RecA mediated homology search finds segregated sister locus in minutes after a double stranded break

Arvid H. Gynnå*, Jakub Wiktor*, Prune Leroy & Johan Elf
Department of Cell and Molecular Biology, SciLifeLab, Uppsala University, Sweden
*equal contribution

Abstract
The search for a homologous template is a central and critical step in the repair of DNA breaks by homologous recombination. However, it is still unclear how the search process is carried out within a cell. Here, we image double-stranded break (DSB) repair in living *E. coli* growing in a microfluidic device and show that two segregated homologous sequences find each other less than 9 min after an induced double-stranded break. To characterize the mechanisms of the search process, we use a new RecA fluorescent fusion that rapidly forms structures after DSBs. Initially, RecA forms a bright cluster at the site of the DNA damage and then extends to form a thin, flexible, and dynamic filament. Based on our observations, we propose a model where the 1D ssDNA-RecA filament stretches along the cell, and the homology search is mediated by diffusion of the repair template that can interact with any segment of the filament. This model reduces the complexity of the search from 3D to 2D and quantitatively predicts that genome-wide search for homology can be finished in minutes, in agreement with our measurements.

Introduction

A DNA double-stranded break (DSB) is a potentially catastrophic event for any cell, and evolution has provided several pathways to repair such damage (Stephen C. Kowalczykowski 2015; Bertrand et al. 2019). In *Escherichia coli* these lesions are repaired exclusively by homologous recombination (HR), in which a homologous locus on the sister chromatid is used as the template for accurate DNA repair (Dillingham and Kowalczykowski 2008). HR repair of a DSB is initiated when the RecBCD complex binds to the free DNA ends. First, RecBCD degrades several kilobases of DNA, but upon a recognition of a chi site, an 8 bp DNA motif, it switches to recombinogenic 5' resection activity (Dillingham and Kowalczykowski 2008). After being activated by a chi site, RecBCD creates 3' single-stranded DNA (ssDNA) tails and assists in loading of RecA proteins onto them. Next, the RecA-ssDNA complex, called presynaptic filament, searches for its sister locus using base-pair homology. At the homologous loci, the 3' ssDNA end invades the double stranded DNA and serves as a primer for DNA replication. The genetic information lost during the process of DSB repair can now be restored by copying it from the undamaged template (S. C. Kowalczykowski 2000).

While the majority of DSBs arise at the replication fork where the homologous template is adjacent (Kuzminov 1995; Kowalczykowski 2000; Pennington and Rosenberg 2007), a DSB may also occur after the loci have segregated and are separated by micrometer-scale distances (Mehta and Haber 2014). In the latter case, HR repair has to perform a long-range, genome-wide search for a homologous repair template. Finding the homology among an overwhelming amount of heterologous sequence is a major challenge for the recombination machinery, even within the relatively small *E. coli* genome (Renkawitz, Lademann, and Jentsch 2014). Search driven by diffusion of the presynaptic complex is not suitable to support the necessary timescale of recombination. An example of such a search comes from the CRISPR system that relies on normal diffusion in 3D to locate a short sequence within a double stranded chromosome based on homology to the guide RNA. There, a single CRISPR-Cas complex spends hours searching for the correct site (Jones et al. 2017; Vink et al. 2019). A similar mechanism for HR would be too slow, since a further complication in HR is that the diffusivity of the RecA-ssDNA filament is limited due to its large molecular mass and its tethering to the rest of the chromosome. Considering this challenge, it has been suggested that search and repair may not be possible at all between segregated sequences (Barzel & Kupiec, 2008; Weiner et al., 2009) Still it appears that long range homologous recombination can be performed within one

generation time in *Caulobacter crescentus* (Badrinarayanan, Le, and Laub 2015) and slower in *E. coli* (Lesterlin et al. 2014). Thus half a century after the discovery of RecA (Clark and Margulies 1965), the mechanism of search is still puzzling.

Several explanations have been proposed to explain the mechanism of search carried out by the RecA-ssDNA complex. It has been shown in vitro that the search for homology can be greatly improved by dividing the task into multiple smaller, parallel searches. A single RecA-ssDNA filament can be involved in interactions with different segments of DNA, allowing for probing of multiple, potentially distant DNA sites at once (Forget and Kowalczykowski 2012). Another in vitro experiment showed that RecA filament prefers homologies of 8 bps or longer, binding shorter ones only transiently (Qi et al. 2015). Such a property could reduce the complexity and speed up the search. However, it is still a question whether those features described in vitro play a key role in the search for homology in the context of a living cell.

One of the difficulties in studying the spatial DNA dynamics of the search process in live *E. coli* cells using fluorescent labels is the potent nuclease activity of RecBCD complex which can degrade many kilobases of DNA before switching activity at a chi site. This makes it impossible to put a DNA binding fluorescent markers in close proximity to the DSB site, as they would be ejected during the initial degradation (Wiktor et al. 2017). Moreover, fluorescent protein fusions to RecA tend to decrease or abolish its activity (Renzette et al. 2005). Recently, DNA repair ability has been retained in cells by expressing wild type RecA alongside a fluorescent fusion. Those studies showed that in *E. coli* the labeled RecA can form either relatively thick and stable "bundles" or short-lived spots at the site of damage (Lesterlin et al. 2014; Amarh, White, and Leach 2018). In *C. crescentus* and *Bacillus subtilis* both spots and elongated RecA structures have also been observed (Badrinarayanan, Le, and Laub 2015; Kidane and Graumann 2005).

Here, we first introduce a system to make a single, and therefore repairable, break in one of the segregated sister chromosomes in *E. coli*. We use it to characterize the repair process in many cells growing in a microfluidic device. We track the cell lineages from the time prior to DSB formation until subsequent cell divisions, confirming the successful repair using a segregated sister as a template. Moreover, using a new RecA-mVenus fusion we observe fast dynamics of the homology search, where the process is facilitated by a thin and dynamic RecA structure.

Results

Tightly controlled formation of DSBs in a microfluidic chip

Straight-forward expression of nucleases such as I-SceI or Cas9 can be used to study DNA repair mechanisms (Roukos et al. 2013; Lesterlin et al. 2014; Anand et al. 2017). However, when those nucleases are expressed at low levels in bacteria, cell-to-cell variation in the nuclease concentration leads to cutting of all or no restriction sites, resulting in no repair templates or no breaks (Cui and Bikard 2016; Wiktor et al. 2017). To increase control of the DSB induction, we prevent access to the introduced chromosomal I-SceI restriction site (referred from now on as *CS*) by surrounding it with two lactose operators (*lacO*), one partly overlapping the restriction site (**Fig 1A-B**). This design allows for rapid accessibility switching of the *CS* depending on the presence or absence of the LacI inducer IPTG. We tested the system in cells with a fluorescent SOS-reporter (*PsulA'-CFP,* **Fig S1B**). Importantly, the experiments are run at 37°C, since we find significant non-specific activity of I-SceI at 30°C (**Fig S1C-E**, see Supplementary material). Expression of the SOS-reporter, as well as cell filamentation, was strongly pronounced in the presence of ATC induced expression of I-SceI and IPTG together. On the other hand, DNA damage was strongly reduced when IPTG was omitted (**Fig 1C**), showing that LacI sequesters the *CS* effectively.

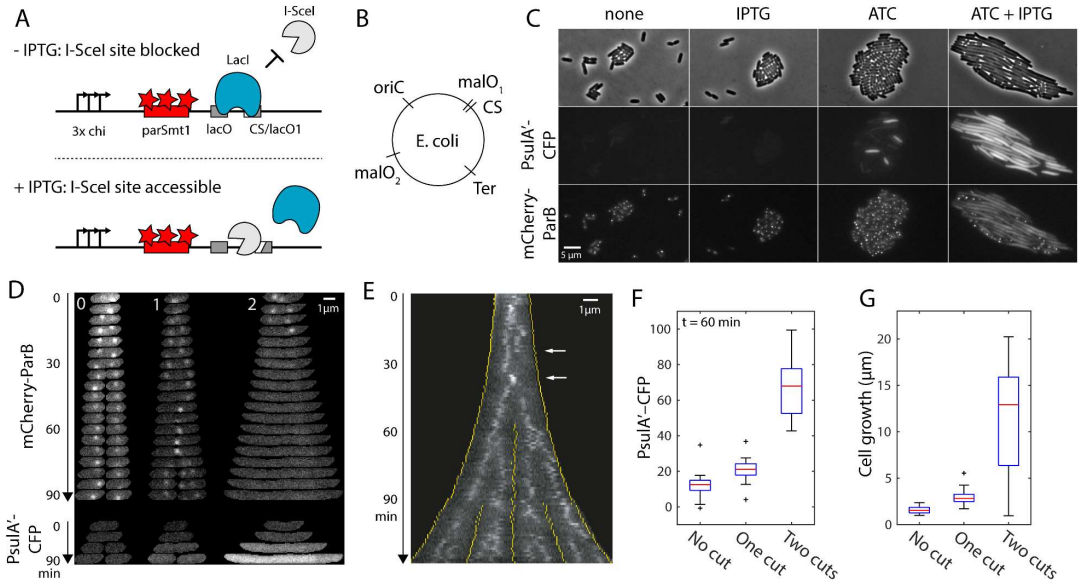

Figure 1. A. *Genetic construct inserted between codA and cynR. Top: With LacI protecting the restriction site, CS. Bottom: Without LacI.* ***B.*** *Approximate genomic positions of the CS construct and the two alternate malO arrays shown on the E. coli chromosome.* ***C.*** *Effect of I-SceI expression with the construct in A, after 3 h incubation with IPTG (1 mM), ATC (0.5 ng/ml) and with ATC & IPTG. ATC induces I-SceI expression, IPTG releases the restriction site (CS). CFP is the SOS response reporter* ***D.*** *Top: mCherry-ParB in cells with 0, 1 and 2 lost mCherry foci. Bottom: SOS-induced CFP expression in the cells above.* ***E.*** *Kymograph of mCherry-ParB in cell undergoing DSB and subsequent repair (indicated by arrows), followed by two cell divisions. Yellow line indicates cell poles.* ***F.*** *CFP fluorescence measured by microscopy 60 min after DSB induction. (n = 20, 20 & 20)* ***G.*** *Cell length extension, from birth until maximum length. (n = 20, 20 & 19)*

Optimizing the level of DSB induction

To determine when a locus was cut by the I-SceI and processed by RecBCD, we integrated a *parS-Mt1* site 300 bp from the *CS*. *ParS-Mt1* is bound by several mCherry-ParB and forms a fluorescent focus (Nielsen et al. 2006). After a DSB is formed, the RecBCD degrades the *parS* site and ejects the mCherry-ParB molecules causing a rapid loss of the focus (Wiktor et al. 2017).

To track bacteria for several generations after the break induction, we use a customized version of the mother-machine microfluidic chip (Baltekin et al. 2017). Automated image analysis of thousands of cells let us observe different outcomes after the DSB formation (Methods). For example, cutting of both chromosomes leads to a strong SOS-response, measured by the *pSulA-CFP* SOS reporter, as well as cell elongation, and eventually inability of a cell to divide (**Fig 1D**).

On the other hand, the cells in which only one of the two *CS* was cut typically survived the DSB, *i.e.* recovered the lost focus and resumed normal cell division (**Fig 1D-E**). A slight induction of the SOS reporter could be measured in these cells **(Fig 1F)** and they also grew to longer length **(Fig 1G)**. To optimize the number of cells undergoing such productive DSB repair, we adjusted the ATC and IPTG induction levels until we achieved a 29 % probability of cutting one out of two sites and a 9% probability of cutting both sites. Since $0.29 \cdot 0.29 = 0.084$, these frequencies suggest that the sites are cut independently.

Rapid pairing between distant homologies

Having the ability to accurately induce DSB in a single chromosome, we needed a method to simultaneously visualize the broken site and its repair template on the sister chromosome. We integrated an array of 12 *malO* sites at 25 kb distance from the *CS* (**Fig 1B**). The array can be visualized by binding of constitutively expressed MalI-mVenus. This MalI-mVenus focus remains intact during repair as RecBCD resection rarely proceeds that far (Wiktor et al. 2017) and thus makes it possible to study the DNA motion after a DSB.

Cells that displayed an increase of CFP fluorescence after the induction of the break and divided afterwards where further manually inspected for loss of the break-proximal mCherry foci and repair, defined by by the reappearance of segregated mCherry foci. We observed the loss of the proximal foci 10.6 ± 6.4 min (mean ± SD, n = 77) after the end of the induction. The delayed loss of foci is likely caused by the delayed restriction activity of I-SceI after DNA binding (Perrin, Buckle, and Dujon 1993). The experiment was performed in parallel with control strains either with uncleavable *CS* or having an inactive enzyme (dI-SceI). When the cells from the cutting and the control strains were inspected together, 56 cells with probable DSBs were identified when active *CS* and I-SceI are present while no DSBs were identified in either of the control strains, giving us confidence that DSB events could be identified with precision.

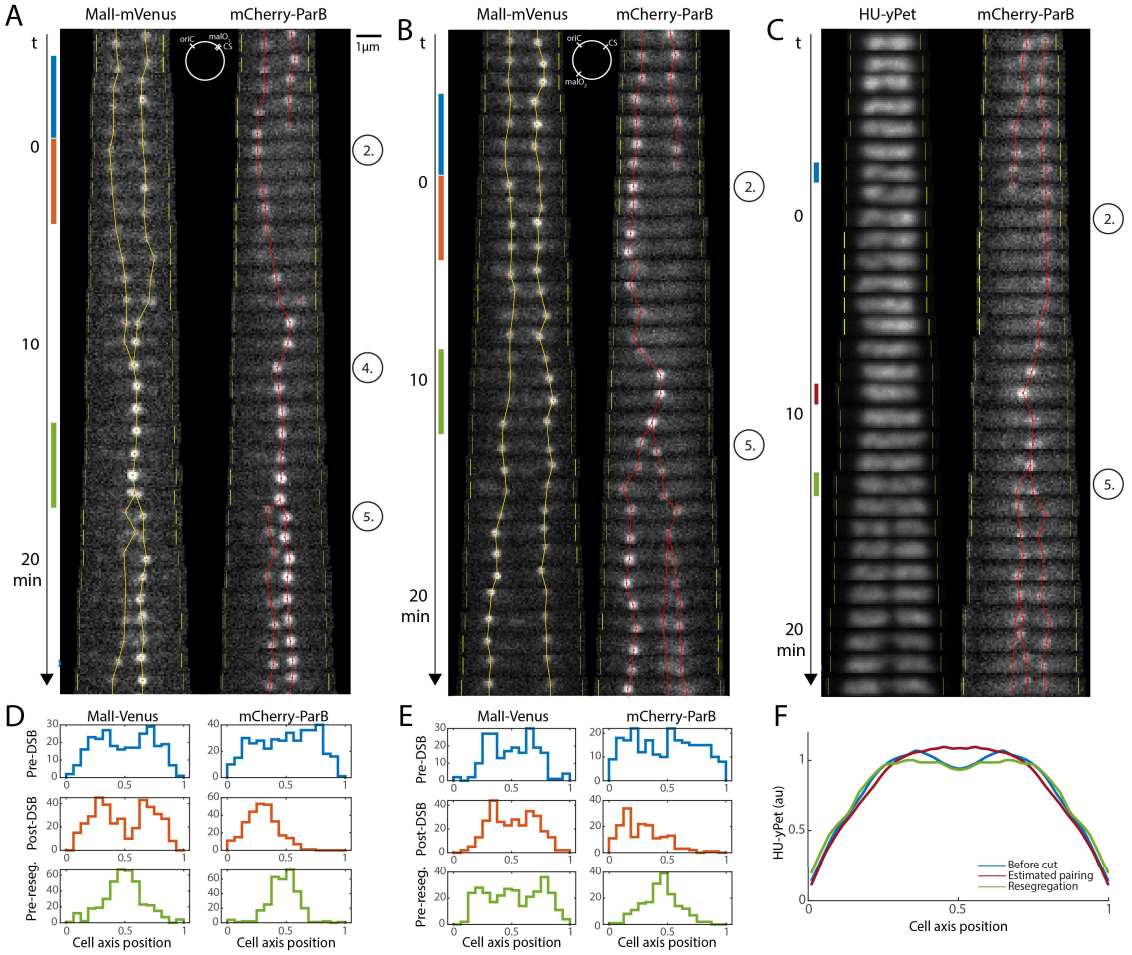

Figure 2: A. *Example cell with a single DSB, displaying MalI-Venus 25 kp from CS, and mCherry-ParB, at CS. mCherry focus loss (2), pairing (4) and resegregation (5) are indicated. Colored bars to the left indicate the time periods used for histograms in D. Cell long-axis boundaries are shown as yellow vertical lines. Inset shows genomic positions of the label binding sites.* ***B.*** *As A, but with MalI-Venus on the opposite chromosomal arm to CS.* ***C.*** *HU-yPet and mCherry-ParB at CS, in cell with one DSB. Colored bars indicate the time points used in F.* ***D.*** *Long-axis distributions of MalI-Venus, 25 kb from CS, and mCherry-ParB during 4 min time windows before the DSB, after the DSB and before mCherry resegregation. (n=77, 77 & 73 cells)* ***E.*** *As D, but with MalI-Venus on the opposite chromosomal arm to CS. (n=42, 42 & 40 cells)* ***F.*** *Long-axis mean HU-yPet density, at the time of a DSB, the estimated time of pairing and at the time of mCherry resegregation, in cells like C. (n=4 cells)*

In the bacteria that lost only one out of two mCherry-ParB foci, we observed that the both MalI-mVenus foci, as well as the mCherry-ParB focus at the uncut *CS*, moved simultaneously towards the center of the cell and merged 8.3 ± 4.4 min (n = 55) after the loss of mCherry focus at the DSB site (**Fig 2A, D**). We observed these pairing events even when MalI-mVenus foci were separated by more than 2 µm beforehand. After the pairing, the break-proximal mCherry-ParB and the 25 kb distant MalI-mVenus foci remained at

the mid-cell for 8.4 ± 6.4 (n = 51) and 8.5 ± 6.5 (n = 48) minutes, respectively, before resegregating into the opposite cell lobes. Reappearance and resegregation of the ParB-mCherry focus, followed by cell division, show that the DSB damage was repaired (**Fig 1E**). During the paired phase, 39 % (n = 77) of the mCherry-ParB foci split temporarily into two foci during 1-3 minutes before merging again (**Fig S2**), suggesting that the break is repaired at this point or that the homologous template has been replicated by a new round of chromosome replication, but that segregation is hindered. In conclusion, we found that the region around the homologous sister locus of a DSB moves to the middle within minutes after a break. However, it was still unclear whether the homology search happens before or after the pairing of the labeled loci and how much of the chromosome that is involved in the pairing.

Pairing between sister chromosomes is limited to the DSB site and its sister locus

To test whether the pairing is specific to sister sites, or is a consequence of large scale chromosomal movements during DNA stress (Odsbu and Skarstad 2014), we moved the *malO* array to the other chromosomal arm, at approximately the same distance from the origin of replication as before. We found that MalI-mVenus foci located on the opposite chromosomal arm to the DSB do not pair with their sister loci during the repair, but follow a typical pattern of undamaged foci (**Fig 2B, E**). We also used a HU-yPet fusion, which abundantly binds to DNA and allows us to visualize the general nucleoid organisation (Wery, Woldringh, and Rouviere-Yaniv 2001; Wu et al. 2015). The longitudinal distribution of HU-yPet following a DSB (**Fig 2C, F**) shows that although the separation between the chromosome lobes transiently disappears, the overall spatial distribution of the chromosome does not change drastically during the repair. Furthermore, we did not observe significant nucleoid compaction, as it was reported for UV treated cells (Odsbu and Skarstad 2014). This data show that the chromosomal movements during the repair of a single DSB is limited only to the region proximal to the break and its homologous sequence. The selective movement of the homologous sequence implies that it is identified by the HR search mechanism before the pairing (**Fig 2A**), i.e. within 9 min.

RecA filaments mediate pairing between segregated homologies

The protein RecA has been implicated in the search phase of HR in a number of imaging studies (Kidane and Graumann 2005; Lesterlin et al. 2014; Badrinarayanan et al. 2015). To validate the timing and localizations of the events we observed with chromosomal markers, we visualised RecA dynamics under our experimental conditions. We constructed a new RecA-mVenus fusion which was expressed directly after the wild type *recA* gene, identically to the tandem *recA-mCherry* construct from (Amarh, White, and Leach 2018). Importantly, the RecA-mVenus fusion did not form persistent structures during unstressed exponential growth as has been observed for other fusions (Renzette et al. 2005), however we did observe occasional short-lived RecA foci and elongated structures even without active I-SceI present (**Fig S3A**). 13% of cells had RecA structures that lasted more than one minute at some point during their lifetime (**Fig S3B**). We found that cells with RecA structures lasting several minutes also expressed more RecA-mVenus, indicating SOS response (**Fig S3C**). These cells also had a longer generation time and divided at larger sizes (**Fig S3D-F**), suggesting that those structures represent spontaneous DNA damage and that RecA-mVenus participates in HR repair.

The *recA-mVenus* fusion was transduced into a strain with the *CS* and the proximal mCherry-ParB label. When DSBs were induced and one of the ParB-mCherry foci disappeared, we found that almost all RecA-mVenus is rapidly attracted to the site of a break, corroborating the interpretation of a ParB-mCherry focus loss as a DSB (**Fig 3A-B**). On average the RecA foci were formed 43 ± 82 s (n = 89) before the loss of ParB-mCherry focus, but an individual focus can be found both before and after the loss of a ParB-mCherry focus. We believe the occasional coexistence of ParB and RecA foci is due to processing and RecA loading on the side of the DSB without the ParB label first.

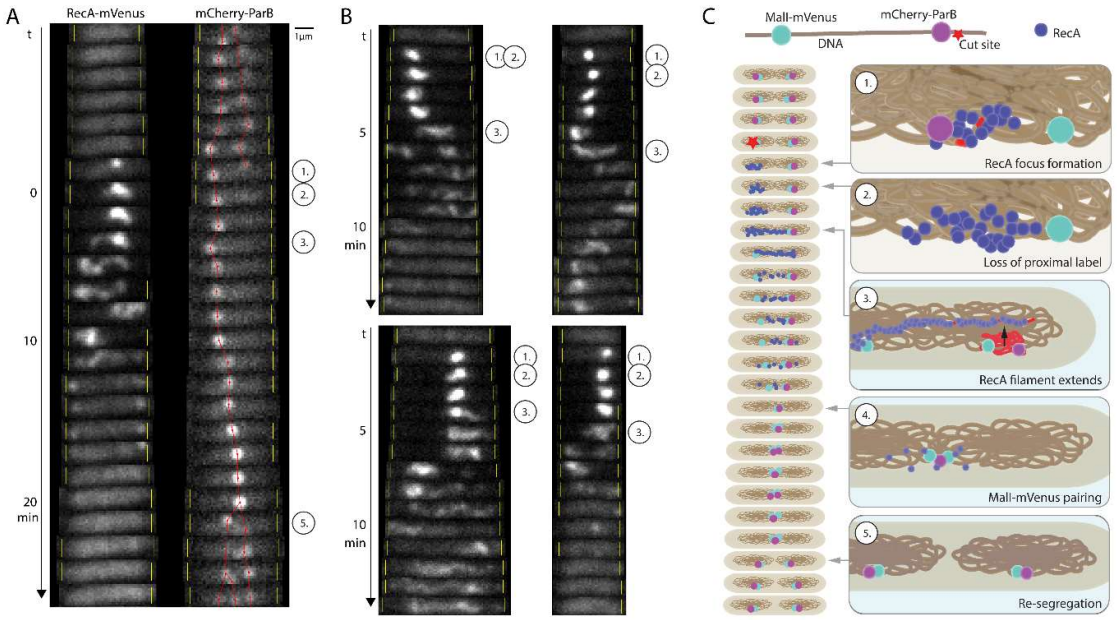

Figure 3: A. *Example cell with one DSB, displaying RecA-mVenus structures and mCherry-ParB foci. RecA focus formation (1), mCherry focus loss (2), RecA filament formation (3) and mCherry resegregation (5) are indicated.* ***B.*** *More examples of RecA-Venus in cells with one DSB. Annotation 1-3 as in A.* ***C.*** *Cartoon showing our interpretation of the fluorescence images observed during and after repair of a DSB.*

A second distinct type of RecA structure could be observed 4.9 ± 2.1 (n = 89) min after the formation of the initial focus. Here, RecA assembled into a dynamic structure, a thin tail protruding from the initial RecA-focus that converts into a filament-like structure extending along the cell length. The elongated structure disappeared in cells that successfully repaired the DNA, typically when the sister ParB-mCherry focus had migrated to mid-cell. In contrast, in the cells where all copies of *CS* were cut, the RecA structures could be observed for more than 30 minutes. We also attempted the same experiment in cells without RecN, a SMC (structural maintenance of chromosomes)-like protein whose role in HR repair is not yet clear, however experimental evidence suggests it acts in an early stage (Meddows et al. 2005; Reyes et al. 2010). When *recN* was deleted, the initial RecA focus and the elongated structure still formed but did not lead to mCherry-ParB pairing and subsequent productive repair in the majority of cases **(Fig S3A)**. Interestingly, a minority of ⊿*recN* cells did repair and returned to cell division, but only after a long delay **(Fig S3B)**. To exclude the possibility that the RecA-mVenus structures observed were related to the nuclease I-SceI, we also generated breaks using Cas9. This resulted in RecA structures indistinguishable from when I-SceI was used, and repair finishing on the same timescale (**Fig S4C**).

The elongated RecA structures observed here are highly dynamic, changing their localization between one-minute frames (**Fig 3A-B**). When the measured timings were summarized (**Table 1**), it was evident that the total repair time, from mCherry foci loss until re-segregation, is similar regardless of whether labeled RecA-mVenus is present or not. The similar time scale suggests that our fusion protein does not have a strong adverse effect on HR search and repair. All together, these results show that a dynamic RecA filament is present during the search phase of HR and dissolves soon after.

Table 1: Measured time intervals between DSB related events, mean ± SD (min).

	MalI-mVenus & mCherry-ParB					RecA-mVenus & mCherry-ParB					
0. Induction											**0.**
1. RecA-mVenus focus	16.6 ± 6.4 n = 77					16.9 ± 6.3 n = 89					**1.**
2. mCherry-ParB focus loss							0.73 ± 1.37 n = 89				**2.**
3. RecA-mVenus structure		8.3 ± 4.4 n = 55			16.2 ± 6.8 n = 73			4.9 ± 2.1 n = 89		18.3 ± 6.1 n = 84	**3.**
4. MalI-mVenus pairing			8.4 ± 6.4 n = 51						13.3 ± 6.3 n = 84		**4.**
5a. mCherry-ParB segregation				8.5 ± 6.5 n = 48							**5a.**
5b. MalI-mVenus segregation											**5b.**

Discussion

Our results give detailed insights into DNA repair by homologous recombination and homology search process in living cells. We observed repair events between segregated sisters lasting shorter than a generation time and involving relocation of the region around both the damaged loci and the repair homologies into mid-cell position. The RecA-mVenus fusion constructed for this study showed surprising behavior, with an initial bright spot formed at the site of the break that grows into short-lived filamentous structure. Our results suggest a new model for a cellular mechanism of homology search, where the RecA filament spans the cell length allowing it to be sampled by many different chromosomal sites in parallel.

Search for a distant homologous sister locus

The most intriguing unresolved question of HR repair is how the RecA-ssDNA complex finds a homologous sister locus within the vast amounts of heterologous DNA on a reasonable time scale. The complexity of the problem can be of different magnitudes depending on the source of a DSB. When breaks arise during DNA replication, e.g. when the replication fork encounters a DNA lesion or collides with a backtracking RNA polymerase complex (S. C. Kowalczykowski 2000; Dutta et al. 2011), the sister locus is nearby and the homology search only needs to scan a small sub-volume of a cell. On the other hand, if the DSB occurs when homologous loci are already segregated, the presynaptic complex must scan the entire volume of the cell and discriminate between millions of heterologous sequences in the process. In a previous *E. coli* study , chromosomal labels adjacent to a DSB site and its sister were reported to pair about one hour after DSB induction by the I-SceI nuclease (Lesterlin et al. 2014). However, it has been demonstrated that such labels are ejected already during end resection (Wiktor et al. 2017) which raises questions about the interpretation of the aforementioned pairing. Our results convincingly show that the pairing between distant, segregated sisters in *E. coli* is rapid and efficient, and involves a loss of the proximal DNA label. The full cycle of search, repair and resegregation takes less than 20 minutes. It is shown that the search for the homologous template finishes in less than 9 minutes, a time which also includes initial RecBCD end processing and transportation of the chromosome region to the cell center. This search commences without major reorganization of the chromosome, excluding search strategies based on global chromosome alignment. A similar HR search time has been observed in *C. crescentus*, where sister pairing after a DSB occurred not in the middle but at the polar-proximal location of the uncut sister (Badrinarayanan, Le, and Laub 2015). The different pairing location in this organism may be due to the polar tethering of the *C. crescentus* chromosome (Mohl and Gober 1997). More importantly, the short time of HR search observed here, and previously in *C. crescentus,* puts very strict demands on any model attempting to explain the mechanism of RecA-ssDNA homology search.

RecA structures

We describe a new spatio-temporal organisation of RecA complex active in the early stages of homologous recombination. Initially, fluorescent RecA forms a cluster at the break site and then extends into an elongated structure. This RecA structure reorganizes rapidly and, when repair is successful, disappears before the homologous loci segregate again. These structures appear critical in the homology search, since their appearance correlates with DSB formation and they disappear approximately when the sister loci pair in the cell center, suggesting that the once the homologies are found the RecA filament is no longer necessary. Importantly, the presence of labelled RecA is not preventing the repair as the measured repair times are nearly identical with or without an additional *recA-mVenus* allele.

Similar rapidly forming, short-lived clusters were previously observed following DSBs between unsegregated loci where no global HR search is needed (Amarh, White, and Leach 2018) There, no filaments were observed and the timescale of repair was comparable to the times observed in our work.

Elongated structures of fluorescently labelled RecA have been reported previously. In *E. coli*, in response to DSB damage by I-SceI, RecA formed thick, mostly immobile, and long-lived bundles stretching along the cell (Lesterlin et al. 2014). The structures observed in our work were remarkably different from the aforementioned bundles, *i.e.* existed for a shorter time, were flexible and dynamic, and lacked the thick central body. On the other hand, in *B. subtilis* fluorescent RecA formed thin thread-like structures, originating from an initial patch of RecA in response to mitomycin C (MMC) induction of DNA breaks. Those threads were still observed even 3 hours after the addition of MMC (Kidane and Graumann 2005).

We have not observed persistent foci of RecA-mVenus similar to those previously reported as RecA storage clusters (Renzette et al. 2005; Ghodke et al. 2019). On the contrary, the RecA-mVenus signal was evenly distributed in exponentially growing cells at most times. We did, however, detect rare spontaneous foci, or less often filaments, in unstressed cells. The spontaneous RecA structures correlate with SOS response indicators, supporting the idea that they are formed as part of DNA repair. The frequency of these RecA structures, excluding the transient ones, agrees with some previous estimates of spontaneous DNA lesions (15-30 % in Cox et al. 2000; Sinha et al. 2018), although assays by other methods have found them to occur at a much lower rate (1-2 % in Pennington and Rosenberg 2007; Shee et al. 2013). Finding spontaneous DNA lesions based on RecA structures has the advantage of detection in real-time without repair inhibition, allowing measurement of the effects of DSBs, such as division delay, in the individual cell.

Implication of the RecA filament for search kinetics

The observed long distance homology search cries out for a quantitative explanation, and one which allows search much faster than diffusion of the bulky and tethered presynaptic complex towards the homologous sister at an unknown location. We will base our model on the thin RecA filament that forms at the DSB site, grows to span the entire length of the cell, and lasts until the repair template starts moving into the mid-cell. We propose a reversed sampling mechanism, where the stretched RecA-ssDNA filament in practice is the target for search by many independently probing short dsDNA stretches. This relies on that the RecA-ssDNA filament can interact with and sample multiple different segments of DNA at the same time as has been shown *in vitro* (Forget and Kowalczykowski 2012). We consider the RecA-ssDNA filament as the target since it will be relatively stiff and immobile on the relevant time and length scale for probing interactions, i.e. <100 ms and 8 bp.

To quantify the situation we need to make a few approximations. Assume that the ssDNA is a thin rod in the center of the cylindrical nucleoid reaching from pole to pole in the z-direction, whereas the homologous dsDNA sequence is coiled up in a small volume at a random position in the nucleoid. The relevant time for the homologous sequences to find each other is the time for a segment in the coiled dsDNA locus to diffuse radially into the rod in the center of the cell. The central aspect of this model is that it does not matter at which z-coordinate it reaches the rod, since there will always be one sequence in the dsDNA that is homologous to the position in the rod that it reaches. This transforms the search problem from 3D to 2D (**Fig 4**), since we can describe the search process from the perspective of the dsDNA fragment that is homologous to the ssDNA sequence that happens to be at the z-position where the rod is reached first. An

equivalent, but slightly more stringent, way to think about the situation, is to consider that the first binding event for many independent searchers (i.e. the chromosomal dsDNA segments), that each can bind to one out of many targets (i.e. the RecA-bound ssDNA segments), has the same rate as one searcher that can bind all targets (see Supplementary material for derivation).

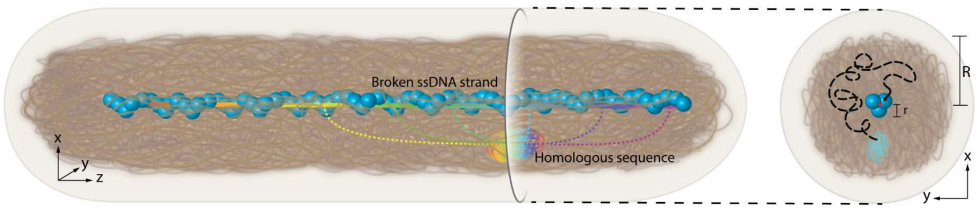

Figure 4: *Cartoon of the suggested search mechanism.* ***Left*** *The ssDNA RecA filament is stretched out in one dimension. The homologous sequence segments (in different colors) can interact with the corresponding segments of the ssDNA. Until the first binding the rates of these individual reactions are independent and the sum of the rates are equal to the rate for one dsDNA segment that can interact with all ssDNA segments.* ***Right*** *This transforms the problem to a two dimensional problem where the challenge is to diffuse from any radial coordinate to the filament in the center.*

In a first order approximation of how long it takes for a chromosomal dsDNA segment to diffuse from a random radial position in the nucleoid to the filament in its center, we can use the diffusion limited rate for reaching a rod in the center of a cylinder of length 2L, i.e. $k = 2\pi \frac{2LD}{\ln R/k}$ (Berg and Blomberg 1976), where r is the reaction radius of the rod, which is assumed to be the width of the RecA-ssDNA filament, and R is the nucleoid radius. The concentration of the searching dsDNA fragment is $c = \frac{1}{V} = \frac{1}{2L\pi R^2}$. The average time to reach the rod is therefore $T = \frac{V}{k} = \frac{R^2}{2D} \cdot \ln\frac{R}{k}$. The nucleoid radius $R \approx 200$ nm and the reaction radius is in the order of $r \approx 1$ nm, although the exact value is inconsequential since only the logarithm of the value enters into the time. The complicated parameter is the diffusion rate constant D, since DNA loci movement is subdiffusive and D is therefore lower at a long length scale than a short. The process will however be limited by the long distance movement corresponding to R. At the length scale of $R \approx 200$ nm, $D \approx 0.001\ \mu m^2 s^{-1}$ (Weber, Spakowitz, and Theriot 2012). The association step of the search process is thus $T \approx \frac{(0.2\ \mu m)^2}{0.001\ \mu m^2 s^{-1}} \cdot \ln\frac{200\ nm}{1\ nm} = 212$ s ≈ 3.5 min.

The RecA filament will however not be accessible for binding all the time since it also needs to probe all other dsDNA segments. If we exclude one minute each for end resection and translocation that leaves 7 min for the search phase. If only half is needed for homologous sequences to meet, 3.5 min is still available for probing other sequences. Is this sufficient to probe all sequences? If the dsDNA is probed in n bp long segments, each of the equally long segment of the 2 kb long ssDNA in the RecA filament will have to interrogate, on average, every $2000/n$ segment of the chromosome. There are 9.6 Mbps (4.8 Mbps per genome and ~2 genomes per cell) of dsDNA in the cell, which corresponds to $9.6 \cdot 10^6/n$ probing segments. This, in turn, means that each ssDNA segment in the RecA filament needs to test $\frac{9.6 \cdot 10^6/n}{2000/n} = 4800$ dsDNA segments. The average time for each test cannot be longer than $\frac{210\ s}{4800} \approx 45$ ms which should be plenty of time considering that Cas9-sgRNA takes at average 30 ms to make a similar task (Jones et al. 2017).

In summary, this model shows that there is at least one physically reasonable description that results in as fast search time as we measure in the experiment. The critical aspect of the model is that the RecA-ssDNA filament is stretched out in 1D which reduces the problem of finding the homologous site from 3D to 2D. In principle all cells can take advantage of the 2D diffusion to achieve fast homology search, all through the advantage is most obvious in rod shaped bacteria.

Methods

Strains

A list of strains used in the study is provided in Supplementary table S1. All strains used for imaging are derivatives of *E. coli* TB28 (Bernhardt and de Boer 2003) where the *rph-1* mutation has been restored to wild type to decrease the growth rate variance, and the *malI* gene and maltose operators have been removed to prevent interference with MalI-Venus. Genomic insertions and deletions were done using lambda red recombination (Datsenko and Wanner 2000) and transferred with P1 transductions. Resistance cassettes have been removed using Flp recombinase on the pCP20 plasmid (Cherepanov and Wackernagel 1995). A selection of strains (see Supplementary table S1) have been whole-genome sequenced to ensure that no unwanted deletions, mutations or genome rearrangements have taken place.

The non-cleavable *CS* construct is identical to the cleavable one, but the recognition sequence is changed from TAGGGATAACAGGGTAAT into TAGGatTAACAacaTAAT. The nuclease I-SceI with a LAA degradation tag was expressed from a tetracycline induced promoter on a p15a origin plasmid with ampicillin resistance (p15a-SceIdeg-amp). The catalytically inactive I-SceI (dI-SceI) was made by introducing the D44A mutation (Niu et al. 2008) in p15a-SceIdeg-amp. A SOS-response reporter was made by integrating *sCFP3a* (Kremers et al. 05/2006) into the *yibU* position preceded by a *sulA* promoter, optimized by replacing the -35 box with TTGaCA. Labels ParBMt1-mCherry (Stouf, Meile, and Cornet 2013) and MalI-mVenus were expressed by constitutive promoters from the closely located *gtrA* and *ygaY* positions, respectively. The binding site for MalI-Venus was an array of 12 maltose operators, binding up to 24 MalI-Venus molecules (Reidl et al. 1989). HupA-YPet was adopted from (Wu et al. 2015). The RecA-mVenus fusion was expressed directly downstream from the endogenous *recA* gene, and made by replacing mCherry with mVenus in the construct by (Amarh, White, and Leach 2018). A PCR fragment was amplified with the primer pair *recA-venus FW* and *DN recA P1 rev* and pKD13Venus as a template was electroporated into the EL715 strain transformed with pKD46 plasmid. The EL715 strain was made by replacing kanamycin resistance from DL5528 (Amarh, White, and Leach 2018) with a chloramphenicol resistance marker from a strain EL665.

Plasmid construction

p15a-SceIdeg-amp was cloned using Gibson assembly protocol (Gibson et al. 2009) by fusing two PCR fragments: (1) pSC101SceI_deg_amp (Wiktor et al. 2017) fragment amplified with Jwu035 and Jwu036, and (2) p15aSceI_deg_Kan (Wiktor et al. 2017) fragment amplified with Jwu037 and Jwu038. p15a-dSceIdeg-amp was cloned with Gibson assembly from two PCR fragments amplified from p15a-SceIdeg-amp template with two primer pairs: (1) Jwu088 and Jwu090, and (2) Jwu085 and Jwu091. p15a-Cas9deg-amp was cloned using Gibson assembly from two fragments: (1) p15a-SceIdeg-amp fragment amplified with Jwu273 and Jwu274, and (2) pCRED (gift from Daniel Camsund) fragment amplified with Jwu263 and Jwu264.

psgRNA-CS1 was cloned in *E. coli* Top10 by blunt-end ligation of a fragment generated with a PCR with Jwu276 and Jwu184 and psgRNA (Cui et al. 2018) plasmid as a template (psgRNA was a gift from David Bikard; Addgene plasmid # 114005).

pKD13Venus was cloned using Gibson assembly using a fragment amplified from pKD13 (Datsenko and Wanner 2000) and primers Jwu027 and Jwu028, and a fragment amplified with primers Jwu025 and Jwu026 and *mVenus* gene as template.

Microfluidic experiments

Microfluidic experiments were performed in a modified variant of the Mother Machine PDMS-glass chip with an opening in the bottom of each cell trap to facilitate cell loading and rapid media exchange (Baltekin et al. 2017). By connecting three types of media to the chip and adjusting their relative pressures in the range of 100-320 mbar, the source of the medium flowing over the cells can be exchanged within seconds. For experiments, cells cultured in minimal medium were flowed into the chip cell channels and then incubated at 37 C in the chip for at least 7 hours before DSBs were induced. Cells were normally grown in M9 medium with glucose (0.4 %), RPMI 1640 amino acid supplement (Sigma-Aldrich R7131, 0.05 %), carbenicillin (20

μg/ml) and the surfactant Pluronic F-108 (Sigma-Aldrich 542342, 21 μg/ml), which gave a generation time of ca 45 minutes. DSBs were induced by switching for three minutes to media also containing anhydrotetracycline (ATC, 20 ng/ml) and isopropyl β-d-1-thiogalactopyranoside (IPTG, 1 mM), and then for three minutes to media with only IPTG. The cells were then imaged while repairing and recovering in the initial medium.

Microscopy

The cells were imaged on a Nikon Ti inverted microscope, using a CFI Plan Apo DM Lambda 100x objective. Epifluorescent imaging was done in YFP, RFP and CFP channels. YFP and RFP were excited by 515 nm (Cobolt 06-01 MLD) and 580 nm (MPB Communications VFL) laser illumination at 11.2 and 1.7 W/cm^2, respectively, during an exposure time of 400 ms. CFP were excited by a LED lamp (Sola Lumencore) also for 400 ms. Images were acquired with a sCMOS camera (Andor Sona 4.2B-11) using the microscope's built-in 1.5 x lens. Circa 40 cell channels could be captured on one image. The growth rate or cell length were not affected by the laser exposure during imaging (**Fig S5**).

The cells were imaged once a minute in phase contrast from one hour before start of induction until three hours afterwards, and in YFP and RFP fluorescence channels once a minute from the start of the induction until two hours afterwards. The CFP channel was imaged at the start of induction and then 30, 60 and 90 minutes later.

Image analysis

Cells were segmented and cell lineages constructed either as described before in (Lawson et al. 2017) or using a specially trained neural network which will be described elsewhere. After filtering for cells with high segmentation quality, we selected cells with a total lifetime of at least 30 min, of which at least 15 min are after the start of the induction. Out of these, cells that successfully divided into daughters and whose CFP fluorescence increased by at least 5 camera counts between the start of the induction and any of the timepoints 30, 60 or 90 minutes were selected for manual inspection. Since maturation of CFP is slow compared to the generation time we also included fluorescence of the descendants of the initial cell. Images of these cells over time were then plotted and whenever a HR repair event (a suddenly lost mCherry focus followed by the remaining mCherry focus segregating into two foci again) was spotted, important events during the cell life were annotated manually. For strains with mCherry-parB and malI-venus, these events include mCherry focus loss, merge of Venus foci and re-segregation of mVenus and mCherry foci. For strains with RecA-mVenus, appearance of the initial focus and the time point when a mVenus structure was visible outside that focus were recorded in addition to the mCherry events mentioned before. Fluorescent foci were detected using à trous wavelet decomposition (Olivo-Marin 2002). Growth rates were calculated by fitting an exponential function to the backbone length of each cell, using a 20 min sliding window.

RecA-mVenus structures were detected using a neural network implemented in Matlab, with an input size of 13 by 42 px, 4 convolutional layers and one fully connected layer. The network was trained and validated on a manually annotated dataset of 7069 unstructured and 1220 structured cells, the latter which were augmented four times by adding random noise. Cell images were extracted from fluorescence images using phase contrast based cell segmentations dilated by 3 px, and then reduced to a height of 13 px by removing the darkest rows at the top or bottom, while the length cropped one pixel in each end and then scaled to 42 px, before being input into the network.

Plate reader experiments

Cells were incubated overnight in the same media as used for microfluidic experiments. The overnight cultures were diluted in triplicates, 1 into 100 ul in a 96 well plate and incubated shaking for 3 (at 37 °C) or 4 h (at 30 °C). Then another 100 ul preheated media with double concentrations of inducers was added, and the plates were incubated while OD600 and fluorescence measurements were taken in a Tecan Spark plate reader with humidity cassette. Measurements from media-only wells were subtracted and the ratios of fluorescence and OD600 after 3 (at 37 °C) or 4 h (at 30 °C) were plotted.

Creating DSBs with Cas9

The experiment in which DSBs were induced by Cas9 were performed as above, but with following exceptions: The EL1217 strain was transformed with plasmids p15a-Cas9deg-amp and psgRNA-CS1. This gRNA targets a region about 200 bp away from I-SceI recognition site. Induction was done with a 90 s pulse of ATC (20 ng/ml). Cells were first selected using a neural network trained to recognize RecA structures, and then manually screened for disappearance of ParB-mCherry focus. Repair times were annotated manually.

Serial-dilution plating assay

Bacterial cultures were grown overnight at 37°C in LB media supplemented with carbenicillin (100 μg/ml). Following day, the cultures were diluted 1:1000 in fresh LB media with carbenicillin (100 μg/ml) and grown at 37°C for 5 hours. Next, 10-fold dilutions were made in PBS buffer and 5 ul of dilutions were plated on LB agarose plates with carbenicillin (100 μg/ml), and different concentrations of ATC (5, 20, 40, or 100 ng/ml). Plates were incubated overnight at 30°C, or 37°C.

Acknowledgements

The authors would like to thank Spartak Zikrin and Praneeth Karempudi for continuous image analysis development, Irmeli Barkefors, David Fange and Daniel Jones for careful reading of the manuscript, and Daniel Camsund for pCRED plasmid. J.W is an EMBO non-stipendiary fellow (EMBO ALTF 1029-2018). This work was funded by H2020 Marie Curie Individual Fellowship grant, Agreement Number 842047 for J.W., the Knut and Alice Wallenberg foundation and the Swedish Research Council.

References

Amarh, Vincent, Martin A. White, and David R. F. Leach. 2018. "Dynamics of RecA-Mediated Repair of Replication-Dependent DNA Breaks." *The Journal of Cell Biology* 217 (7): 2299–2307.

Anand, Ranjith, Annette Beach, Kevin Li, and James Haber. 2017. "Rad51-Mediated Double-Strand Break Repair and Mismatch Correction of Divergent Substrates." *Nature* 544 (7650): 377–80.

Badrinarayanan, Anjana, Tung B. K. Le, and Michael T. Laub. 2015. "Rapid Pairing and Resegregation of Distant Homologous Loci Enables Double-Strand Break Repair in Bacteria." *The Journal of Cell Biology* 210 (3): 385–400.

Baltekin, Özden, Alexis Boucharin, Eva Tano, Dan I. Andersson, and Johan Elf. 2017. "Antibiotic Susceptibility Testing in Less than 30 Min Using Direct Single-Cell Imaging." *Proceedings of the National Academy of Sciences of the United States of America* 114 (34): 9170–75.

Berg, Otto G., and Clas Blomberg. 1976. "Association Kinetics with Coupled Diffusional Flows." *Biophysical Chemistry*. https://doi.org/10.1016/0301-4622(76)80017-8.

Bernhardt, Thomas G., and Piet A. J. de Boer. 2003. "The Escherichia Coli Amidase AmiC Is a Periplasmic Septal Ring Component Exported via the Twin-Arginine Transport Pathway." *Molecular Microbiology* 48 (5): 1171–82.

Bertrand, Claire, Annabelle Thibessard, Claude Bruand, François Lecointe, and Pierre Leblond. 2019. "Bacterial NHEJ: A Never Ending Story." *Molecular Microbiology* 111 (5): 1139–51.

Cherepanov, P. P., and W. Wackernagel. 1995. "Gene Disruption in Escherichia Coli: TcR and KmR Cassettes with the Option of Flp-Catalyzed Excision of the Antibiotic-Resistance Determinant." *Gene* 158 (1): 9–14.

Clark, A. J., and A. D. Margulies. 1965. "Isolation and Characterization of Recombination-Deficient Mutants of Escherichia Coli K12." *Proceedings of the National Academy of Sciences of the United States of America* 53 (February): 451–59.

Cox, M. M., M. F. Goodman, K. N. Kreuzer, D. J. Sherratt, S. J. Sandler, and K. J. Marians. 2000. "The Importance of Repairing Stalled Replication Forks." *Nature* 404 (6773): 37–41.

Cui, Lun, and David Bikard. 2016. "Consequences of Cas9 Cleavage in the Chromosome of Escherichia Coli." *Nucleic Acids Research* 44 (9): 4243–51.

Cui, Lun, Antoine Vigouroux, François Rousset, Hugo Varet, Varun Khanna, and David Bikard. 2018. "A CRISPRi Screen in E. Coli Reveals Sequence-Specific Toxicity of dCas9." *Nature Communications* 9 (1): 1912.

Datsenko, K. A., and B. L. Wanner. 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 (12): 6640–45.

Dillingham, Mark S., and Stephen C. Kowalczykowski. 2008. "RecBCD Enzyme and the Repair of Double-Stranded DNA Breaks." *Microbiology and Molecular Biology Reviews: MMBR* 72 (4): 642–71.

Dutta, Dipak, Konstantin Shatalin, Vitaly Epshtein, Max E. Gottesman, and Evgeny Nudler. 2011. "Linking RNA Polymerase Backtracking to Genome Instability in E. Coli." *Cell* 146 (4): 533–43.

Forget, Anthony L., and Stephen C. Kowalczykowski. 2012. "Single-Molecule Imaging of DNA Pairing by RecA Reveals a Three-Dimensional Homology Search." *Nature* 482 (7385): 423–27.

Ghodke, Harshad, Bishnu P. Paudel, Jacob S. Lewis, Slobodan Jergic, Kamya Gopal, Zachary J. Romero, Elizabeth A. Wood, Roger Woodgate, Michael M. Cox, and Antoine M. van Oijen. 2019. "Spatial and Temporal Organization of RecA in the Escherichia Coli DNA-Damage Response." *eLife* 8 (February): e42761.

Gibson, Daniel G., Lei Young, Ray-Yuan Chuang, J. Craig Venter, Clyde A. Hutchison 3rd, and Hamilton O. Smith. 2009. "Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases." *Nature Methods* 6 (5): 343–45.

Jones, Daniel Lawson, Prune Leroy, Cecilia Unoson, David Fange, Vladimir Ćurić, Michael J. Lawson, and Johan Elf. 2017. "Kinetics of dCas9 Target Search in Escherichia Coli." *Science* 357 (6358): 1420–24.

Kidane, Dawit, and Peter L. Graumann. 2005. "Dynamic Formation of RecA Filaments at DNA Double Strand Break Repair Centers in Live Cells." *The Journal of Cell Biology* 170 (3): 357–66.

Kowalczykowski, S. C. 2000. "Initiation of Genetic Recombination and Recombination-Dependent Replication." *Trends in Biochemical Sciences* 25 (4): 156–65.

Kowalczykowski, Stephen C. 2015. "An Overview of the Molecular Mechanisms of Recombinational DNA Repair." *Cold Spring Harbor Perspectives in Biology* 7 (11). https://doi.org/10.1101/cshperspect.a016410.

Kremers, Gert-Jan, Joachim Goedhart, Erik B. van Munster, and Theodorus W. J. Gadella. 05/2006. "Cyan and Yellow Super Fluorescent Proteins with Improved Brightness, Protein Folding, and FRET Förster Radius ‡ , ‡." *Biochemistry* 45 (21): 6570–80.

Kuzminov, A. 1995. "Collapse and Repair of Replication Forks in Escherichia Coli." *Molecular Microbiology* 16 (3): 373–84.

Lawson, Michael J., Daniel Camsund, Jimmy Larsson, Ozden Baltekin, David Fange, and Johan Elf. 2017. "In Situ Genotyping Of A Pooled Strain Library After Characterizing Complex Phenotypes." *bioRxiv*. https://doi.org/10.1101/142729.

Lesterlin, Christian, Graeme Ball, Lothar Schermelleh, and David J. Sherratt. 2014. "RecA Bundles Mediate Homology Pairing between Distant Sisters during DNA Break Repair." *Nature* 506 (7487): 249–53.

Meddows, Tom R., Andrew P. Savory, Jane I. Grove, Timothy Moore, and Robert G. Lloyd. 2005. "RecN Protein and Transcription Factor DksA Combine to Promote Faithful Recombinational Repair of DNA Double-Strand Breaks." *Molecular Microbiology* 57 (1): 97–110.

Mehta, Anuja, and James E. Haber. 2014. "Sources of DNA Double-Strand Breaks and Models of Recombinational DNA Repair." *Cold Spring Harbor Perspectives in Biology* 6 (9): a016428.

Mohl, D. A., and J. W. Gober. 1997. "Cell Cycle-Dependent Polar Localization of Chromosome Partitioning Proteins in Caulobacter Crescentus." *Cell* 88 (5): 675–84.

Nielsen, Henrik J., Jesper R. Ottesen, Brenda Youngren, Stuart J. Austin, and Flemming G. Hansen. 2006. "The Escherichia Coli Chromosome Is Organized with the Left and Right Chromosome Arms in Separate Cell Halves." *Molecular Microbiology* 62 (2): 331–38.

Niu, Yan, Kristen Tenney, Hongye Li, and Frederick S. Gimble. 2008. "Engineering Variants of the I-SceI Homing Endonuclease with Strand-Specific and Site-Specific DNA-Nicking Activity." *Journal of Molecular Biology* 382 (1): 188–202.

Odsbu, Ingvild, and Kirsten Skarstad. 2014. "DNA Compaction in the Early Part of the SOS Response Is Dependent on RecN and RecA." *Microbiology* 160 (Pt 5): 872–82.

Olivo-Marin, Jean-Christophe. 2002. "Extraction of Spots in Biological Images Using Multiscale Products." *Pattern Recognition* 35 (9): 1989–96.

Pennington, Jeanine M., and Susan M. Rosenberg. 2007. "Spontaneous DNA Breakage in Single Living Escherichia Coli Cells." *Nature Genetics* 39 (6): 797–802.

Perrin, A., M. Buckle, and B. Dujon. 1993. "Asymmetrical Recognition and Activity of the I-SceI Endonuclease on Its Site and on Intron-Exon Junctions." *The EMBO Journal* 12 (7): 2939–47.

Qi, Zhi, Sy Redding, Ja Yil Lee, Bryan Gibb, Youngho Kwon, Hengyao Niu, William A. Gaines, Patrick Sung, and Eric C. Greene. 2015. "DNA Sequence Alignment by Microhomology Sampling during Homologous Recombination." *Cell* 160 (5): 856–69.

Reidl, J., K. Römisch, M. Ehrmann, and W. Boos. 1989. "MalI, a Novel Protein Involved in Regulation of the Maltose System of Escherichia Coli, Is Highly Homologous to the Repressor Proteins GalR, CytR, and LacI." *Journal of Bacteriology* 171 (9): 4888–99.

Renkawitz, Jörg, Claudio A. Lademann, and Stefan Jentsch. 2014. "Mechanisms and Principles of Homology Search during Recombination." *Nature Reviews. Molecular Cell Biology* 15 (6): 369–83.

Renzette, Nicholas, Nathan Gumlaw, Jared T. Nordman, Marlee Krieger, Su-Ping Yeh, Edward Long, Richard Centore, Ruethairat Boonsombat, and Steven J. Sandler. 2005. "Localization of RecA in Escherichia Coli K-12 Using RecA-GFP." *Molecular Microbiology* 57 (4): 1074–85.

Reyes, Emigdio D., Praveen L. Patidar, Lee A. Uranga, Angelina S. Bortoletto, and Shelley L. Lusetti. 2010. "RecN Is a Cohesin-like Protein That Stimulates Intermolecular DNA Interactions in Vitro." *The Journal of Biological Chemistry* 285 (22): 16521–29.

Roukos, V., T. C. Voss, C. K. Schmidt, S. Lee, D. Wangsa, and T. Misteli. 2013. "Spatial Dynamics of Chromosome Translocations in Living Cells." *Science*. https://doi.org/10.1126/science.1237150.

Shee, Chandan, Ben D. Cox, Franklin Gu, Elizabeth M. Luengas, Mohan C. Joshi, Li-Ya Chiu, David Magnan, et al. 2013. "Engineered Proteins Detect Spontaneous DNA Breakage in Human and Bacterial Cells." *eLife* 2 (October): e01222.

Sinha, Anurag Kumar, Christophe Possoz, Adeline Durand, Jean-Michel Desfontaines, François-Xavier Barre, David R. F. Leach, and Bénédicte Michel. 2018. "Broken Replication Forks Trigger Heritable DNA Breaks in the Terminus of a Circular Chromosome." *PLoS Genetics* 14 (3): e1007256.

Stouf, Mathieu, Jean-Christophe Meile, and François Cornet. 2013. "FtsK Actively Segregates Sister Chromosomes in Escherichia Coli." *Proceedings of the National Academy of Sciences of the United States of America* 110 (27): 11157–62.

Vink, Jochem N. A., Koen J. A. Martens, Marnix Vlot, Rebecca E. McKenzie, Cristóbal Almendros, Boris Estrada Bonilla, Daan J. W. Brocken, Johannes Hohlbein, and Stan J. J. Brouns. 2019. "Direct Visualization of Native CRISPR Target Search in Live Bacteria Reveals Cascade DNA Surveillance Mechanism." *Molecular Cell*, November. https://doi.org/10.1016/j.molcel.2019.10.021.

Weber, Stephanie C., Andrew J. Spakowitz, and Julie A. Theriot. 2012. "Nonthermal ATP-Dependent Fluctuations Contribute to the in Vivo Motion of Chromosomal Loci." *Proceedings of the National Academy of Sciences of the United States of America* 109 (19): 7338–43.

Wery, M., C. L. Woldringh, and J. Rouviere-Yaniv. 2001. "HU-GFP and DAPI Co-Localize on the Escherichia Coli Nucleoid." *Biochimie* 83 (2): 193–200.

Wiktor, Jakub, Marit van der Does, Lisa Büller, David J. Sherratt, and Cees Dekker. 2017. "Direct Observation of End Resection by RecBCD during Double-Stranded DNA Break Repair in Vivo." *Nucleic Acids Research*, December. https://doi.org/10.1093/nar/gkx1290.

Wu, Fabai, Erwin Van Rijn, Bas G. C. Van Schie, Juan E. Keymer, and Cees Dekker. 2015. "Multi-Color Imaging of the Bacterial Nucleoid and Division Proteins with Blue, Orange, and near-Infrared Fluorescent Proteins." *Frontiers in Microbiology* 6 (June): 607.

Supplementary material for

RecA mediated homology search finds segregated sister locus in minutes after a double stranded break

Arvid H. Gynnå*, Jakub Wiktor*, Prune Leroy & Johan Elf
Department of Cell and Molecular Biology, SciLifeLab, Uppsala University, Sweden
*equal contribution

Meganuclease I-SceI has non-specific activity

Surprisingly, we found that I-SceI expression at 30 °C causes activation of the SOS response even when no restriction site is present, while a catalytically inactive I-SceI enzyme (dI-SceI) did not (**Fig S1A-D**). This finding suggested that I-SceI has activity on DNA outside of its known restriction site. When measuring cell survival by plating on LB agar plates with ATC, we found that cells with intact DNA repair systems were unaffected except at very high concentrations (**Fig S1E**). RecA deficient cells were however much more susceptible to this activity, while RecN deficient cells showed intermediate survival. This indicates that the unspecific damage caused by I-SceI is repairable by HR. We believe this activity is double stranded cleavage at other sites than canonical restriction sites. We have not determined if the activity is truly without specificity or if it, perhaps more likely, occurs at DNA sequences following a more or less well-defined pattern. For the purpose of minimizing this non-specific cleavage we induced I-SceI at different temperatures and found that the toxicity was lower at 37 °C then at 30 °C (**Fig S1A-B**). This non-specific activity of I-SceI has not been reported previously and we believe it may have been a confusing factor in earlier studies where this enzyme was used.

The rate of first binding of many searchers to individual targets equals the rate of binding of a single searcher to many targets

The total rate of binding is $r = \Sigma_i r_i$, where r_i is the rate for template segment i to find its homologous ssDNA segment. If we write out the dependence of at which z-coordinate z_j, the filament is reached, the total rate can be expressed as $r = \Sigma_j \Sigma_i r_i(z_j) p(z_j)$, where $p(z_j)$ is the probability to reach the filament at position z_j and $r_i(z_j)$ is the conditional rate for segment i of binding given that the filament is reached at z_j. Here, the rate of binding is zero unless the template segment matches the ssDNA that is at the specific z-position, i.e. $r_i(z_{j \neq i}) = 0$ which means that as $r = \Sigma_i r_i(z_i) p(z_i)$, i.e. the total binding rate is the same as the binding rate for a single dsDNA segment that can bind at any position at the filament, irrespective of z-position, and for which each position is homologous.

Figure S1: Non-specific I-SceI activity

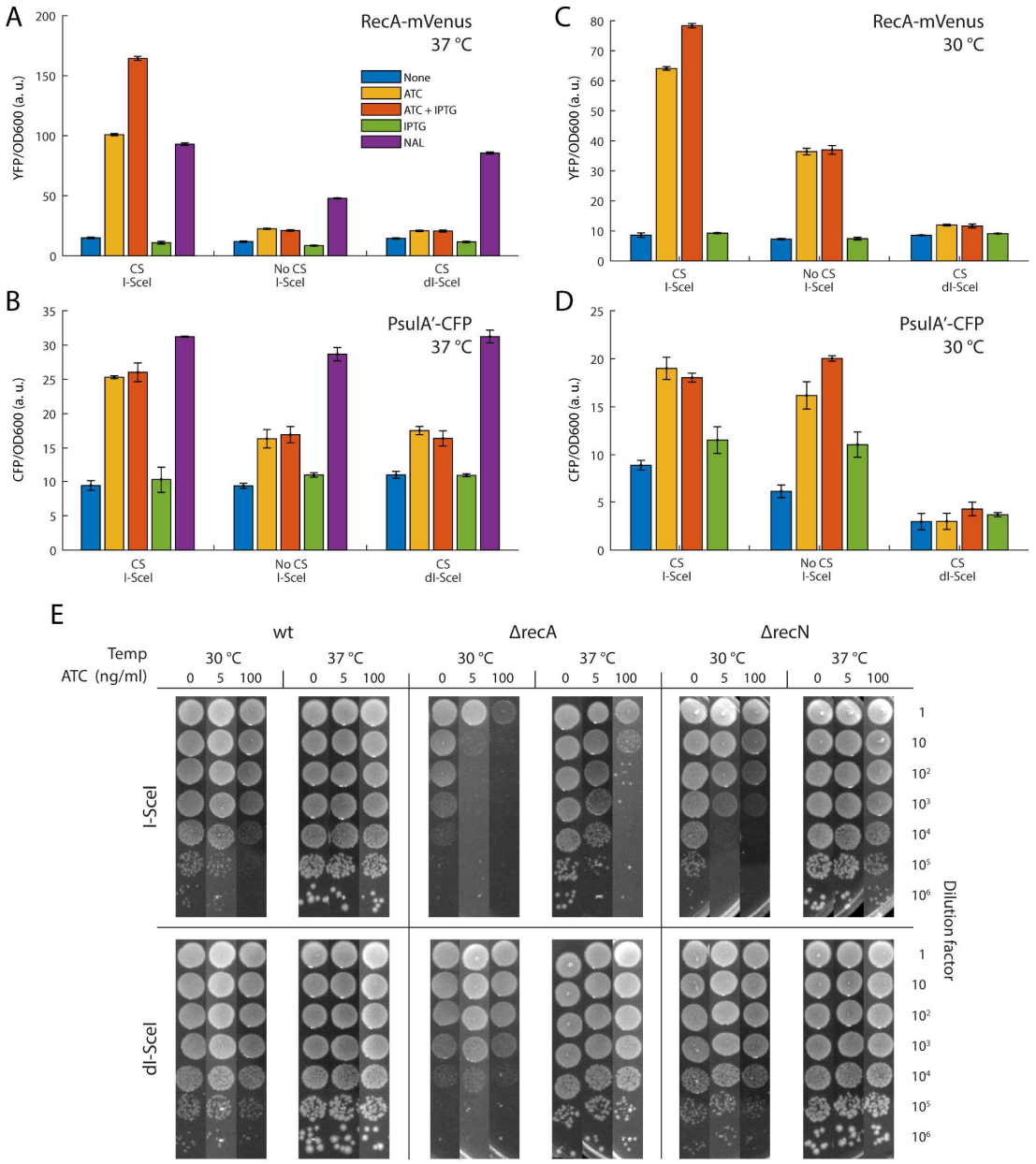

Figure S1: A. RecA-mVenus fluorescence in plate reader after incubation at 37 C for 3 hours with ATC (2 ng/ml), IPTG (1 mM) or nalidixic acid (NAL, 10 ng/ml). ATC induces the nuclease I-SceI, IPTG releases the CS, NAL is a DNA gyrase inhibitor. **B.** As A, but PsulA'-CFP fluorescence. **C.** As A. but after incubation at 30 C for 4 hours. **D.** As C, but PsulA'-CFP fluorescence. **E.** Serial dilutions of wild type, ΔrecA and ΔrecN cells without I-SceI restriction sites, plated on different concentrations of ATC and incubated at 30 °C or 37 °C. Strains were transformed with p15a-SceIdeg-amp (I-SceI), or p15a-dSceIdeg-amp (dI-SceI).

Figure S2: Temporary sister loci segregation

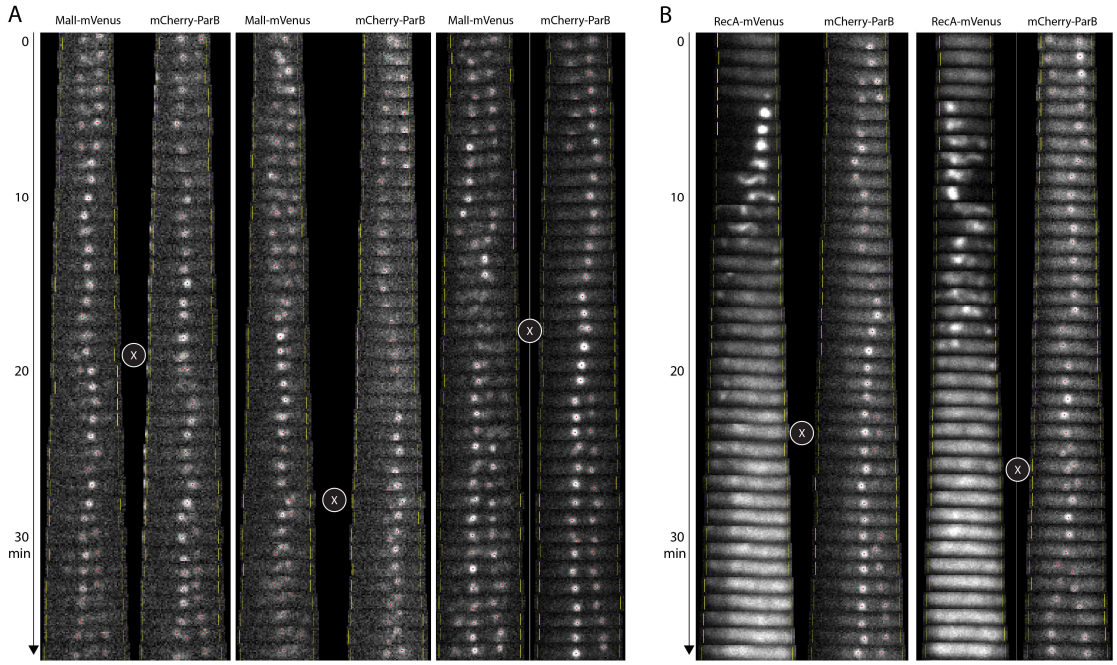

Figure S2: Examples of cells showing temporary split of mCherry-ParB foci after sister pairing but before definitive re-segregation. An X indicates when a temporary split takes place. **A.** In cells with a MalI-Venus marker 25 kb from the CS. **B.** In cells expressing RecA-mVenus.

Figure S3: Spontaneous RecA-mVenus structures

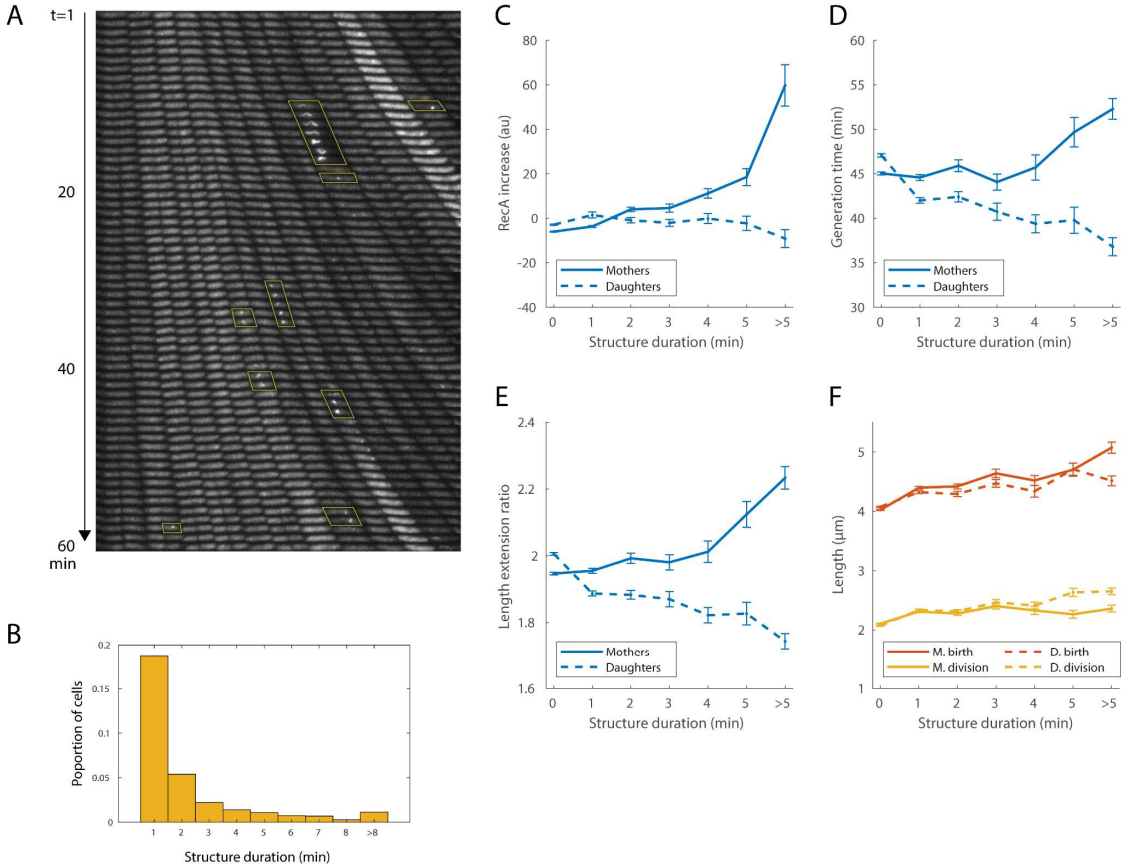

Figure S3: A. An example of a microfluidic channel with cells carrying the RecA-Venus gene during steady state growth in a mother machine cell trap. Short-lived RecA-Venus structures are annotated. **B.** Proportion of cells with spontaneous RecA-mVenus structures of different duration. (n=8220) **C.** SOS induction, measured by RecA-mVenus increase, in cells with RecA-mVenus structures and their daughters. **D.** Generation time of the cells in C. **E.** Length extension, as ratio between division and birth length, of the cells in C. **G.** Birth and division length of the cells in C. Error bars represent standard error. (mothers n=5577, 1505, 436, 179, 111, 87 & 221, daughters n=4709, 1440, 419, 192, 110, 75 & 177)

Figure S4: ΔrecN mutants and Cas9 induced DSBs

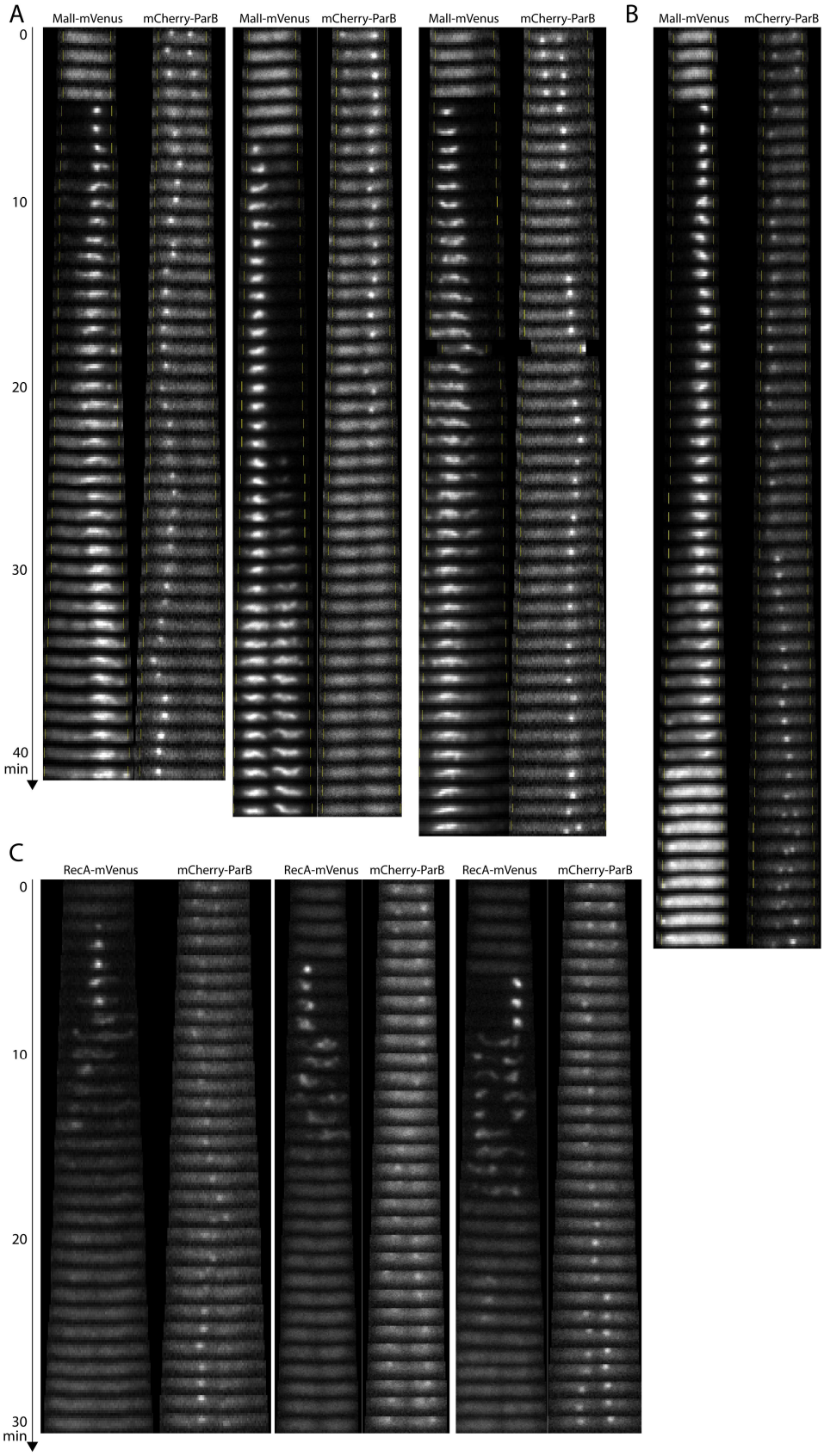

Figure S4: (previous page) **A.** Examples of ⊿*rec*N cells with single DSB which do not repair. **B.** Example of ⊿*rec*N which does repair a single DSB. In A-B, every RecA-mVenus image is individually intensity scaled due to the strong induction of the RecA promoter over the course of the imaging**. C.** Examples of single DSBs induced by Cas9.

Figure S5: Growth with and without laser exposure

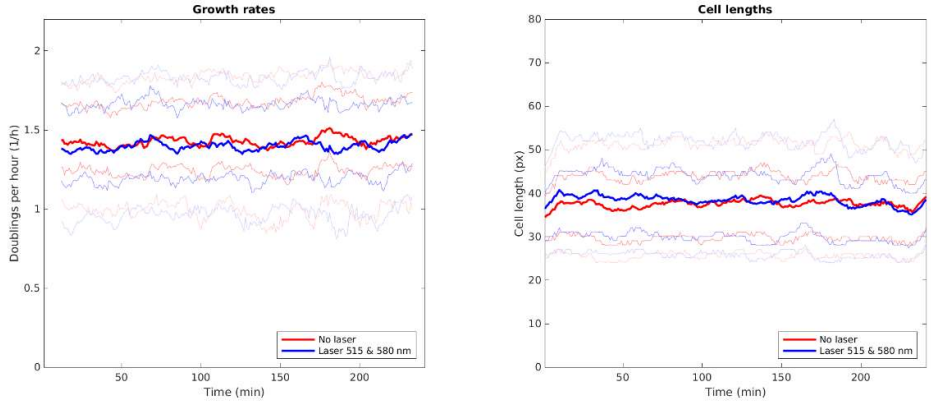

Figure S5: Growth rate and length of cells exposed to 515 nm and 580 nm laser, 400 ms once every minute from 60 to 180 minutes, while growing without DSB induction. Growth rate is estimated from cell elongation in a 20-minute sliding window.

Table S1: Strains used

Strain	Comment	Genotype	Used in	Whole genome sequenced
EL662	Base strain	E. coli MG1655 rph+ Δlac ΔmalI::frt-TetR-frt		
EL665		Eco TB28 codA::I-SceI_array-parSP1-frt-CmR-frt		
EL715		Eco codA::I-SceI_array-frt-CmR-frt recA-recX::recAmwg-mCherry		
EL1217		EL662 lac+ gtrA::P58-mCherry:parB-SpR recA-recX::recAmwg-venus-frt codA::parBMt1-lacOsym-lacOCSmut-frt yibU':PsulA'-SCFP3a-frt	Fig S4C	
EL1219	CS-, recA-mVenus	EL662 lac+ gtrA::P58-mCherry:parB-SpR recA-recX::recAmwg-venus-frt codA::parBMt1-lacOsym-lacOCSmut-frt yibU':PsulA'-SCFP3a / p15a-ISceIdeg-amp	Fig S1A-D	Yes
EL1220	CS+, recA-mVenus	EL662 lac+ gtrA::P58-mCherry:parB-SpR recA-recX::recAmwg-venus-frt codA::parBMt1-lacOsym-lacOCS-frt yibU':PsulA'-SCFP3a / p15a-ISceIdeg-amp	Fig 3A-B, S1A-D, S2B, Table 1	Yes
EL1226	CS+, -25 kb malO	EL662 lac+ intC::P59-malIvenus-frt gtrA::P58-mCherry:parB-SpR codA::parBMt1-lacOsym-lacOCS yibU':PsulA'-sCFP3a yahA::malOx12 / p15a-ISceIdeg-amp	Fig 1C-G, 2A, 2D, S2A, S5, Table 1	
EL1227	CS-, -25 kb malO	EL662 lac+ intC::P59-malIvenus-frt gtrA::P58-mCherry:parB-SpR codA::parBMt1-lacOsym-lacOCSmut yibU':PsulA'-SCFP3a yahA::malOx12 / p15a-dISceIdeg-amp		
EL1338	CS+, hupA-YFP	EL662 lac+ gtrA::P58-mCherry:parB-SpR codA::lacOsym-lacOCS-parBMt1-frt hupA-yPet-frt yibU'::PsulA'-scfp3a-frt / p15a-ISceIdeg-amp	Fig 2C, 2F	
EL1900	CS+, left arm malO	EL662 lac+ intC::P59-malIvenus-frt gtrA::P58-mCherry:parB-SpR ygaY'::malOx12 yibU'::PsulA'-scfp3a codA::parBMt1-lacOsym-lacOCS / p15a-ISceIdeg-amp	Fig 2B, 2E	
EL1910	CS+, -25 kb malO, dI-SceI	EL662 lac+ intC::P59-malIvenus-frt gtrA::P58-mCherry:parB-SpR codA::parBMt1-lacOsym-lacOCS yibU':PsulA'-SCFP3a yahA::malOx12 / p15a-dISceIdeg-amp		
EL1926	recA-venus, dI-SceI	EL662 lac+ gtrA::P58-mCherry:parB-SpR recA-recX::recAmwg-venus-frt codA::parBMt1-lacOsym-lacOCS-frt yibU':PsulA'-SCFP3a-frt / p15a-dISceIdeg-amp	Fig S3A-F	
EL1926	CS+, recA-mVenus, dI-SceI	EL662 lac+ gtrA::P58-mCherry:parB-SpR recA-recX::recAmwg-venus-frt codA::parBMt1-lacOsym-lacOCS-frt yibU':PsulA'-SCFP3a-frt / p15a-dISceIdeg-amp	Fig S1A-D	
EL2051	ΔrecN, I-SceI	Eco MG1655 Δlac ΔrecN::frt-CmR-frt / p15a-ISceIdeg-amp	Fig S1E	
EL2052	ΔrecN, dI-SceI	Eco MG1655 Δlac ΔrecN::frt-CmR-frt / p15a-dISceIdeg-amp	Fig S1E	
EL2053	I-SceI	Eco MG1655 Δlac / p15a-ISceIdeg-amp	Fig S1E	
EL2054	dI-SceI	Eco MG1655 Δlac / p15a-dISceIdeg-amp	Fig S1E	
EL2055	ΔrecA, I-SceI	Eco MG1655 Δlac ΔrecA::frt-kan-frt / p15a-ISceIdeg-amp	Fig S1E	
EL2056	ΔrecA, dI-SceI	Eco MG1655 Δlac ΔrecA::frt-kan-frt / p15a-dISceIdeg-amp	Fig S1E	

Table S2: Primers

Primer	Purpose	Sequence
recA-venus FW	recA-mVenus	TAGCCGAAACCAACGAGGACTTTGGGAGCATCGTTTCCAAGGGCGAGGAG ctgttcaccgg
DN recA P1 rev		Gcaaaagggccgcagatgcgacccttgtgtatcaaacaagacga TTATTTGGTGTAGGCTGGAGCTGCTTC
Jwu035	p15a-SceIdeg-amp	cttaacagggaagtgagaggtagggtacgggttttgctgcc
Jwu036		GATATCGAGCTCGCTTGGACgcttggactcctgttgatagatc
Jwu037	p15a-SceIdeg-amp	GTCCAAGCGAGCTCGATATC
Jwu038		cctctcacttccctgttaagtatc
Jwu088	p15a-dSceIdeg-amp	gatagatccagtaatgacctcag
Jwu090		GTATCGGTCTGATCCTGGGTGCGGCTTACATCCGTTCTCGTGATG
Jwu085	p15a-dSceIdeg-amp	ctgaggtcattactggatctatc
Jwu091		ACCCAGGATCAGACCGATAC
Jwu273	p15a-Cas9deg-amp	tgagtcagctaggaggtgacGCTGCTAACGACGAAAACTACG
Jwu274		gctaggtgtagagcagcctac
Jwu263	p15a-Cas9deg-amp	cactttacttttatctaatc
Jwu264		gtcacctcctagctgactcaaatc
Jwu276	psgRNA-CS1	AGCAAAAGGGCCGCAGATGCGACCCTTGTGTATCAAACAAGACGA ctcatcgcagtactgttgtattc
Jwu184		actagtattatacctaggactgag
Jwu027	pKD13Venus	gacgagctgtacaaggaattgtgaCCGTCGACCTGCAGTTCGAAG
Jwu028		ACCTGAACCTTGCCCCGATCCCCGGAATTAATTCTCATGTTTGACAG
Jwu025	pKD13Venus	GGATCGGGGCAAGGTTCAGGTCCTGGATCGGGGCAGGGAAGCGGTCCGGGA gtgagcaagggcgaggagctg
Jwu026		tcacaattccttgtacagctcgtc