation would thus occur after enzymes, such as topoisomerases, have disentangled the structure (Badrinarayanan *et al.* 2015).

While the process behind achieving movement of the replicated chromosome in the cell is not fully understood, the different models propose various explanations. While many arguments against the replication factory model have been presented, the replisome could provide a part of the energy necessary to move the chromosome and thus contribute to the segregation process (Badrinarayanan *et al.* 2015). Entanglement between sister chromatids has been suggested to cause the cohesion between them, which could result in tension accumulating in the entangled structure. The release of this tension could mediate the segregation process. This is somewhat supported by results which show that mutations in the coding sequences of proteins involved in relieving tension and entanglement can interfere with successful chromosome segregation. Segregation of certain loci with long cohesion periods has been observed to occur in an abrupt manner, consistent with tension release contributing to the segregation process (Badrinarayanan *et al.* 2015).

Active transport processes have also been described, with the suggested centromere-like mechanism being an example. While the MukBEF complex has an important role in segregation, different nucleoid associated proteins (NAPs), such as HU, H-NS, IHF and Fis could be involved as well (Wang *et al.* 2013, Badrinarayanan *et al.* 2015). These NAPs have been described as analogous to eukaryotic histones in *E. coli*. They are bound to the chromosome to provide a structure which facilitates compaction of it and they could thus have a role in the segregation process as well (Wang *et al.* 2013). It has been proposed that the segregation could be driven by entropy. This model describes the chromosome as a self-avoiding polymer and by trying to achieve maximal entropy the sister chromatids become segregated between the sister cells (Jun & Mulder 2006). However, while the model is promising the chromosome has been described as a self-adhering structure and not a self-avoiding one. The nucleoid has also been shown to have a defined structure and proposing that the chromosome would randomly fill the volume of the cell contradicts this (Kleckner *et al.* 2014).

2.2 Microfluidics

The development of microfluidic systems has allowed for detailed analysis of both single cells as well as single molecules. These systems make it possible to easily capture and observe the analysed organisms with various microscopy techniques. The miniaturization allows for high throughput measurements to be performed. The use of microfluidics makes it possible to achieve separation, quantification and sorting of cells or molecules. This has been achieved by manipulating, among other physical properties, the flow rate in the various microfluidic channels, using acoustics or electrophoretic methods (Reece *et al.* 2016). Other applications involve phenotyping and analysis of gene expression (Lawson *et al.* 2017) as well as diagnostic studies (Baltekin *et al.* 2017). The use of these systems also makes it possible to effectively apply different types of growth media or reagents to the cells. Repeated experiments in the same growth conditions become easily reproducible. Since a large number of cells, depending on the design of the system, can be analysed in the same growth conditions it is possible to achieve statistical significance in the analysis. Certain designs can allow for analysis of cells during several generations as well (Baltekin *et al.* 2017, Lawson *et al.* 2017).

With regards to this project, the use of microfluidic chips allows for the capture of genetically modified bacterial strains in the channels of the chip through which growth medium can be applied. The cells can grow exponentially and divide in these conditions for several days, while simultaneously the growth and cell division processes can be observed with fluorescence microscopy in order to analyse them. These microfluidic systems are fabricated as chips, often made from a transparent polymer with channels embedded in them, ranging between the micrometre to the nanometre scale. As these systems have made it possible to miniaturize laboratory platforms they have been referred to as Lab-On-Chip devices and thus also microfluidic chips (Reece *et al.* 2016). While various polymers have seen applications in

this field, polydimethylsiloxane (PDMS) has often been used as it can easily be cast using designed moulds (Duffy *et al.* 1998).

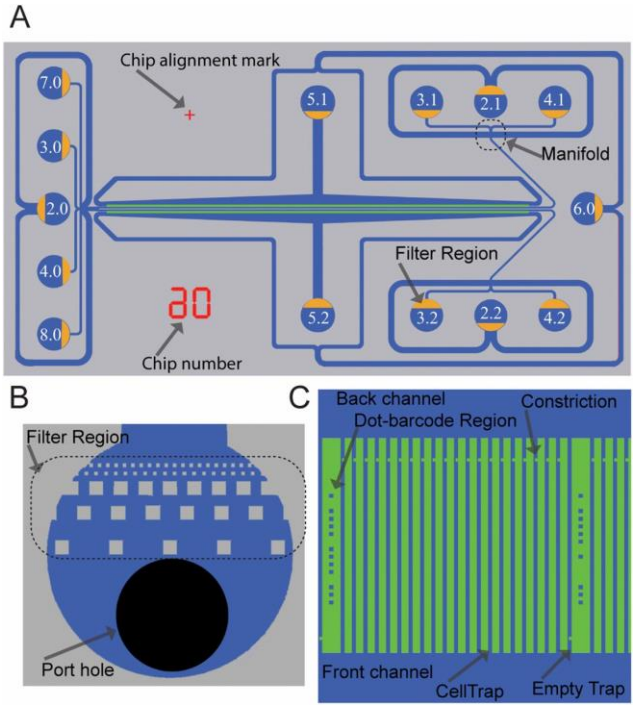


Figure **1 A sketch describing (A) the microfluidic chip design, (B) the port region of the microfluidic chip and (C) the traps used to capture cells, which are located near the centre of the chip. The figure has been previously published by Baltekin *et al.* (2017). The illustration is used with permission from Johan Elf.**

The microfluidic chips used in this project are based on designs presented by Baltekin *et al.* (2017). As depicted in the sketch in Figure 1, the microfluidic chip has a region containing 4000 channels where cells can be captured, thus called traps. Each chip contains two rows of traps 150 µm apart. At one end of these traps there is a 300 nm constriction which allows for the applied solutions, such as growth medium, to flow while preventing the cells from moving through the channels. Medium can flow over the cells allowing them to grow without escaping from the traps.

Based on the moulds used for the preparation of these chips, the dimensions of the traps can vary. They can be chosen based on the organism that is studied and its size, which can also depend on the growth medium used. In this project, chips with 50 µm long traps with a 1.0 × 1.0 µm cross-sectional area were used. The traps are divided into sets which are identified by a dotted binary barcode, based on 12 dots next to each set of traps. Each set also contains an empty trap, with the aforementioned constriction placed on the opposite side as compared to the other traps, preventing cells from entering this channel. This design is used in the image analysis in order to subtract the background signal from the signal measured in the other channels. At each of the ports there is also a filter region. It has been designed to prevent large particles from interfering with the cells and components in the chip. They also stop large air bubbles form entering the chip.

2.3 The impact of growth conditions on replication

As it has been previously shown, the growth rate of bacteria affects the point at which replication is initiated (Wallden *et al.* 2016). Organisms that achieve generation times that are shorter than the chromosome replication times, such as *E. coli*, need to initiate a replication round for the cells in the next generation before chromosome replication of the dividing cell has been terminated. This results in overlapping replication cycles (Cooper & Hemstetter 1968). If the bacterial chromosome is visualised by tracking a fluorescent signal at a locus it can appear as several fluorescent foci because of the several rounds of replication taking place. The overlapping replication cycles can thus make the analysis challenging (Reyes-Lamonthe *et al.* 2008). In fast growing *E. coli* several separate origins have been observed due to this process (Wallden *et al.* 2016). Since replication initiation depends on the growth rate of the bacteria the number of replication initiation events occurring in the dividing cell can be affected by the growth conditions. The use of a minimal growth medium can contribute to fewer rounds of replication and thus the presence of only one or two fluorescent foci (Wallden *et al.* 2016), making it less difficult to track the loci.

2.4 Tracking of molecules using phase contrast and fluorescence microscopy

Phase contrast microscopy is often used when analysing biological samples, such as cells. The method transforms variations in the refractive properties of the structures in the cells to variations in contrast which can be observed in the microscope. Compared with illumination using white light, often termed bright-field microscopy, phase contrast microscopy allows for visualisation of structures which would have otherwise required staining of the cells (Sanderson 2002). As this type of treatment can be damaging, phase contrast microscopy is thus usually used to study live cells. Figure 2 shows a comparison between a bright-field and phase contrast image of cells on an agarose pad. Light that is applied to a cell becomes diffracted which leads to a shift in its phase and a decreased amplitude compared to the light surrounding the cell. This phase difference cannot be observed directly but with the phase contrast method its visualisation is made possible by segregating and focusing the surrounding and diffracted light to create an image (Sanderson 2002). This is achieved by placing a condenser annulus between the lamp and the sample and a phase plate between the sample and the objective.

The phase annulus splits the wave of light into two waves, which become diffracted by the studied sample and thus results in the phase shift. The diffracted and non-diffracted light waves travel through the phase ring, which is attached to the phase plate. The surrounding, non-diffracted light becomes shifted in phase in order to become realigned with the phase of the diffracted light. Its amplitude is decreased, as the surrounding light initially has a higher amplitude than the diffracted light (Sanderson 2002). Using this method, the difference in phase between the two light waves becomes transformed to a visible difference in amplitude.

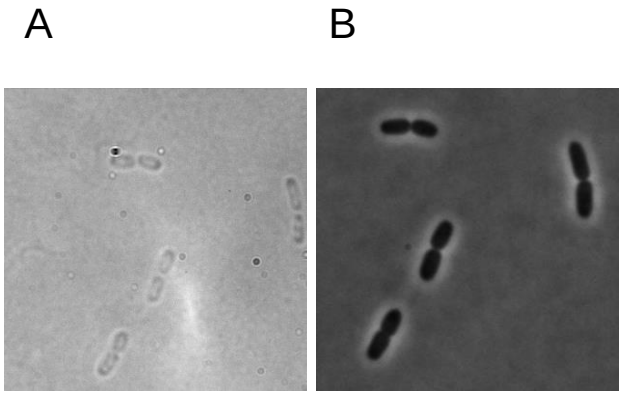


Figure 2 A bright-field (A) and phase contrast (B) image of *E. coli* cells on an agarose pad.

Fluorescence is based on the emission of light following absorption of light which leads to excitation. As energy is lost between the absorption and emission, the wavelength of the absorbed light is usually shorter than that of the emitted light (Lichtman & Conchello 2005). This property can be used to track cells or molecules by labelling them with a fluorescent compound. Since the absorption and emission light wavelengths are different, the light coming from the sample can be filtered in order to only follow the emitted light. Labelled molecules can be tracked over time in various structures such as cells. These molecules can be imaged through microscopy by applying laser light with a wavelength that is absorbed by the fluorescent molecule followed by observation of the emitted light (Lichtman & Conchello 2005). It can be combined with bright-field and phase contrast microscopy to localize the tracked signal in the cells.

An important property to consider when fluorescence is used for tracking is photobleaching of the fluorescent molecules (Lichtman & Conchello 2005). Fluorescent molecules can cycle between excitation by the absorbed light and emission of light, but after a number of cycles the intensity of the fluorescent signal decreases until no light emission occurs. This is termed photobleaching. While the details behind the process are not clear for all fluorescent molecules, it has been proposed that the absorbed energy that excites the molecules can be transferred to molecular oxygen (Lichtman & Conchello 2005). This leads to the formation of reactive singlet oxygen which damages the fluorescent molecule and thus affects its ability to fluoresce. It also affects cells that are being imaged by interacting with various molecules, which can damage them as well (Lichtman & Conchello 2005). In order to avoid this process when tracking molecules, the laser light is not applied continuously to the sample and is instead shuttered. It is also applied at relatively low intensities (Lichtman & Conchello 2005).

2.5 Image processing and analysis

Image analysis and tracking of cells and molecules involve several steps. In this project the algorithms are similar to those presented by Lawson *et al.* (2017) and Baltekin *et al.* (2017), with certain variations. The initial processing of the images involves identification of a region of interest, detection of the traps where cells are growing as well as empty ones and

subtraction of background signal. Region of interest identification is performed by analysing the 16-bit gray value of the pixels and summing them along the y axis. This results in an intensity profile for each analysed image. The process also involves detecting the maximal x values in the profile to identify the bright vertical lines that can be observed at the edges of the trap region. These are used to crop the image and define the region of interest.

For trap detection, the pixel values along the x-axis are summed to calculate an intensity profile for the y-axis of the image. This can be used to detect the maximal values along the y-axis, which will correspond to the traps as they are brighter than the PDMS background signal. While similar to the region of interest detection, this process is performed as a first estimate of the trap positions. For further images a cross-correlation between the image and a cell trap mask image is calculated to identify the specific trap locations. The mask image consists of a trap-sized white rectangle with a black background. The detection of empty traps is performed by summation of pixel values along the x-axis. Since the empty traps will have the smallest deviations in intensity, they are identified as the traps with the lowest summed derivative. The identified empty trap is subsequently used for background subtraction.

The cells in the images are identified through cell segmentation. The approach is based on the method present by Sadanandan *et al.* (2014). The algorithm analyses the curvature in the images and performs thresholding with an ellipse fitted to the objects identified in the image (Sadanandan *et al.* 2014). Once the cells have been segmented they are tracked during growth by the Baxter algorithm (Magnusson *et al.* 2015). Generally, the algorithm uses the results from the segmentation algorithm to track the cells in an automated manner. It connects the images to achieve tracking by using a score function to identify the most probable track for a given image sequence. It adds cell tracks sequentially based on maximizing the score function, which achieves the most probable tracking for the images (Sadanandan *et al.* 2014). The addition is terminated when no new tracks can increase the scoring function. While this algorithm is based on previously presented methods (Magnusson *et al.* 2015), it has been modified to be able to analyse prokaryotic cells and cells that are dividing.

The identification of fluorescent dots is performed by using the à trous wavelet transform of the images. This approach is based on filtering the images and detecting wavelet coefficients that have high values in these images. These maximal values have been shown to correspond to significant irregularities in the images, which can be identified as fluorescent dots (Olivo-Marin 2002). As these coefficient values do not occur in the presence of large amounts of background signal, the wavelet transform provides an accurate way of detecting the dots (Olivo-Marin 2002).

3 Materials and methods

3.1 Design of a template plasmid and identification of *parS* integration sites

A literature study was performed to identify 13 sites on the chromosome of *E. coli* at which *parS* could be integrated. In order to analyse the native cell division process, the choice of these 13 sites was based on avoiding changes in the phenotype that the integrations could cause. Thus, positions that had previously been used for introduction of DNA sequences with λ-Red recombination were seen as appropriate candidates for the integration of *parS*. Certain sites are pseudogenes or non-essential genes, which have been used previously without reports of phenotypic changes (Lawson *et al.* 2017, Zerbini *et al.* 2017). All integration sites and their respective position on the chromosome can be found listed in Table 1 in the Appendix. The 13 sites are referred to as position 1-13, according to the description in the table. The sites have also been illustrated in Figure 3 in order to show their approximate relative position on the *E. coli* chromosome.

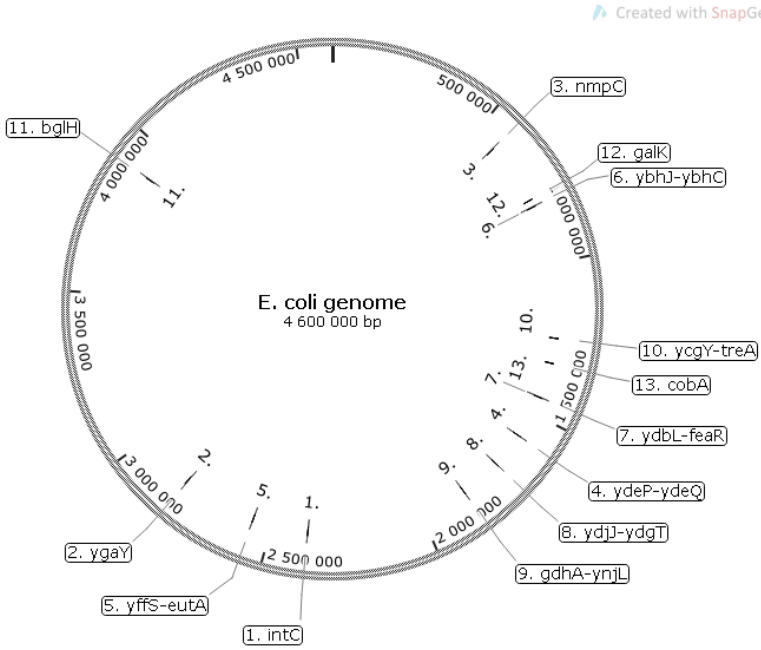


Figure 3 An illustration of the circular *E. coli* genome with the 13 sites that were chosen for the introduction of *parS*.

In order to introduce the *parS* sequence at different positions on the chromosome a template plasmid was designed. The plasmid design contained the *parS* sequence, a kanamycin resistance cassette, a SacB coding sequence as well as an ampicillin resistance cassette and a pUC19 replicon (Yanisch-Perron *et al.* 1985). Since the same sequence fragment containing *parS*, the kanamycin resistance cassette and a SacB coding sequence would be integrated at all 13 positions the plasmid would be used as a PCR template to amplify this fragment. Different homologous ends would be introduced in the primer sequences depending on the

chromosomal integration site. The kanamycin resistance cassette was included in the fragment to select for successful integration of the fragment. The introduction of a SacB coding sequence can be used for counter-selection as it expresses levansucrase, which makes the bacteria sensitive to sucrose. This would make it possible to remove the kanamycin resistance cassette and SacB coding sequence and analyse chromosomal dynamics only with the *parS* sequence present.

3.2 Golden gate assembly

The construction of a template plasmid used for the introduction of the *parS* sequence onto the chromosome was performed by assembling three DNA sequence fragments using Golden gate assembly. The fragment containing *parS* was amplified by PCR from the genome of *E. coli* strain *codA-parS-cynR*. All oligonucleotides used for PCR and sequencing are listed in Table 2, Table 3, Table 4 and Table 5, which can be found in the Appendix. The oligonucleotides were ordered from Integrated DNA Technologies, unless otherwise noted. Similarly, the second DNA fragment containing the pUC19 plasmid replicon and an ampicillin resistance gene were amplified by PCR from the pGuide8 plasmid. Finally, the fragment containing a kanamycin resistance gene as well as the coding sequence for SacB, which causes the bacteria to become sensitive to sucrose, was amplified from the genome of an *E. coli* KanR-SacB strain. See Table 6 and Table 7 respectively in the Appendix for details about all the strains and plasmids used. A strong synthetic terminator, termed L3S2P24 (Chen *et al.* 2013), was introduced in the reverse primer used to amplify the fragment with the kanamycin resistance cassette and SacB coding sequence. The terminator was introduced to avoid errors in transcription termination, which could have affected the phenotype of the strains. The amounts of different reagents and the protocol for these PCRs were adjusted based on the Q5 High Fidelity DNA polymerase protocol (New England Biolabs). The subsequent PCRs were based on this protocol unless otherwise stated.

The lengths of the fragments were verified by 0.8 % agarose gel electrophoresis. For the *parS* fragment specifically, a 2 % agarose gel electrophoresis was used. Following PCR the reactions were purified using the GeneJet PCR purification kit (Invitrogen) by following the manufacturer's instructions. For the DNA sequence fragment containing the ampicillin resistance cassette and the pUC19 replicon, the template plasmid had to be removed by DpnI restriction in order to avoid subsequent false positive colonies. The restriction enzyme digestion was performed according to the manufacturer's instructions using the FastDigest DpnI (Thermo Scientific). This fragment was run on an 0.8 % agarose gel followed by gel extraction using the PureLink Quick Gel extraction kit (Invitrogen). The fragments were then used in a Golden gate assembly with the BpiI restriction enzyme (New England Biolabs) and T4 DNA ligase (Thermo Scientific) as described previously by Engler & Marillonnet (2013).

Golden gate assembly uses type IIS restriction enzymes, such as BpiI, and a ligase in order to efficiently assemble a number of DNA sequence fragments in a one-sample thermocycling reaction (Engler & Marillonnet 2013). Type IIS restriction enzymes cleave the restriction

sequence adjacent to their 6 bp recognition sites. The digestion results in formation of sticky ends which can be ligated enzymatically. Since the recognition site is not the same as the restriction site, the sequences of the sticky ends will depend on how the sequences adjacent to the recognition sites are designed. Thus, the recognition and restriction sites can be introduced at the ends of a DNA sequence fragment in order for them to be digested by the IIS restriction enzyme, followed by ligation. The restriction-ligation reaction occurs at 37 °C and 16 °C. Thermocycling occurs 50 times between these temperatures, followed by a digestion step at 50 °C and a heat-inactivation step at 80 °C (Engler & Marillonnet 2013).

The Golden gate reaction mix was subsequently used in a transformation with chemically competent One Shot Top10 *E. coli* cells (Invitrogen), following the manufacturer's instructions. The transformed cells were plated on agar plates with 50 µg/ml kanamycin to select for bacteria that had successfully taken up the plasmid. The colonies were also re-streaked on agar plates with 50 µg/ml kanamycin and 5 % sucrose. This re-streak was performed to see which of the colonies also contained a functional SacB coding sequence. The colonies were analysed with colony PCR using primers listed in Table 3. All colony PCRs used DreamTaq DNA polymerase (Thermo Scientific) unless otherwise noted, following the manufacturer's instructions.

The lengths of the colony PCR products were verified with 0.8 % agarose gel electrophoresis and purified as described before. The purified PCR products were analysed with sequencing using the Mix2Seq kit (Eurofins Genomics) with the primers listen in Table 3. Colonies which grew on agar plates with 50 µg/ml kanamycin and did not grow on agar plates with 50 µg/ml kanamycin and 5 % sucrose showed the desired phenotype. All agar plates and LB overnight cultures with kanamycin contained 50 µg/ml of the antibiotic, unless otherwise noted. The colonies that also showed the expected sequencing results were grown overnight in liquid LB culture with kanamycin. The LB was based on 10 g of Tryptone, 10 g of NaCl and 5 g of yeast extract dissolved in water, while the pH was adjusted to 7.3. The plasmid was extracted from these liquid cultures using the PureLink Quick Plasmid Miniprep kit (Invitrogen) according to the manufacturer's instructions. The plasmid was also analysed with sequencing to further confirm that the assembled plasmid was correct. The same primers were used for this validation as for the colony PCR products. Once colonies containing the desired plasmid, referred to as pParS, had been identified liquid cultures of these colonies were stored as cryostocks, with a 20 % final concentration of glycerol at -80 °C. All subsequent preparations of cryostocks were performed similarly, unless otherwise noted.

3.3 Introduction of *parS* onto the chromosome

The assembled pParS plasmid was used as a template for a PCR with primers that would introduce 40 bp long homologous sequences at the ends of the resulting PCR product. The primers used in this PCR are listed in Table 4. The lengths of the PCR products were confirmed with 0.8 % agarose gel electrophoresis. The PCR products were purified, followed by DpnI restriction enzyme digestion and gel extraction from a 0.8 % agarose gel, all as

described previously. The fragments were concentrated and purified using the SureClean Plus kit (Bioline), with the precipitate being resuspended in UltraPure DNase/RNase-Free distilled water (Invitrogen).

The target strain for the fragment, a wildtype *E. coli* BW25993 (Datsenko & Wanner 2000), was electroporated with the pSIM6 plasmid, a temperature-sensitive plasmid from which the necessary enzymes for λ-Red recombination could be expressed (Datta *et al.* 2006). The wildtype strain was grown from a cryostock overnight in LB, without antibiotics, followed by removal of salts from the cells. This was achieved with repeated centrifugation at 3000 × g and removal of the resulting supernatant, followed by resuspension in 50 ml of cold sterile water. In the last centrifugation step the cell pellet was resuspended in 150 µl of 10 % glycerol. The cell suspension was subsequently mixed with the concentrated DNA fragment and electroporated in cuvettes with a MicroPulser Electroporator (Biorad), followed by addition of LB and recovery at 30 °C on a shaker. After two hours of recovery the cell culture was plated on agar plates with 100 µg/ml ampicillin which were incubated at 30 °C overnight. A colony was subsequently grown overnight in LB with 100 µg/ml carbenicillin and stored as a cryostock. All agar plates with ampicillin and LB overnight cultures with carbenicillin contained a 100 µg/ml concentration of the antibiotics, unless otherwise noted.

Similarly, the *parS* sequence fragment with homologous ends was introduced into the cells by electroporation and λ-Red recombination. The target strain with pSIM6 was grown overnight from the cryostock at 30 °C in LB with carbenicillin. The culture was diluted 1:100 and kept at 30 °C until the optical density at 600 nm had reached 0.2-0.4. The λ-Red protein expression was then induced by growing the cells at 42 °C for 15 min. The cultures were chilled on ice, followed by removal of salts and electroporation, as described before. After the recovery the cultures were plated on agar plates with kanamycin.

The resulting colonies were subsequently analysed with colony PCR as before, using primers that flank the insertion site on the chromosome. These primers are listed in Table 5. The colonies used in the PCR were also re-streaked on agar plates containing kanamycin as well as kanamycin and 5 % sucrose, to confirm that the strains showed the correct phenotype. This would indicate that the correct sequences should be present. The lengths of the colony PCR products were verified using 0.8 % agarose gel electrophoresis. The products were then purified as described before. The colony PCR products that were amplified from colonies that showed the desired phenotype on the plates and the correct band length on the agarose gel were sequenced using the Mix2Seq kit (Eurofins Genomics). The *parS* confirmation reverse primer shown in Table 5 was used for sequencing the PCR products. The corresponding colonies were grown overnight in LB with kanamycin at 30 °C on a shaker. They were then stored as cryostocks.

3.4 Curing of pSIM6 and transformation of pMS11

The stored cultures were cured of the pSIM6 plasmid in order to avoid interference that it could cause with subsequent experiments. This was performed by growing cultures in LB with kanamycin from cryostocks at 42 °C on a shaker for three hours. The cultures were spread on agar plates containing kanamycin and incubated at 42 °C overnight. The resulting colonies were screened by re-streaking them on agar plates containing kanamycin and ampicillin respectively. As pSIM6 contains an ampicillin resistance cassette, cells that have been successfully cured of pSIM6 should not grow on plates with ampicillin. Colonies that grew on the agar plates with kanamycin but not on the plates with ampicillin were thus used in subsequent experiments. These colonies were also analysed with colony PCR, 0.8 % agarose gel electrophoresis and sequencing as before to confirm that no changes in *parS* had occurred. The colonies for which the results were correct were stored as cryostocks.

Overnight cultures with LB and kanamycin were prepared from the cryostocks. The salts in the medium were washed from the cells, as before which was followed by electroporation of the pMS11 plasmid. This plasmid contains the ParB coding sequence with the ParB expression being regulated by the *lac* promoter (Stouf *et al.* 2013). ParB is expressed in fusion with the red fluorescent protein mCherry. The plasmid was purified as before from the *codA-parS-cynR* strain, see Table 6 for details. Following electroporation, the cells were plated on agar plates with 50 µg/ml chloramphenicol. All subsequent use of agar plates and growth medium with chloramphenicol contained 50 µg/ml of the antibiotic, unless otherwise noted. The resulting colonies were analysed with colony PCR, 0.8 % gel electrophoresis, sequencing as well as plating on agar plates with kanamycin and kanamycin with 5 % sucrose. The colony PCR was performed using primers listed in Table 5, while sequencing was performed with the Mix2Seq kit (Eurofins Genomics) using the *parS* confirmation reverse primer in Table 5. The colonies which showed the desired phenotype and whose colony PCR products showed bands that corresponded to the correct length were used in subsequent experiments. If the colony PCR products for these colonies were confirmed by sequencing as before, the colonies were grown overnight in LB with chloramphenicol and subsequently stored as cryostocks.

3.5 Validation with agarose pads

Before performing the microfluidic chip experiments the cells were analysed with fluorescence microscopy on agarose pads. The pads were prepared by applying melted 2 % low melting point agarose mixed with M9 glycerol medium onto microscope slides, pressed down with cover slips to shape them. The medium contained 100 µM $CaCl_2$, 2 mM $MgSO_4$, 1x concentration of M9 salts, 0.4 % glycerol (v/v), 0.85 g/l Pluronic F108 and 1x concentration of an RPMI amino acids solution (Sigma-Aldrich). Cultures were grown overnight from cryostocks in LB medium with chloramphenicol, followed by a 1:200 dilution in the M9 glycerol medium, also with chloramphenicol. After growing the cultures at 30 °C on a shaker for 1-3 hours, 1 ml of each was centrifuged for 5 min at 3000 × g. Once the

resuspended culture solution had been applied to the agarose pad it was again pressed down with a cover slip. The slides with the agarose pads were subsequently analysed under the microscope to confirm the presence of a fluorescent signal resulting from the interaction between *parS* and ParB. A BW25993 strain with no *parS* present was also electroporated with the pMS11 plasmid as before. This strain was analysed on agarose pads as a negative control experiment. Bright-field and fluorescence images were taken for comparison for all of the described strains.

3.6 Microfluidic chip preparation

The microfluidic chips used for the fluorescence microscopy experiments were prepared similarly to Baltekin *et al.* (2017). These chips were made with moulded PDMS which is covalently bonded to a cover slip (No. 1.5). The PDMS was prepared by mixing a silicone elastomer base known as Sylgard 184 (Dow Corning) and a curing agent in a 10:1 weight ratio followed by mixing in a FastPrep-24 homogenizer (MP Biomedicals). The solution was subsequently centrifuged for 30 s at 4000 rpm. It was then poured onto the moulds followed by de-gassing under vacuum for 10 min. It was then cured at 100 °C overnight. The moulds used have been produced by NMetis as described by Baltekin *et al.* (2017). The designs for the moulds have been made previously by the Elf research group. With the described design the cured PDMS contains an array of chip features which can be diced into individual PDMS chips. The chips were punched at the marked ports to which tubes can be connected. The chips were subsequently rinsed in isopropanol (≥99.8 %, Sigma-Aldrich) to remove possible contaminations. They were dried by blowing pressurized air on them.

The cover slips to which the PDMS chips were bonded were cleaned by rinsing with deionised water, ethanol (96 %, Sigma-Aldrich) and again with deionised water. They were also cleaned by sonication in 2 % (v/v) Hellmanex III for 45 min. This was followed by several rounds of rinsing with Milli-Q water. The cover slips were then stored in a container with Milli-Q water.

Similarly to the chips, the cover slips were blow-dried as well and both were treated with plasma using the HPT-200 Benchtop Plasma System (Henniker Plasma). This results in an activated surface which allows for the bonding between the surfaces exposed to the plasma treatment to occur. The isopropanol rinsing, blow-drying and plasma treatment were all performed in a Scanlaf Mars safety cabinet (Labogene). Following the plasma treatment, the surface of the chip exposed to the treatment was applied to the cover slip in order for them to bond. The bonded chips and cover slips were incubated at 100 °C for one hour after which Scotch Magic Tape was put on the chip surface to keep it free from dust. They were then stored in room temperature until use. A bonded chip can be seen in Figure 4.

3.7 Cell loading and growth in microfluidic chips

The microfluidic chip used for the analysis of the strains was attached at the moveable stage of the microscope. Solutions that were applied to the chip were stored in 15 ml Falcon tubes. Application of the solutions and control of the flow rates was performed using pressure with the OB1-Mk3 (Elveflow). Tubing was connected to the chip with custom-made metal connectors with a 90° bend in the middle (New England Small Tubing), as depicted in Figure 4A.

A B

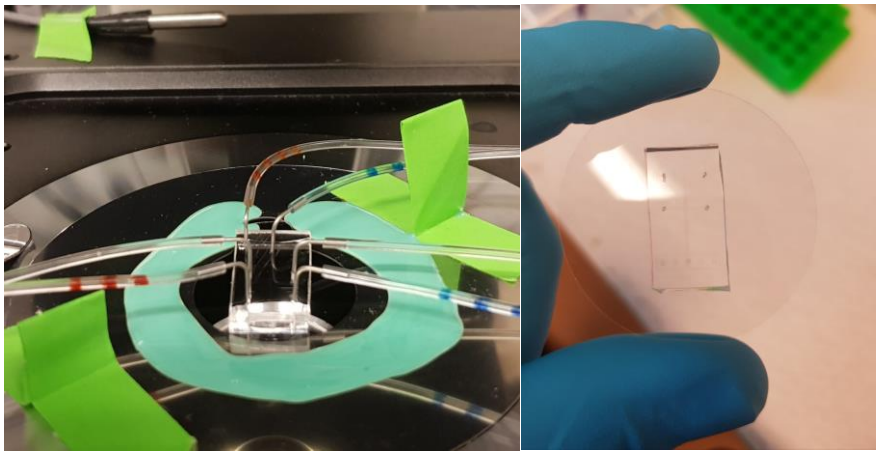


Figure 4 (A) A bonded microfluidic chip with tubing connected, shown at the stage of the microscope. (B) A microfluidic chip bonded to a coverslip, shown from underneath.

Prior to application of the cultures and media to the chip all tubing and connectors were washed with Milli-Q water, once with filtered 70 % ethanol (Sigma-Aldrich) followed by another wash with Milli-Q water. The chip was subsequently wet with M9 glycerol medium (0.4 % v/v) with chloramphenicol. Wetting of the chip was performed by applying medium to the back (5.1 and 5.2, see reference numbers in Figure 1), waste (2.0) and medium (8.0) ports. In order to confirm that the flow was directed away from the medium port, thus avoiding contamination, beads were used in the applied medium. This made it possible to visualize the flow in the channel from the medium port. Once the chip was filled with medium, cell cultures were applied at the loading ports (2.1 and 2.2) with 80 mBar. Pressure in the back ports was lowered to 30 mBar, 60 mBar in the medium port and no pressure was applied to the waste port. This allowed for cells to enter the traps. Once a majority of the traps contained cells, the pressure on the loading and back ports was lowered to 0, while the pressure on the medium port was increased to 120 mBar. Cells were then continuously supplied with medium, allowing them to grow in the chip. They were imaged at nine trap positions during growth.

When growing the cells in the chip the temperature was adjusted to 30 °C in a custom-made microscope cage incubator (Okolab), because of the temperature-sensitive replicon of pMS11

(Stouf *et al.* 2013). The temperature in the incubator was controlled by the Airtherm-Atx temperature controller (World Precision Instruments). Falcon tubes with media and cultures, which were kept outside the incubator in room temperature, were connected to the chip with tubing through which the solutions were applied, as it has been previously performed (Baltekin *et al.* 2017).

3.8 Microscope setup

The microscopy experiments were all performed using an inverted microscope with a Nikon Ti-E setup. For these experiments the 100x CPI Plan Apo Lambda objective from Nikon was used. For taking phase contrast images a CMOS camera DMK 23U274 (The Imaging Source) was used, while the bright field and fluorescence images were taken with an Andor Ixon EMCCD camera. A 561 nm Genesis MX laser from Coherent was used for the fluorescence imaging of cells, both on agarose pads and in microfluidic chips. The laser light effect was set to 15 mW for the agarose pad experiments and 10 mW for the imaging performed in microfluidic chips. All laser light shuttering was performed with the AOTFnC (AA Opto Electronics), while the TLED+ (Sutter Instruments) was used as a white light source. The experiments were performed with an Apo TIRF/1.49 100x oil-immersion objectives. The fluorescence and bright-field images of cells on agarose pads were taken with a 300 ms exposure time.

With regards to the cells analysed in the microfluidic chips, the phase contrast images were acquired every 30 s, while bright-field and fluorescence images were acquired every 3 min. All imaging of the microfluidic chips was performed with a 300 ms exposure time. The camera and the microscope were controlled using an open-source microscopy software (MicroManager 1.4.20, Edelstein *et al.* 2010), which was also used for all image acquisition. Bright-field, phase contrast and fluorescence images of the cells were acquired automatically at specified trap positions for eight hours using a custom-made plugin for this software.

3.9 Optimization of fluorescently labelled ParB expression in microfluidic chips

In order to determine which conditions would be suitable for tracking of ParB-mCherry in the cells that were analysed in microfluidic chips, a number of imaging experiments were performed with different concentrations of IPTG added to the medium. The microfluidic chips were wet and loaded with cells as described before. Cells were grown in the chips and imaged. The growth medium in the experiments contained increasing concentrations of IPTG in order to induce the expression of ParB-mCherry. The cells were grown in medium supplied with 0, 10 µM, 100 µM and 1 mM of IPTG. This analysis was only performed on the strain containing *parS*, a kanamycin resistance cassette and a SacB coding sequence at position 1 due to time constrains. The strain was grown overnight from a cryostock at 30 °C in LB with of chloramphenicol followed by 1:200 dilution in M9 glycerol (0.4 % v/v) with

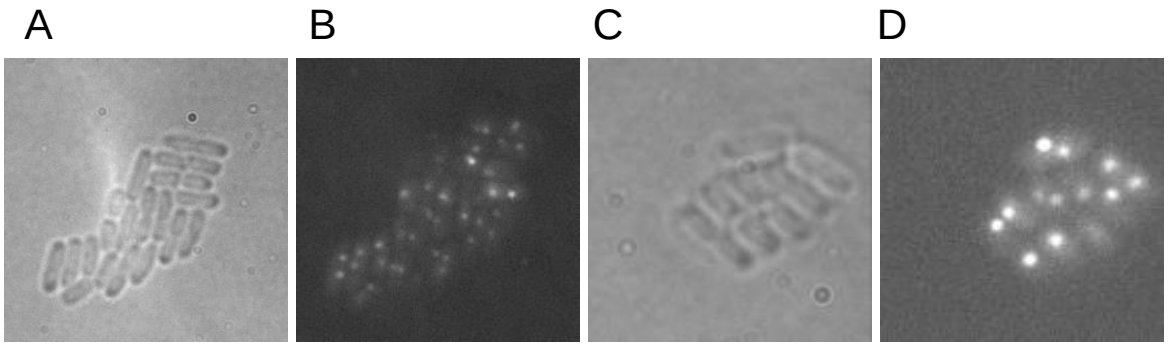
chloramphenicol and growth for several hours at 30 °C before loading the cells into the chip. Following this analysis, other strains where grown in growth medium supplied with 1 mM IPTG the microfluidic chips.

Variation in fluorescent signal due to the addition of the RPMI amino acids solution in the growth medium was analysed on agarose pads for the strain that had the *parS* sequence fragment integrated at position 9. The agarose pads were prepared as described previously. The strain was grown as described for the validation experiments but the growth medium was supplied with 0, 1 and 1.25 mM of IPTG as well as with or without the RPMI amino acids solution. The *codA-parS-cynR* strain, which has the *parS*/ParB system, was analysed similarly in the presence and absence of the RPMI amino acids solution for comparison. No IPTG was added to the growth medium of the *codA-parS-cynR* strain.

4 Results

4.1 The *parS*/ParB strains show fluorescent foci on agarose pads

Successful recombination of the *parS* sequence fragment was achieved at eight of the thirteen chosen sites on the *E. coli* chromosome. The results of the validation experiments with agarose pads seen in Figure 5A-5P show that fluorescent foci can be observed in all of the analysed strains. The results from the background strain with pMS11 and no *parS* present, seen in Figure 5Q and 5R, appear to emit a fluorescent signal from the whole volume of the cell. The validation was done to confirm that the expression of ParB from pMS11 and the presence of *parS* on the chromosome would result in fluorescent foci before analysing them in microfluidic chips.



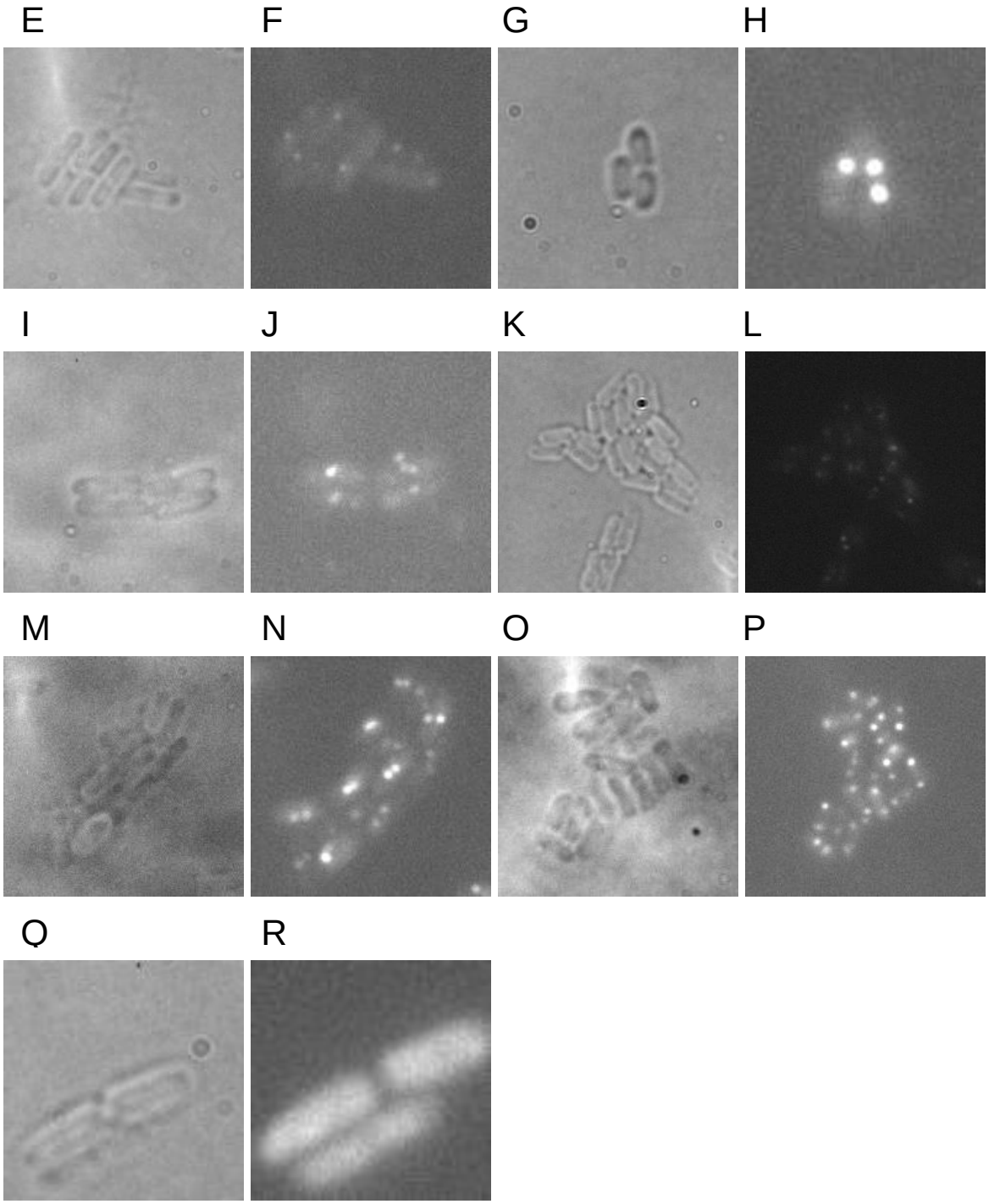


Figure 5 Bright-field and fluorescence images of cells with pMS11 and *parS* sequence present on the chromosome (A-P), as well as cells with pMS11 and no *parS* (Q, R) analysed on agarose pads. The *parS* sequence has been introduced at positions 1 (A, B), 2 (C, D), 5(E, F), 7 (G, H), 8 (I, J), 9 (K, L), 10 (M, N), 11 (O, P).

4.2 Induction with 1 mM IPTG is required to observe fluorescent foci in microfluidic chips

As suggested previously (Nielsen *et al.* 2006), over-expression of ParB can result in fluorescent foci that are challenging to track. While the expression of ParB-mCherry was controlled by the *lac* promoter it was not initially induced, in order to avoid tracking difficulties. The expression of ParB-mCherry from pMS11 was also not induced when presented in previous studies (Stouf *et al.* 2013). As seen in Figure 6A and 6B, showing the strain with the *parS* sequence fragment integrated at position 1 in microfluidic chips, growing the cells in the chip without an inducer in the growth medium resulted in a signal that was similar to the background signal. This strain was subsequently analysed with increasing concentrations of IPTG to achieve a stronger fluorescent signal. As seen in Figure 6, the increasing IPTG concentration results in an increase in the intensity of the fluorescent dots.

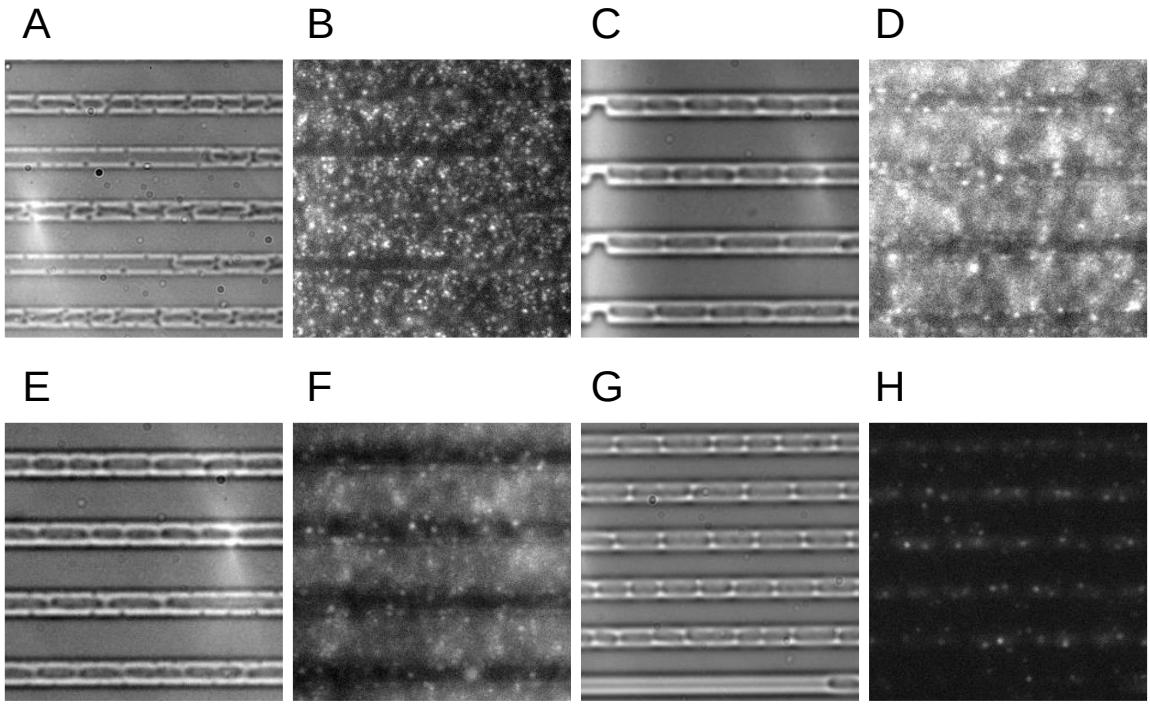
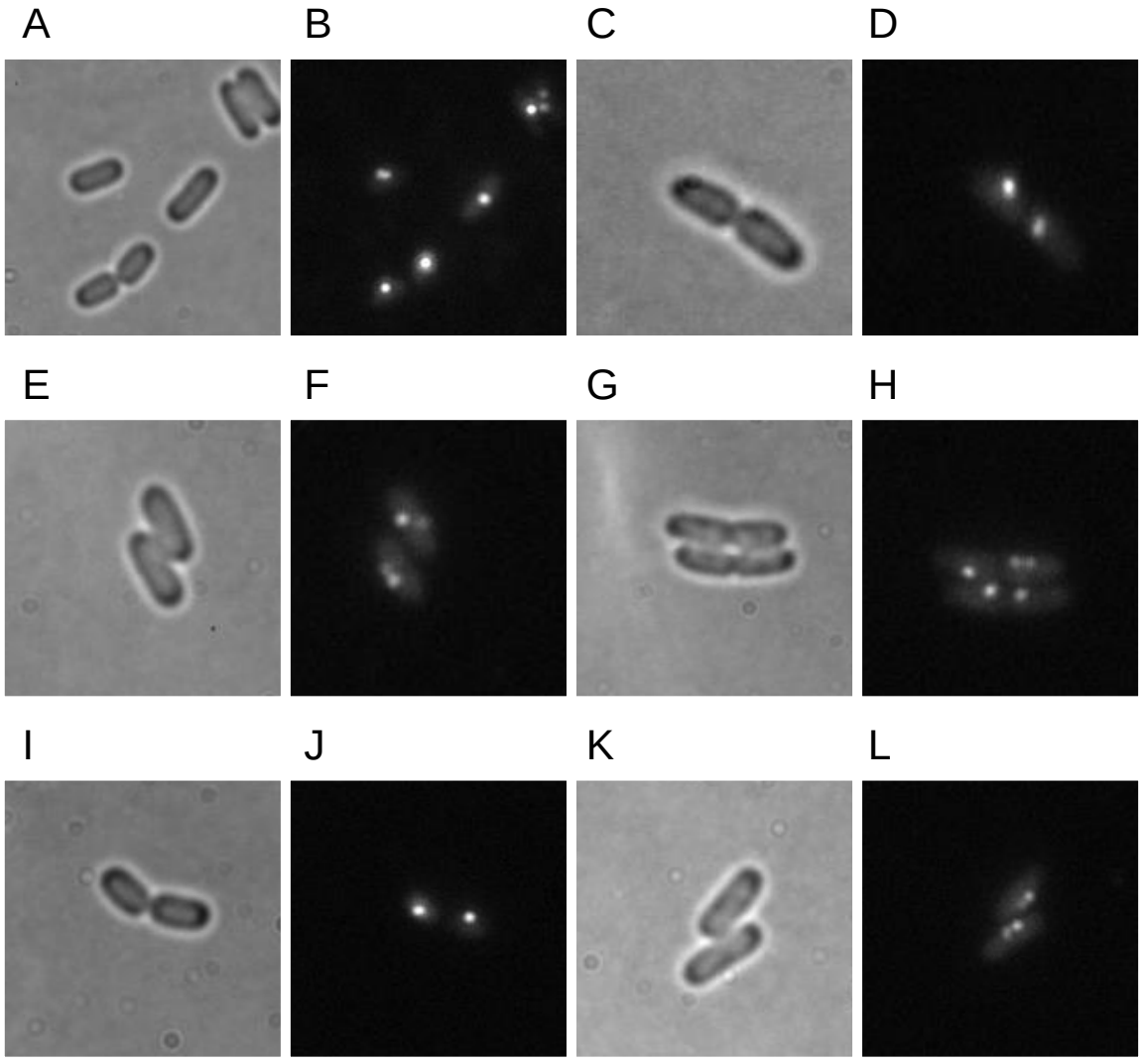


Figure 6 Bright-field and fluorescence images in microfluidic chips of a strain with the *parS* sequence fragment introduced at position 1. The strain was grown with 0 (A, B), 10 µM (C, D), 100 µM (E, F) and 1 mM (G, H) IPTG in the growth medium.

The *codA-parS-cynR* strain from which the *parS* sequence was amplified also has the *parS*/ParB system. It had been previously analysed by the Elf research group where the cells were grown in microfluidic chips. No RPMI amino acids solution was used in the growth medium when it was grown due to overlapping rounds of replication observed in the bacteria. Following initial microfluidic chip experiments of the strains shown in Figure 6 it appeared that the expression levels of ParB-mCherry were low. A similarly low fluorescent signal was not observed when the *codA-parS-cynR* strain with the same system had been studied previously. A comparison between this strain and the strain with *parS* introduced at position 9

was performed by analysing their fluorescent signal on agarose pads. The strain with *parS* at position 9 was grown with and without RPMI amino acids solution and 0, 1 and 1.25 mM IPTG in the growth medium. The *codA-parS-cynR* strain was grown with and without the RPMI amino acids solution as well but with no IPTG added to the growth medium. This was performed to compare the fluorescent signal in the respective strains and to determine if the presence of the amino acids solution would affect it. The presence of RPMI amino acids solution was the only difference in growth conditions between the two strains and the comparison was based on this. As seen in Figure 7 the difference in the fluorescent signal based on the presence and absence of RPMI is difficult to determine. In the absence of IPTG and in the case of induction with 1.25 mM IPTG, Figure 7A-7D and 7I-7L respectively, the signal appears to be stronger in the absence of the amino acids solution. The same relationship appears to be true for the *codA-parS-cynR* strain in Figure 7N and 7P. Based on Figure 7B and 7N, the fluorescent signal in the position 9 *parS* strain also appears to be relatively similar to that of the *codA-parS-cynR* strain when analysed in similar conditions.



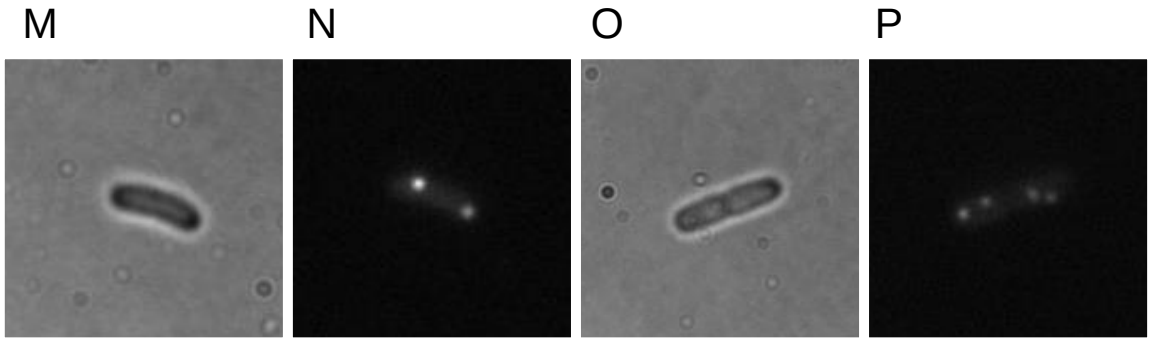


Figure 7 Bright-field and fluorescence images of the strain with the *parS* sequence fragment integrated at position 9 (A-L) and the *codA-parS-cynR* strain (M-P) imaged on agarose pads. The strain with *parS* at position 9 was grown without RPMI amino acids solution and with 0 (A, B), 1 (E, F) and 1.25 mM (I, J) IPTG in the growth medium. It was also grown with RPMI amino acids solution and with 0 (C, D), 1 (G, H) and 1.25 mM (K, L) IPTG in the growth medium. The *codA-parS-cynR* strain was grown without (M, N) and with (O, P) RPMI amino acids solution in the absence of IPTG.

4.3 The dynamics of *parS* at various loci can be analysed over the cell cycle

Three of the eight strains with the *parS* sequence fragment integrated were successfully imaged in microfluidic chips. Figure 8 shows the bright-field, phase contrast and fluorescence images of the strains. The strains which were analysed had *parS* integrated at position 1, 9 and 11 respectively. It should be noted that this figure shows some of the images taken in a longer image sequence, resulting from eight hours of imaging. The strains were all grown with 1 mM IPTG in the growth medium. The fluorescent dots seen in Figure 8 were successfully detected with the image analysis algorithms. The image sequences that were acquired were used in the image analysis to calculate the distribution of the *parS* sequence in the cells over the cell cycle, as shown in Figure 9A, 9C and 9E. The corresponding growth rate distributions for each strain were calculated as well, which can be seen in Figure 9B, 9D and 9F. Based on these distributions the growth rates of the three strains are similar, with approximately equal mean growth rates and standard deviations.

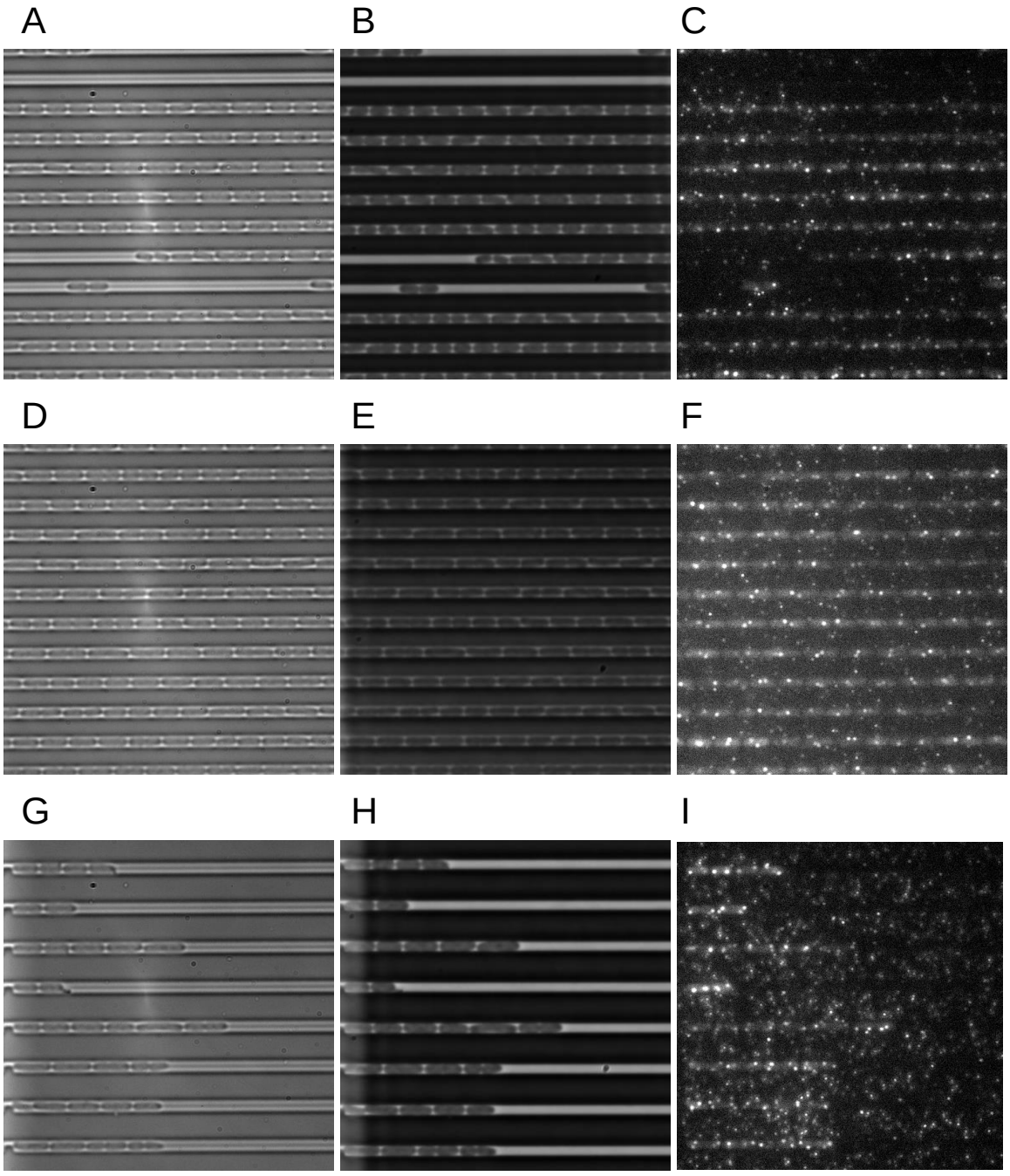


Figure 8 Bright-field, phase contrast and fluorescence images of strains with *parS* at position 1 (A, B, C), 9 (D, E, F) and 11 (G, H, I) grown in microfluidic chips. These images are part of an image sequence which was acquired during automated fluorescence microscopy imaging of the cells for eight hours.

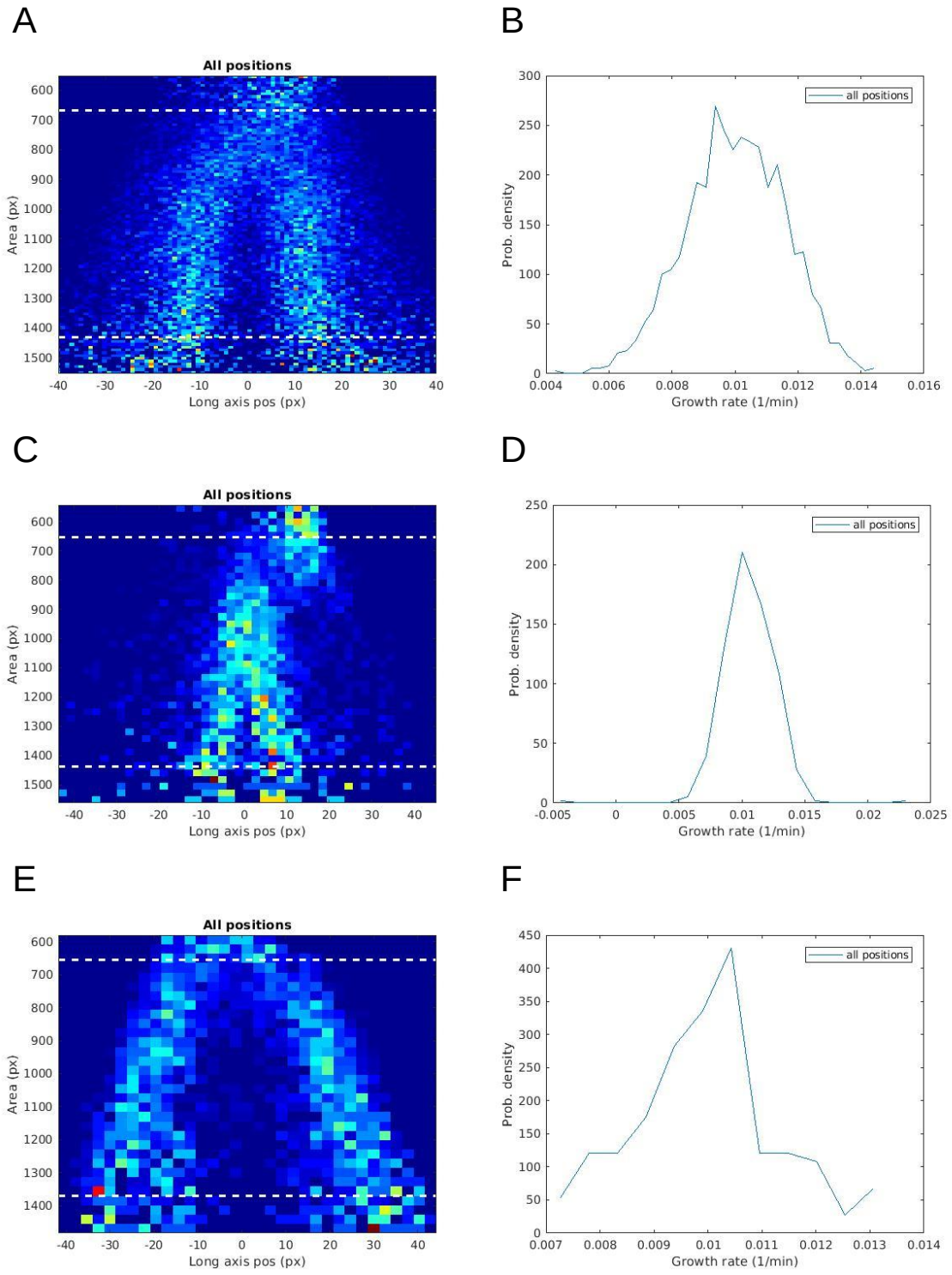


Figure 9 Distributions of the localisation of the *parS* sequence fragment along the long axis of the modified *E. coli* strains at different cell areas and their corresponding average growth rate distributions. The distributions were calculated for cells with *parS* integrated at positions 1 (A, B), 9 (C, D) and 11 (E, F). The number of cells analysed to calculate the distributions were 1392, 283 and 146 for positions 1, 9 and 11 respectively. The dashed lines in the localisation distributions indicate the birth and division areas of the cells.

5 Discussion

5.1 Validation of the strains with the *parS*/ParB system

The validation experiments with agarose pads resulted in fluorescent foci that could be observed in the analysed cells, as seen in Figure 5A-5P. The results seen in Figure 5 indicate that these foci occur due to the presence of *parS* and ParB-mCherry. Figure 5Q and 5R show that no foci form in the background strain. The fluorescent signal is instead present in the whole volume of the cell. Since *parS* is absent in these strains, the expressed ParB-mCherry become distributed in the cytosol. The distribution of the fluorescent proteins results in a fluorescent signal from the whole cell, instead of distinct fluorescent foci.

Some variation was observed in the intensity of the signal seen in the strains with *parS*, which could be due to variations in the expression of ParB-mCherry. While the fluorescent protein fusion was expressed with the *lac* promoter, its expression was not induced as it has been previously reported that this can result in difficulties with foci tracking (Nielsen *et al*. 2006). The use of non-induced expression could have contributed to the variation in expression and therefore the fluorescence as well. The results presented in Figure 5 suggest that the *parS*/ParB system functions as expected (Nielsen *et al*. 2006, Stouf *et al*. 2013) in the modified cells, which makes it possible to analyse the localisation of the different parts of the chromosome. However, the variation in expression could make the imaging challenging, as the tracking requires a distinct fluorescent signal.

5.2 Optimization of ParB-mCherry expression

As the expression of ParB-mCherry was not induced in the validation experiments this was also omitted during the initial experiments performed in microfluidic chips. However, in the microfluidic chips the fluorescent signal appeared to be very weak due to a comparably large amount of background signal, as seen in Figure 6. A possible explanation for this is that PDMS is weakly fluorescent (Cesaro-Tadic *et al*. 2004) and could have contributed to the background signal, suggesting that the signal observed in the cells was relatively low. As the expression of ParB-mCherry is regulated by the *lac* promoter, IPTG was used to induce the expression assuming that this would result in an increase of the signal. The experiments performed with the strain with *parS* at position 1 in the microfluidic chips with different concentrations of IPTG show an increasing fluorescent signal, as seen in Figure 6. This result suggests that low expression levels of ParB-mCherry contributed to the low signal intensity. Following these results, 1 mM of IPTG was used in the growth medium when subsequently imaging cells in the chips. Only one of the strains was analysed when optimizing the IPTG concentration due to time constraints. The results were also assumed to be similar for the other strains, as the only difference between them was the position of the *parS* sequence fragment.

The optimization experiment which involved analysing the effect of the RPMI amino acids solution on the fluorescent signal was based on comparing the analysed strain with the *codA-parS-cynR* strain. This strain also has a *parS* sequence on the chromosome and the pMS11 from which ParB-mCherry is expressed. When analysed previously by the Elf research group in microfluidic chips the signal was higher than what was observed in the microfluidic chip experiments with the strain with *parS* at position 9. Both strains were grown in M9 glycerol (0.4 % v/v) medium without the use of induction. One difference in the composition of the medium was the absence of RPMI when growing the *codA-parS-cynR* strain and thus this difference was analysed on agarose pads. Again, the analysis was performed for only one strain due to time constraints and assuming that the results would be similar for the other strains with the *parS* sequence fragment.

As seen in Figure 7, the fluorescent signal appears to be stronger in the absence of RPMI for the strain with *parS* at position 9 and the *codA-parS-cynR* strain. However, the difference is small and could be attributed to variations in gene expression, making it difficult to determine the effect of the amino acids solution. While the fluorescent signal for both strains appears to be relatively strong on the agarose pads, it could be difficult to analyse it in the microfluidic chips due to the background fluorescence of PDMS. The difference in fluorescent signal intensity between the strains could be more apparent if they would have been analysed in microfluidic chips. A more thorough comparative analysis of these strains in microfluidic chips is required to understand why the large difference in growth conditions is needed to obtain a similar fluorescent signal.

The *codA-parS-cynR* strain had previously been used as a recipient strain in a P1 phage transduction experiment (Thomason *et al*. 2007) where a sequence was transduced in the intergenic region between *codA* and *cynR*. The infected strain from which the sequence was transduced was an *E. coli* TB28, an MG1655 variant which does not have the *lac* operon (Berhardt & De Boer 2004). Since *codA* and *cynR* are close to the *lac* operon, it is possible that the *lac* operon of the recipient *codA-parS-cynR* strain was affected by the transduced sequence from the TB28 strain during the recombination and parts of it were removed (Thomason *et al*. 2007). If the *codA-parS-cynR* strain does not have the *lac* operon, the expression of ParB-mCherry from pMS11 would not be repressed, which should result in a strong fluorescent signal even in the absence of an inducer. The possible absence of the *lac* operon in the *codA-parS-cynR* strain could have contributed to the different fluorescent signal intensities observed when comparing it with the strain with *parS* at position 9 on agarose pads.

A possible explanation for the large difference between the results observed when analysing the strain with *parS* at position 9 in agarose pads and microfluidic chips could also be attributed to oxygen availability. While oxygen has been implied in causing photobleaching of the fluorescent proteins (Lichtmann & Conchello 2005), in the case of the fluorescent protein mCherry it has also been shown to be necessary for the maturation process of the protein (Hebisch *et al*. 2013). In order to fluoresce the maturation of mCherry requires two

oxidation steps (Hebisch *et al*. 2013). The oxygen availability could be higher in the agarose pads than in the microfluidic chips, which are closed systems in which air is initially displaced by the applied medium. Considering this difference between the pads and chips it is possible that lack of oxygen could contribute to the observed difference. However, this does not explain why a similar difference has not been observed for the *codA-parS-cynR* strain previously.

5.3 The distribution of *parS* over the cell cycle

Figure 9A, 9C and 9E show the distributions of ParB-mCherry bound to *parS* at position 1, 9 and 11 respectively. The similarity in the corresponding average growth rate distributions in Figure 9B, 9D and 9E can be expected since the strains were grown in similar conditions in the microfluidic chips. A comparison of the *parS* localisation during cell division would have been more difficult if the growth rates would differ significantly. The growth rate affects the replication time and overlapping replication cycles could occur at high growth rates. Differences in growth rates could thus result in different numbers of replication cycles occurring in the bacteria, making the comparison between them difficult. Also, since the replication time would be shorter for strains that grow relatively fast the loci would segregate faster because of this. The growth rate distributions for the analysed strains suggest that the introduced sequences do not disturb the growth of the bacteria. Thus, the *parS* localisation results can be used to interpret how the loci are organised during the cell division process.

As seen in the illustration in Figure 1, position 1 and 9 are close to the terminus, while position 11 is closer to the origin. The appearance of two foci and their separation appears to occur later for position 1 and 9 than for position 11. This result is expected based on the sequential segregation of the chromosome after replication (Nielsen *et al.* 2006, Reyes-Lamonthe *et al.* 2008). Loci that are close to the origin become replicated before those that are further away from the origin. These origin proximal loci also become segregated before more distant loci. Since position 11 is close to the origin it is replicated and segregated before the loci that are terminus proximal, such as position 1 and 9. In accordance with this, following the birth of a cell, indicated in Figure 9A, 9C and 9E with a dashed line, the fluorescent signal at position 11 becomes separated into two parts at a smaller cell area than the corresponding signal for position 1 and 9.

It appears that positions 1 and 9 become localized at different positions along the long axis of the cell. This result can be seen on the x-axis of the distributions of the fluorescent foci following division in Figure 9. While there are approximately 600 kb between the two positions, see Table 1 for reference, the results in Figure 9 suggest that these loci are spatially far apart. Also, the loci appear to become segregated at different times following birth, indicated by the separation of the fluorescent signal in Figure 9C and 9E. While the loci are relatively close, it is possible that the point at which they are replicated and segregated differs significantly. However, the number of data points for the distributions of position 1 is larger than for position 9, 1392 and 283 respectively. Analysing the dynamics at position 9 with

more data points would provide stronger support for the comparison with the dynamics at position 1.

It is worth noting that the algorithms used for the generation of the plots showing the distributions of *parS* localisation over the cell cycle and the growth rates of the cells use certain criteria regarding the cells included in the analysis. The cell tracks used by the algorithm need to identify a mother cell and at least two daughter cells. These mother and daughter cells also need to be identified in five consecutive frames. If these criteria are not fulfilled, the cell track will not be included in the analysis. The results presented in Figure 9 are thus based on cell tracks that could fulfil these criteria, meaning that many cells in the images were filtered and not included. While a large number of cells can be captured in the microfluidic chips, all of them will not be analysed. The filtering is necessary to use high quality cell tracks, but it can result in only a subset of the cells contributing to the analysis. The results from the analysis could thus be biased towards the dynamics observed in a small number of cells. In order to achieve higher statistical significance more cells can be included in the analysis by imaging more traps in the microfluidic chip as well as ensuring that the acquired images contain cells that can easily be tracked.

The results presented in Figure 9 show that the dynamics of the loci at which *parS* has been introduced can be analysed over the cell cycle using the described methodology. A similar approach can thus be applied to a large-scale study to analyse the chromosome organisation with higher resolution. This can make it possible to study the dynamics of the whole chromosome and analyse if the organisation is the same between generations. Certain considerations need to be taken into account when using the presented methodology, with regards to the expression of ParB and the growth conditions used for the analysis.

5.4 Further considerations regarding the *parS*/ParB strains

Certain considerations should be noted with regards to the *parS*/ParB system. The SacB coding sequence can be used in the removal of the kanamycin resistance cassette and the coding sequence, which were introduced on the chromosome with *parS*, by selecting for colonies that lack sucrose-sensitivity. This would allow for the analysis of the chromosome dynamics with only *parS* present. While a similar chromosomal integration with *parS* and a selectable marker has been used before (Nielsen *et al*. 2006), it is possible that the introduced sequences affect the dynamics. The resistance cassette and SacB coding sequence are approximately 2.5 kb long in total and removing a sequence this long could affect the organisation of the chromosome at the locus where it has been introduced. This could lead to different results being observed.

ParB has been suggested to bind at *parS* but also at sequences adjacent to it. This property is used to achieve a strong fluorescent signal. However, since adjacent sequences are affected as well, this mechanism could affect the expression of certain genes, depending on where *parS* is introduced. It has been stated that when using the *parS*/ParB system, over-expression of ParB

should be avoided as it could interfere with replication (Reyes-Lamonthe *et al*. 2008). This stands in contrast to the induction that was required to achieve a sufficiently high fluorescent signal in this project. While the expression of ParB was induced, the distributions seen in Figure 9 show that cells could replicate successfully. The effect of ParB expression on replication and *parS* proximal gene expression should be investigated further to assess the possible phenotypic change that it could cause. It is also possible that the organisation of the chromosome at different loci could affect the fluorescent signal intensity. Certain loci could be part of a domain which is not easily accessible for the binding of ParB, resulting in a lower intensity at such loci. The analysis would then be biased towards loci that are accessible. The fluorescent signal intensity at different loci where *parS* has been introduced should be measured to determine the impact that this could have on the analysis.

6 Acknowledgements

I would like to thank Daniel Camsund for his work as the sub-supervisor of my project. Thank you for all of the help during the cloning phase of the project and insightful comments on the report, they were very helpful. I want to thank my supervisor Johan Elf for giving me the opportunity to do this project in the Elf research group and for all the valuable feedback during the project. I also want to thank my subject reader Disa Larsson Hammarlöf for all of the support and great feedback on the report, I appreciate it a lot. Many thanks to David Fange as well, for the all the help with the imaging, microscopy and image analysis as well as suggestions on how to interpret the results of the analysis. Thank you to everyone in the Elf research group for making me feel welcome in the group. I also want to thank my friends and family who have been extremely supportive during the project.

References

Badrinanyanan A, Le TBK, Laub, MT. 2005. Bacterial chromosome organization and segregation. Annual Review of Cell and Developmental Biology. 31: 171-199.

Bates D, Kleckner N. 2005. Chromosome and Replisome Dynamics in *E. coli*: Loss of Sister Cohesion Triggers Global Chromosome Movement and Mediates Chromosome Segregation. Cell. 121: 899-911.

Baltekin Ö, Boucharin A, Tano E, Andersson DI, Elf J. 2017. Antibiotic susceptibility testing in less than 30 minutes using direct-cell imaging. Proceedings of the National Academy of Sciences of the United States of America. 114: 9170-9175.

Berhardt TG, De Boer PAJ. 2004. Screening of synthetic lethal mutants in *Escherichia coli* and identification of EnvC (YibP) as a periplasmic septal ring factor with murein hydrolase activity. Molecular Microbiology. 49: 1255-1269.

Cesaro-Tadic S, Dernick G, Juncker D, Burrman G, Kropshofer H, Michel B, Fattinger C, Delamarche E. 2004. High-Sensitivity Miniaturized Immunoassays for Tumor Necrosis Factor α using Microfluidic Systems. Lab on a Chip. 4: 563-569.

Chen YJ, Liu P, Nielsen AKA, Brophy JAN, Clancy K, Peterson T, Voigt CA. 2013. Characterization of 582 natural and synthetic terminators and quantification of their design constraints. Natura methods. 10: 659-664.

Cooper S, Helmstetter CE. 1968. Chromosome replication and the division cycle of *Escherichia coli* B/r. Journal of Molecular Biology. 31: 519-540.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences of the United States of America. 97: 6640-6645.

Datta S, Constantino N, Court DL. 2006. A set of recombineering plasmids for gram-negative bacteria. Gene. 379: 109-115.

Duffy DC, McDonald JC, Schueller OJA, Whitesides GM. Rapid Prototyping of Microfluidic Systems in Poly(dimethylsiloxane). 1998. Analytical Chemistry. 70: 4974-4984.

Edelstein AD, Amodaj N, Hoover KH, Vale RD, Stuurman N. 2011. Computer Control of Microscopes using μManager. Current Protocols in Molecular Biology. 92: 14.20.1-14.20.17.

Engler C, Marillonnet Sylvestre. 2013. Combinatorial DNA Assembly Using Golden Gate Cloning. Synthetic Biology. 1073: 141-156.

Funnel B. 2016. ParB Partition Proteins: Complex Formation and Spreading at Bacterial and Plasmid Centromeres. Frontiers in Molecular Biosciences. 3: 44.

Fisher JK, Bourniquel A, Witz G, Weiner B, Prentiss M, Kleckner N. 2013. Four-Dimensional Imaging of *E. coli* Nucleoid Organization and Dynamics in Living Cells. Cell. 153: 882-895.

Hebisch E, Knebel J, Landsberg J, Frey E, Leisner M. 2013. High Variation of Fluorescence Protein Maturation Times in Closely Related *Escherichia coli* strains. 8: e75991.

Kleckner N, Fisher JK, Stouf M, White MA, Bates D, Witz G. 2015. The Bacterial Nucleoid: Nature, Dynamics and Sister Segregation. Current Opinion in Microbiology. 22: 127-137.

Reyes-Lamonthe R, Nicolas E, Sherratt D. 2012. Chromosome Replication and Segregation in Bacteria. Annual Review of Genetics. 46: 121-143.

Reyes-Lamonthe R, Wang X, Sherratt D. 2008. *Escherichia coli* and its chromosome. Trends in Microbiology. 16: 238-245.

Lawson MJ, Camsund D, Larsson J, Baltekin Ö, Fange D, Elf J. 2017. In situ genotyping of a pooled strain library after characterizing complex phenotypes. Molecular Systems Biology. 13: 947.

Lemon KP, Grossman AD. 1998. Localization of Bacterial DNA Polymerase: Evidence for a Factory Model of Replication. Science. 282: 1516-1519.

Lemon KP, Grossman AD. 2001. The extrusion-capture model for chromosome partitioning in bacteria. Genes & Development. 15: 2031-2041.

Lichtman JW, Conchello JA. 2005. Fluorescence Microscopy. Natura Methods. 2: 910-919.

Magnusson KEG, Jaldén J, Penney MG, Blau HM. Global Linking of Cell tracks using the Viterbi Algorithm. Transacations on Medical Imaging. 34: 911-929.

Nielsen HJ, Ottesen JR, Youngren B, Austin SJ, Hansen FG. 2006. The *Escherichia coli* chromosome is organized with left and right chromosome arms in separate cell halves. Molecular Microbiology. 62: 331-338.

Olivo-Marin J. 2002. Extraction of spots in biological images using multiscale products. Pattern recognition 35: 1989-1996.

Reece A, Xia B, Jiang Z, Noren B, McBride R, Oakey J. 2016. Microfluidic techniques for high throughput single cell analysis. Current Opinion in Biotechnology. 40: 90-96.

Sadanandan SK, Baltekin Ö, Magnusson KEG, Boucharin A, Ranefall P, Jaldén J, Elf J, Wählby J. 2014. Segmentation and Track-Analysis in Time-Lapse Imaging of Bacteria. Journal of Selected Topics in Signal Processing. 10: 174-184.

Sanderson JB. 2002. Phase Contrast Microscopy. Encyclopedia of Life Sciences. doi: 10.1038/npg.els.002635.

Stouf M, Meile JC, Cornet F. 2013. FtsK actively segregates sister chromosomes in *E. coli*. Proceedings of the National Academy of Sciences of the United States of America. 110: 11157-11162.

Thomason LC, Constantino N, Court DL. *E. coli* Genome Manipulation by P1 Transduction. Current Protocols in Molecular Biology. 1.17.1-1.17.8.

Toro E, Shapiro L. 2010. Bacterial chromosome organization and segregation. Cold Spring Harbor Perspectives in Biology. 2: a000349.

Wallden M, Fange D, Lundius E, Baltekin Ö, Elf J. 2016. The Synchronization of Replication and Division Cycles in Individual *E. coli* cells. Cell. 166: 729-739.

Wang X, Llopis PM, Rudner DZ. 2013. Organization and segregation of bacterial chromosomes. Nature Reviews Genetics. 14: 191-203.

Yanisch-Perron C, Vieira J, Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. Gene. 33: 103-119.

Zerbini F, Zanella I, Fraccascia D, König E, Irene C, Frattini LF, Tomasi M, Fantappiè L, Ganfini L, Caproni E, Parri M, Grandi A, Grandi G. 2017. Large scale validation of an efficient CRISPR/Cas-based multi gene editing protocol in *Escherichia coli*. Microbial Cell Factories. 16: 68.

Appendix

Table 1 Description of the chromosomal position of the 13 sites at which the *parS* sequence fragment with a kanamycin cassette and SacB coding sequence was integrated.

Position	Chromosomal locus position	Description
1	2.47 Mbp	Removal of the pseudogene *intC*
2	2.81 Mbp	Removal of the pseudogene *ygaY*
3	0.57 Mbp	Integration into the pseudogene *nmpC*
4	1.58 Mbp	Integration in the intergenic region between *ydeP* and *ydeQ*
5	2.56 Mbp	Integration in the intergenic region between *yffS* and *eutA*
6	0.8 Mbp	Integration in the intergenic region between *ybhJ* and *ybhC*
7	1.44 Mbp	Integration in the intergenic region between *ydbL* and *feaR*
8	1.70 Mbp	Integration in the intergenic region between *ydgJ* and *ydgT*
9	1.84 Mbp	Integration in the intergenic region between *gdhA* and *ynjL*
10	1.24 Mbp	Integration in the intergenic region between *ycgY* and *treA*
11	3.90 Mbp	Removal of the nonessential gene *bglH*
12	0.78 Mbp	Removal of the nonessential gene *galK*
13	1.32 Mbp	Removal of the nonessential gene *cobA*

Table 2 Oligonucleotides used for the introduction of BpiI restriction sites used in the Golden gate assembly.

Description	Sequence (5' – 3')
parS-GG-forward	GGAGAAGACAAGTCAGCGAAATTATGAGTCACGAAGAGG
parS-GG-reverse	TCCGAAGACTTCTATAGGATGCCGAAGAGCATCTT
KanR-SacB-GG-forward	GGAGAAGACATATAGTGTAGGCTGGAGCTGCTTC
KanR-SacB-GG-reverse	TCCGAAGACGCATAAGGACCAAAACGAAAAAACACC CTTTCGGGTGTCTTTTCTGGAATTTGGTACCGAG TTATTTGTTAACTGTTAATTGTCCTTGTTC

pGuide8-GG-forward	GGAGAAGACGGTTATCGCTCCTTTCGCTTTCTTCC
pGuide8-GG-reverse	TCCGAAGACGCTGACTAAAACTGGATGGCTTTCTTGCCG

Table 3 Oligonucleotides used for confirmation of the pParS plasmid being successfully assembled. The oligonucleotides were also used for sequencing of the fragments amplified from the plasmid.

Description	**Sequence (5' – 3')**
Seq-pGuide8 forward	ACGGAAATGTTGAATACTCATGC
Seq pParS reverse	GCCTTTTTACGGTTCCTGGCC

Table 4 Oligonucleotides used for the introduction of homologous ends to the *parS* sequence fragment with a kanamycin resistance cassette and a SacB coding sequence.

Description	**Sequence (5' – 3')**
intC forward	AAAACTTTGGTCGCAATGAGCGATACGATACTTCCTGAAA GCGAAATTATGAGTCACGAAGAGG
intC reverse	GCAGACCAGATCCCGGCATTCAACAAAGCACTGGCAACAT GGACCAAAACGAAAAAACACCC
ygaY forward	TCCTCGGTACGGCATTAACCGGTTTATTCTCAGTCGTGGC GCGAAATTATGAGTCACGAAGAGG
ygaY reverse	ACCGCCAATAAAGTAGCTAGTCATGTAACCTGCGGTCAGG GGACCAAAACGAAAAAACACCC
nmpC forward	AAGGTGAAACCCAAATCAACGATCAACTGACTGGTTTCGG GCGAAATTATGAGTCACGAAGAGG
nmpC reverse	TCTGGCAGGACGTCAGTCCACGCACCGATGTCGTATGCTA GGACCAAAACGAAAAAACACCC
ydeP-ydeQ forward	AAGACTCTGGCTTCAATTGTGCGCGGATTTTCTTACAGGT GCGAAATTATGAGTCACGAAGAGG
ydeP-ydeQ reverse	AATTGCGCGGATAATTATTTTGTGAAGGCTATTAGCCTAC GGACCAAAACGAAAAAACACCC
yffS-eutA forward	CGCCCGGCCCCGCTGCCGGGTTTTTGCTATGCACCACAAT GCGAAATTATGAGTCACGAAGAGG

yffS-eutA reverse	GCTGGTGGAATTGTATCCGGCTGTGTATCCGGTTGGGGTA GGACCAAAACGAAAAAACACCC
ybhJ-ybhC forward	CTACGAATCGCATCGAATCTGTAGGCCAGATAAGGCATTT GCGAAATTATGAGTCACGAAGAGG
ybhJ-ybhC reverse	GCGCAAGCATCGCATCCGACAATAAGTGCCGGATGCTGCGA GGACCAAAACGAAAAAACACCC
ydbL-feaR forward	GCGATTGATCTATTTTCCTGAAACAAGGTGAATATTCAAA GCGAAATTATGAGTCACGAAGAGG
ydbL-feaR reverse	CTATGCATTTTTCAGGGCAAAAGGCAATTTGACAGGAGTT GGACCAAAACGAAAAAACACCC
ydgJ-ydgT forward	TAAATATCGCAAAAACCTCAGTAAAAATCTTGCTGGAGCT GCGAAATTATGAGTCACGAAGAGG
ydgJ-ydgT reverse	ATCCATTAACTTCAGGGGGTAAATGTTACTTAGCAATAAT GGACCAAAACGAAAAAACACCC
gdhA-ynjL forward	TGTAAATGCCTGATGGCGCTACGCTTATCAGGCCTACAAA GCGAAATTATGAGTCACGAAGAGG
gdhA-ynjL reverse	CCGGCCTACATTAGAGCGTAACTGCAATGAATTGTGCCCA GGACCAAAACGAAAAAACACCC
ycgY-treA forward	AAAAAACCATAGCTATGGTGTTATCTGCCGACAACGCCAT GCGAAATTATGAGTCACGAAGAGG
ycgY-treA reverse	AACCAGCGCTTACTCCGACAACGTCATTCGGCTTCTCCGG GGACCAAAACGAAAAAACACCC
bglH forward	CACCAGACAAATCCCAATAACTTAATTATTGGGATTTGTT GCGAAATTATGAGTCACGAAGAGG
bglH reverse	CGCATCGCATCCAGACTGTTCTGAATGCGACGATAATTAA GGACCAAAACGAAAAAACACCC
galK forward	AGTCAGCGATATCCATTTTCGCGAATCCGGAGTGTAAGAA GCGAAATTATGAGTCACGAAGAGG
galK reverse	ACGCAAAAAGCCCCGAGCGGTTAAACTCAGGGCTTTATTTTTA GGACCAAAACGAAAAAACACCC
cobA forward	ACGCCCAACCGGGCCGTAAACCAGGAATTTCCCAATGAGT GCGAAATTATGAGTCACGAAGAGG

cobA reverse	GTATGCCGGACGTCATATCCGGCATTTTTACAGATTAATA GGACCAAAACGAAAAAACACCC

Table 5 Oligonucleotides used for the confirmation of successful integration of the *parS* sequence with a kanamycin resistance cassette and a SacB coding sequence. The oligonucleotides were also used for sequencing of the fragments amplified from the plasmid. The *parS* confirmation reverse oligonucleotide, which binds in the kanamycin resistance cassette was used for sequencing to confirm that the *parS* sequence specifically had been successfully integrated.

Description	**Sequence (5' – 3')**
intC conf forward	AAGCTCTTTCGTTCTTAGGGC
intC conf reverse	TACCGCAAGAAGAACTTCCC
ygaY conf forward	TACTGTTTCTTGTTCCCCTCGGTG
ygaY conf reverse	AAAACCTCGCCACCAGACCAG
nmpC conf forward	TGATGGTGATACTACTTATGCCC
nmpC conf reverse	TCCAAGTGTCACCACCGAAT
ydeP-ydeQ conf forward	TGTGAAGGCTATTAGCCTAC
ydeP-ydeQ conf reverse	TCTATGCGCTCCATTCACGA
yffS-eutA conf forward	GCCTTCAACCCCATCTCA
yffS-eutA conf reverse	TGAATGCGGTATTCGCCA
ybhJ-ybhC conf forward	TGTGAATGTAGGTCGCATTCGG
ybhJ-ybhC conf reverse	CGGTTCGATGCGATTTGTAG
ydbL-feaR conf forward	CCGCAATATCGGGATAACGC
ydbL-feaR conf reverse	CTGCAACGCCAATTAGTTAAAAC
ydgJ-ydgT conf forward	CATTAACTTCAGGGGGTAAATGTTAC
ydgJ-ydgT conf reverse	ATCGCAAAAACCTCAGTAAAAATCT
gdhA-ynjL conf forward	TGGCGCAGGGTGTGATTTA
gdhA-ynjL conf reverse	GCCTGAAATTTTGCCGGG
ycgY-treA conf forward	CGCCAGGACTTGAAAAACCG

ycgY-treA conf reverse	CTCAACCAAAGAGGCACAAC
bglH conf forward	TGGCGATGAGCTGGATAAA
bglH conf reverse	AGATGGCAGAGGTAATAAGA
galK conf forward	GTTATGAAATGCTGGCAGAG
galK conf reverse	GAAAGTAAAGTCGCACCC
cobA conf forward	CTGATATCATGCCGCTCTAC
cobA conf reverse	CTGGCGGACGTAATAATAAC
parS confirmation reverse	GTGAGATGACAGGAGATCCT

Table 6 A Summary of all strains used, excluding those that were developed during the project and commercial strains.

Strain	**Description**
BW25993	A wildtype BW25993 *E. coli* strain, based on the MG1655 strain (Datsenko &Wanner, 2000).
KanR-SacB	A modified DH5α *E. coli* strain with a kanamycin resistance cassette and SacB coding sequence on the chromosome.
codA-parS-cynR	A modified BW25993 *E. coli* strain with the *parS* sequence in the intergenic region between *codA* and *cynR*.

Table 7 A summary of all plasmids used, describing their use, copy number and temperature-sensitivity. Selectable markers are specified for all plasmids.

Plasmid	**Description**	**Selection marker**
pParS	Used as a template for the introduction of homologous end sequences onto a DNA sequence fragment containing *parS*, a kanamycin resistance cassette and a SacB coding sequence. It is a high copy number plasmid.	Kanamycin and ampicillin resistance, sucrose sensitivity.
pGuide8	Used as a template for the amplification of the pUC19 origin of replication and ampicillin	Ampicillin resistance.

	resistance cassette for introduction of end sequences used in Golden gate assembly.	
pSIM6	Contains the coding sequences for Beta, Exo and Gam, which are used in λ-Red recombination. Their expression is repressed at 30 °C and induced at 42 °C. It is a low copy number plasmid with a temperature-sensitive origin of replication (Datta *et al.* 2006).	Ampicillin resistance.
pMS11	Contains the coding sequence for ParB, with the expression being controlled by the *lac* promoter. It is a low copy number plasmid.	Chloramphenicol resistance.