A chromosomal loop anchor mediates bacterial genome organization

Dugar, G.; Hofmann, A.; Heermann, D.W.; Hamoen, L.W.

DOI
10.1038/s41588-021-00988-8

Publication date
2022

Document Version
Final published version

Published in
Nature genetics

ARTICLES
https://doi.org/10.1038/s41588-021-00988-8
nature genetics

A chromosomal loop anchor mediates bacterial genome organization

Gaurav Dugar[1]✉, Andreas Hofmann[2], Dieter W. Heermann[2] and Leendert W. Hamoen[1]✉

Nucleoprotein complexes play an integral role in genome organization of both eukaryotes and prokaryotes. Apart from their role in locally structuring and compacting DNA, several complexes are known to influence global organization by mediating long-range anchored chromosomal loop formation leading to spatial segregation of large sections of DNA. Such megabase-range interactions are ubiquitous in eukaryotes, but have not been demonstrated in prokaryotes. Here, using a genome-wide sedimentation-based approach, we found that a transcription factor, Rok, forms large nucleoprotein complexes in the bacterium *Bacillus subtilis*. Using chromosome conformation capture and live-imaging of DNA loci, we show that these complexes robustly interact with each other over large distances. Importantly, these Rok-dependent long-range interactions lead to anchored chromosomal loop formation, thereby spatially isolating large sections of DNA, as previously observed for insulator proteins in eukaryotes.

Structural organization of genomes facilitates DNA compaction, replication and segregation, and gene regulation[1–7]. Eukaryotes show a hierarchical organization of their chromosomes: from nucleosomes at the kilobase (kb) scale to higher order, physically isolated, domains at the megabase (Mb) scale[5,8]. Key players in this higher order structural organization of the genome are chromatin insulators, which are DNA elements associated with specific insulator proteins that mediate the formation of long-range chromatin loops, to physically isolate large regions of the chromosome[8–11]. Genomic regions within such loops self-interact more frequently and are called topologically associating domains, or TADs. These domains are physically and functionally isolated genome units that help co-regulate genes and block the spread of regulatory activity to adjacent domains[12,13]. Although insulator proteins are not present in bacteria, TAD-like domains have been observed at sites of highly expressed genes in bacteria and are termed chromosomal interaction domains (CIDs)[1,14,15].

Condensin/cohesin complexes composed of structural maintenance of chromosomes (SMC) proteins are major players in DNA compaction and organization in all organisms, including bacteria[15,16]. In *Bacillus subtilis* and *Caulobacter crescentus*, SMC complexes have been shown to tether the chromosome arms together using loop extrusion[1,17]. The SMC homologs MukB in *Escherichia coli* and MksB in *Pseudomonas aeruginosa* have been shown to organize chromosomes by promoting *cis* contacts between chromosomal loci[4,18], and recently an SMC-like protein termed coalescin (ClsN) in archaea was found to be involved in compartmentalization of its chromosomes[19]. However, chromosome conformation capture (3C) techniques, including Hi-C, have provided no evidence that these prokaryotic SMC complexes are directly involved in the formation of CIDs in bacteria[1,4,5,20].

In eukaryotes, promoters are often regulated by enhancer regions that are located several kb to Mb away on the chromosome, and insulator proteins block interactions between enhancers and promoters by formation of physically isolated TADs[21,22]. In prokaryotes, genes are regulated by transcription factors that bind close to the promoter region, generally within several hundred bp of the transcription start site[23]. The absence of long-range promoter–enhancer interaction in prokaryotes has led to the assumption that protein-mediated long-range anchored loops may have primarily evolved in eukaryotes[5,11]. Here, aided by a chromatin sedimentation assay and Hi-C, we show that the transcription factor Rok forms large (Mb range) and well-defined anchored chromosomal loops leading to the formation of CIDs in the bacterium *B. subtilis*. By spatially isolating large regions of the genome, Rok displays similarities with insulator proteins found in eukaryotes.

Results

Sedimentation-based approach reveals large DNA complexes. Sedimentation-based approaches have been applied routinely to interrogate macromolecular complexes like protein–protein and protein–RNA interactions[24,25]. We reasoned that a similar approach may reveal vital clues about higher order chromosome organization in bacteria. To test this, we mildly treated the lysate of *B. subtilis* with DNase and subsequently added it on top of dense sucrose steps. After ultracentrifugation, the top fraction was sequenced to identify the absence of DNA regions that had entered the sucrose, and hence could be part of large complexes potentially involved in chromosome organization (Fig. 1a and Extended Data Fig. 1a). Using this approach, we found eight genomic regions that partially disappeared from the top fraction after centrifugation (Fig. 1b and Extended Data Fig. 1b). These chromosomal complexes could be recovered from the sucrose-dense fraction (Fig. 1b). To find potential DNA-binding proteins involved in the formation of these large DNA complexes, we analyzed existing chromatin immunoprecipitation (ChIP) datasets for *B. subtilis*. We found that all eight sites overlapped with the binding sites for the transcription factor Rok[26] (Supplementary Table 3). To examine whether Rok is responsible for formation of these complexes, we performed the DNA sedimentation analysis using a *rok* deletion strain. As shown in Fig. 1b, this prevented the sedimentation of all eight sites, and these regions were subsequently termed Rok clusters.

Higher density of Rok motifs are observed in Rok clusters. Rok is a small (191 amino acid) DNA-binding protein and its role as a

[1]Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, the Netherlands. [2]Institute for Theoretical Physics, Heidelberg University, Heidelberg, Germany. ✉e-mail: g.dugar@uva.nl; l.w.hamoen@uva.nl

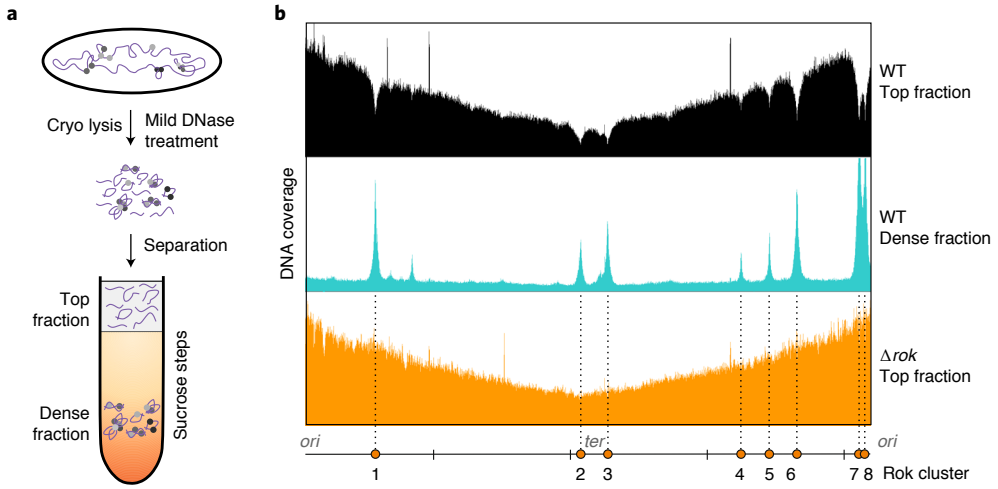


Fig. 1 | Identification of DNA associated with large complexes in *B. subtilis*. **a**, Schematic representation of the sedimentation-based DNA coverage analysis. Large DNA–protein complexes are removed from cell lysate using ultracentrifugation over sucrose steps after mild DNase treatment. The DNA retained in the cleared top fraction and dense fraction are sequenced. Lower coverage at some regions of the genome in the top fraction indicates potential formation of large nucleoprotein complexes at these sites. **b**, DNA coverage maps of the top fraction (black) along with the dense fraction (cyan) obtained from the wild-type (WT) strain. DNA coverage map of the top fraction of the Δ*rok* strain (orange) is shown below. Rok clusters are marked using orange dots. The higher DNA coverage around the origin compared to the terminus is due to multifork DNA replication in exponentially growing *B. subtilis* cells. Genome positions in Mb intervals are marked below.

transcription repressor has been well defined. Apart from acting as a repressor of competence and cell surface genes[27,28], Rok is also known to silence xenogeneic or horizontally acquired genes, which typically have a higher AT content compared to the host genome[29,30].

Previously, ChIP analysis of Rok revealed that it binds to more than 250 sites dispersed over the *B. subtilis* genome[26]. To determine why only a small subset of the Rok-binding sites lead to the formation of Rok clusters, we analyzed the distribution of high-affinity Rok-binding motifs across the genome. Rok is known to bind to an array of A/T-rich DNA, but shows the highest affinity toward AACTA- or TACTA-containing motifs[30]. We found that the 10-kb regions spanning the Rok clusters were significantly enriched in such high-affinity Rok-binding motifs (WWACTAW) when compared to other Rok-binding sites that do not participate in Rok cluster formation (Fig. 2a,b).

We wondered whether these Rok clusters might be visible in cells using a fluorescent Rok-fusion. Indeed, during the exponential growth phase we observed several discrete signals of Rok-GFP distributed over the cells (Fig. 2c). Notably, only two to four fluorescent spots could be observed in each cell during stationary growth phase (Fig. 2c). This suggests that either Rok dissociates from some of the Rok clusters, or the individual clusters join together, raising the striking possibility of long-range interactions between distant Rok clusters and the formation of large anchored chromosomal loops.

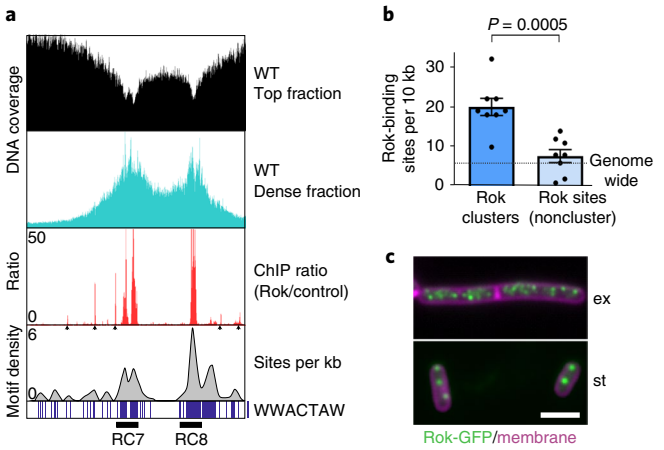


Fig. 2 | Rok clusters have more high-affinity Rok-binding motifs. **a**, Sedimentation-based DNA coverage maps near Rok clusters (RC) 7 and 8 along with Rok ChIP data[26]. High-affinity Rok-binding motif (WWACTAW) density (in gray) along with the individual sites marked as vertical blue lines is shown below. Other Rok-enriched ChIP sites near the Rok clusters are marked with arrow heads. **b**, Quantification of high-affinity Rok-binding motifs (WWACTAW) at the eight Rok clusters compared to eight other Rok-binding sites (mean values, s.e., two-tailed *t*-test). The average genome-wide distribution of the motif is shown using the dotted line. **c**, Visualization of Rok-GFP clusters during exponential (ex) and stationary (st) growth phases using fluorescence microscopy. Membrane was stained using Nile red. Scale bar, 2 µm.

Rok clusters interact to form large chromosomal loops. To test for possible interactions between Rok clusters, we performed Hi-C on wild type and a Δ*rok* strain grown to stationary phase (Fig. 3a). The Hi-C maps revealed a clear juxtaposition of the two chromosome arms by the *B. subtilis* SMC complex observed previously[1,17,20]. Importantly, we also observed specific interactions between several Rok clusters. These interactions were seen as peaks of interaction in the contact matrix (Fig. 3a, insets). The interactions between Rok clusters were completely lost in the Δ*rok* strain and were subsequently restored after *rok* complementation (with native promoter) at an ectopic locus (Fig. 3a and Extended Data Fig. 2).

Rok binds to the chromosome via its C-terminal DNA-binding domain, whereas the N-terminal domain is involved in multimerization[29,30]. This multimerization activity is likely important for the observed interactions between the Rok clusters. To test this, we performed Hi-C with two known multimerization-deficient Rok mutant strains (Rok mutants with truncations of the N-terminal domain) that still retain their ability to bind DNA using the C-terminal domain[29]. Indeed, both Rok mutant strains showed no long-range interactions between the Rok clusters (Extended Data Fig. 2).

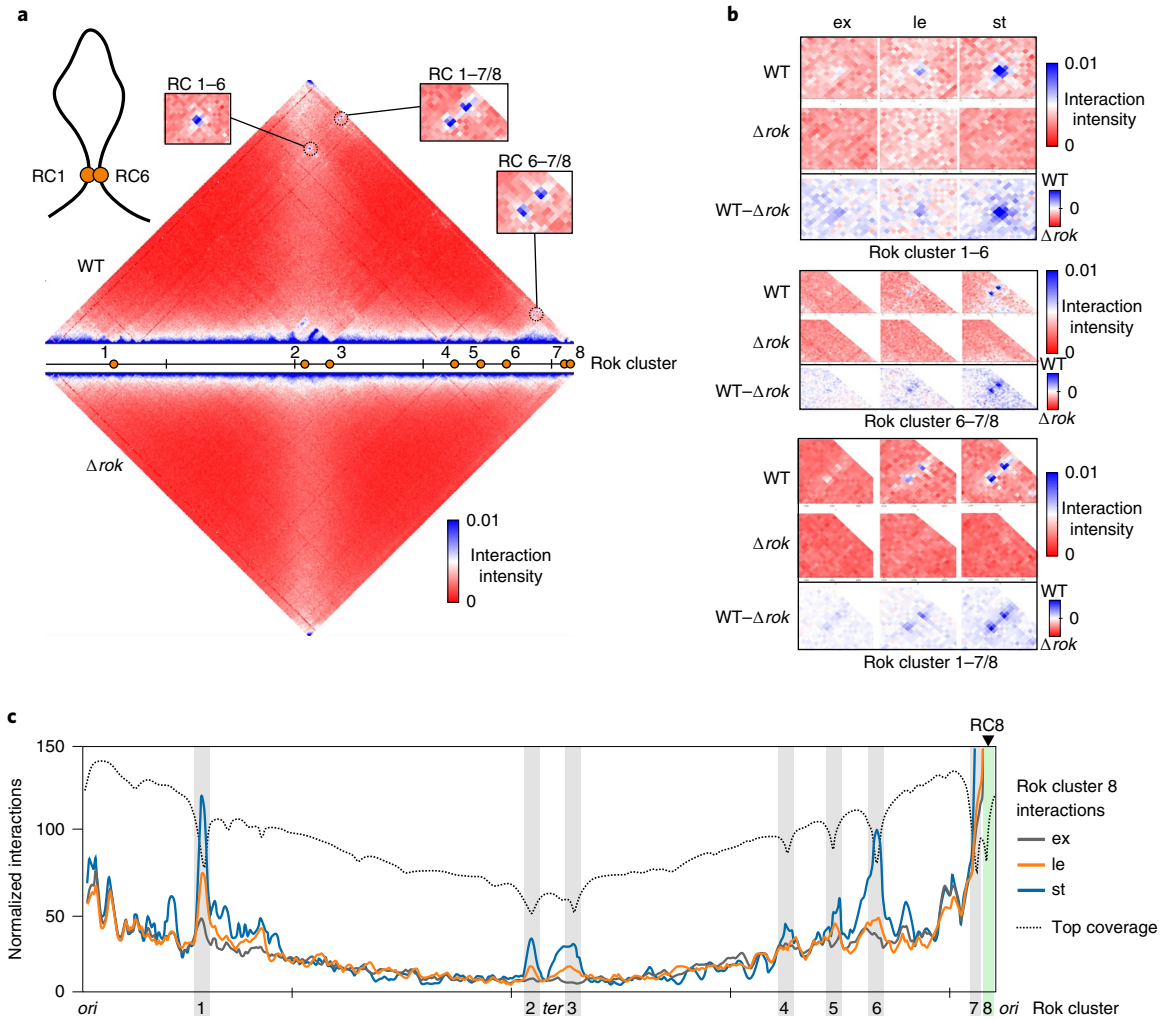


Fig. 3 | Rok clusters specifically interact with each other. a, Normalized Hi-C contact maps of wild type (top) and Δ*rok* strains (below) at stationary phase. Some of the corner peaks formed as a result of interaction between Rok clusters are shown in the insets. These peaks of interaction were absent in the Δ*rok* strains. **b**, Normalized Hi-C contact maps showing interactions between indicated Rok clusters in wild type and Δ*rok* strains along with the difference plot (WT – Δ*rok)* at indicated growth phases (ex, exponential; le, late exponential; st, stationary). **c**, Virtual 4C analysis to study interaction of Rok cluster 8 (marked using a green bar) with the whole genome at different growth phases. The other Rok clusters along the genome are marked using gray bars. DNA coverage of the top fraction obtained from the wild-type strain is shown as a dotted line. Genome positions in Mb intervals are marked below.

Furthermore, we also performed Hi-C on cells collected at exponential and late-exponential growth phase to check for changes in interaction between Rok clusters over growth. This revealed that the contact frequency between the Rok clusters increased gradually from exponential to stationary growth phase, complementing the fluorescence microscopy observations (Figs. 2c and 3b and Supplementary Fig. 1). Virtual 4C analysis uses Hi-C data to plot the interaction of a specific locus of interest (bait) with the whole genome. When we applied virtual 4C analysis with Rok cluster 8 as bait, a clear increase in interaction frequency with the other clusters from exponential to stationary phase was observed (Fig. 3c). No such changes in interactions were observed in the Δ*rok* strain (Supplementary Fig. 2).

It is possible that the active replisome complex can dissociate Rok clusters and thereby regulate chromosomal loop formation at different growth phases. To test this, we performed Hi-C on exponentially growing cells after arresting the replisome with the replication inhibitor hydroxyurea[31]. Indeed, clear interactions between Rok clusters were observed when replication was blocked (Extended Data Fig. 3), suggesting a potential role for DNA replication in modulating the long-range interactions by Rok.

Rok clusters interact within 'superclusters'. Virtual 4C analysis revealed that Rok clusters 1, 6, 7 and 8, which are near the origin, primarily interact with each other to form an 'origin supercluster', even though Rok clusters 1 and 6 are located around 1 Mb apart (Figs. 3a and 4a and Extended Data Fig. 4). Rok clusters 2 and 3 on the other hand are close to the terminus and interact with each other to form a 'terminus supercluster' (Fig. 4b and Extended Data Fig. 5). Rok cluster 4 lies between the origin and terminus, and showed no interaction with the other Rok clusters (Supplementary Fig. 3). In addition, we noticed that four other Rok-binding sites[26] were also recruited to their nearby supercluster during stationary growth, three at the origin supercluster and one at the terminus supercluster (Extended Data Figs. 5 and 6). It is likely that these sites also form Rok clusters, but our sedimentation approach was not sensitive enough to clearly identify them.

Relocated Rok cluster interacts with other Rok clusters. As a final validation that these chromosome loops were indeed formed by interactions between Rok clusters, we deleted Rok cluster 8 (3.4 kb region) and inserted it at an ectopic locus (*amyE* gene) located

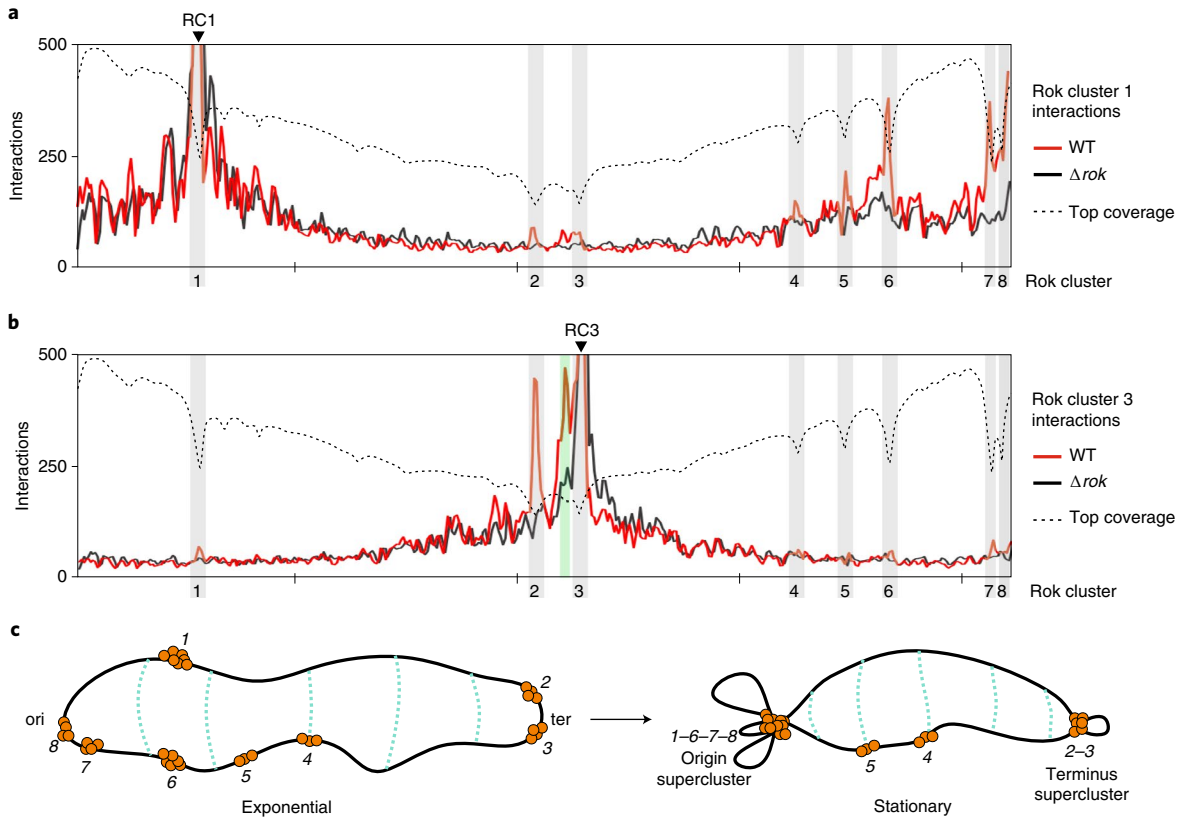


Fig. 4 | Rok clusters interact within superclusters. a,**b**, Interactions of Rok cluster 1 from origin supercluster (**a**) and Rok cluster 3 from terminus supercluster (**b**) with the whole genome during stationary phase in wild type and Δ*rok* strains. Rok clusters are marked using gray bars. A Rok-bound site (noncluster) also interacted with Rok clusters 2 and 3 and is marked using a green bar in (**b**), see Extended Data Fig. 5. **c**, Schematic illustration of Rok-dependent association of Rok clusters at exponential and stationary growth phases. Orange dots represent individual Rok proteins multimerized at the eight Rok clusters along the genome. Dotted cyan lines represent the SMC-mediated juxtaposition of the *B. subtilis* chromosome arms[17]. It is unknown if more than two Rok clusters can simultaneously interact within a cell.

~330 kb away (Fig. 5a). Rok cluster 8 encompasses only one operon encoding five genes (*yybN*, *yybM*, *yybL*, *yybK*, *yybJ*) of unknown function (Extended Data Fig. 7). We removed the promoter along with the start codon of the first gene in the operon (*yybN*) to disrupt transcription at the complemented locus and exclude its role in mediating interaction between Rok clusters. Hi-C on this strain showed that the relocated Rok cluster could still interact with the other Rok clusters, similar to the wild type situation (Fig. 5b,c). RNA-seq revealed transcription primarily for the first gene in the operon (*yybN*) in wild-type cells, which was abrogated in the relocated strain, excluding a direct role of transcription in mediating interaction between Rok clusters (Extended Data Fig. 7c). Furthermore, we also inserted the same Rok cluster 8 region at the *amyE* locus in the reverse orientation. Complementation of Rok cluster 8 in the reverse orientation was also found to interact with other Rok clusters, thereby excluding the role of binding polarity in mediating long-range interactions by Rok (Extended Data Fig. 7).

Validation of interaction between Rok clusters in live cells. Hi-C is performed on a large population of cells. To confirm that the Rok clusters interact in single live cells, we marked the DNA near Rok cluster 2 and 3 in the terminus supercluster using two fluorescent repressor operator systems (FROS) arrays (LacO and TetO), and visualized them individually using LacI and TetR fused to YFP and CFP, respectively. Indeed, Rok clusters 2 and 3 were found significantly closer together in wild-type cells compared to the *rok* deletion strain grown to stationary growth phase (Fig. 5d), supporting the Hi-C data. Introduction of Rok-mCherry in the above strain revealed that in a cell the signal for one of the Rok spots closely overlapped with Rok clusters 2 and 3 during stationary growth phase (Supplementary Fig. 4).

The SMC complex influences interaction between Rok clusters. In *B. subtilis*, the SMC complex is recruited to the ParB–*parS* nucleoprotein complexes near the origin of replication, where it tethers the two chromosome arms using its loop extrusion activity[1,17,20]. The SMC complexes are then unloaded near the replication terminus by the recombinase XerD[32]. The symmetrical interaction of Rok clusters along the origin–terminus axis suggests that the SMC complex may also influence the interactions between the Rok clusters by juxtaposition of the two chromosome arms[17] (Fig. 4c and Extended Data Fig. 8a). Hi-C on a strain lacking the functional SMC complex (*scpB* deletion strain) abrogated the long-range interactions between Rok clusters at stationary phase when compared with wild-type cells grown under the same conditions (Extended Data Fig. 8b), indicating that the alignment of chromosomal arms by the SMC complex is potentially important for mediating interactions between Rok clusters.

Previously, Rok was found to interact with DnaA, the initiator protein of DNA replication in *B. subtilis*[26]. To deduce a potential role of DnaA in Rok-mediated chromosomal interactions, Hi-C was performed on a *dnaA* deletion strain, where DNA replication was initiated by a plasmid-derived origin of replication (*oriN*) and its replication initiator protein (RepN)[33]. As several Rok clusters retained

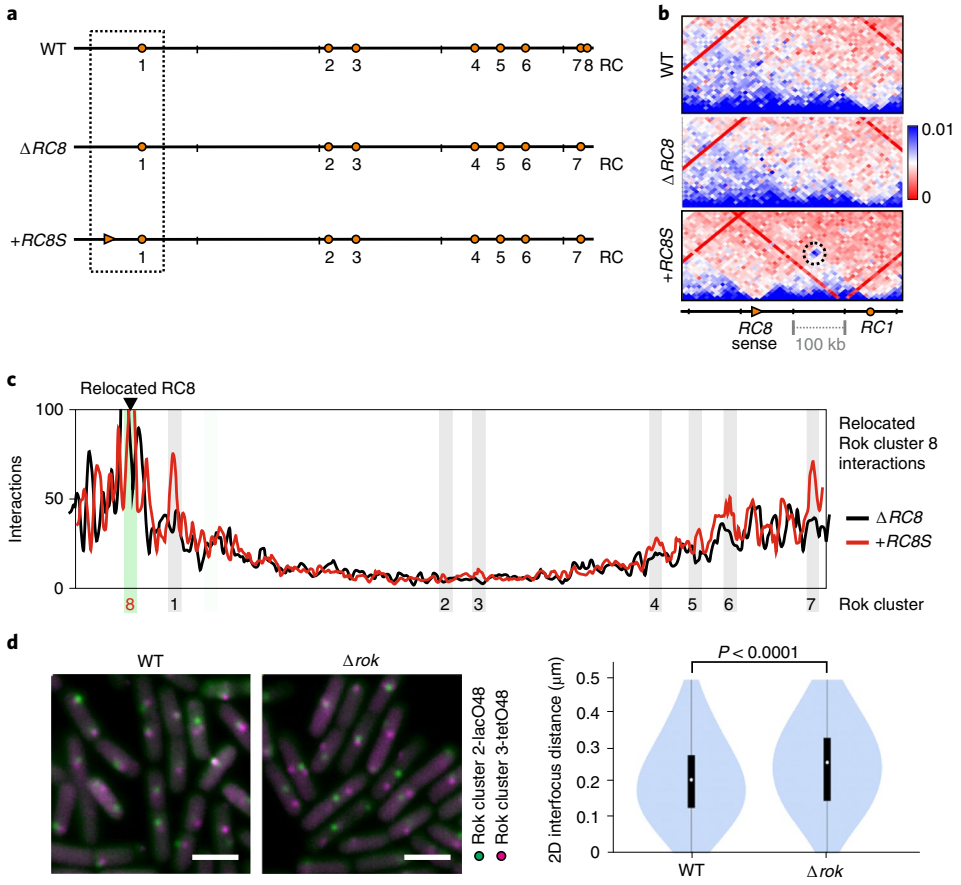


Fig. 5 | Verification of interaction between Rok clusters. a, Location of Rok clusters in the wild type and modified genomes. RC8 (Rok cluster 8) was deleted (Δ*RC8*) and complemented at an ectopic locus (*amyE*) within the origin supercluster in sense (+*RC8S*) orientation. **b**, Normalized Hi-C contact maps of wild type and the RC8 mutant strains near the complementation locus at stationary phase. **c**, Virtual 4C analysis to determine the interactions of the *amyE* locus (containing the complemented RC8) with the whole genome during stationary phase in Δ*RC8* and the complemented strain (+*RC8S*). Rok clusters are marked using gray bars and the RC8-complemented *amyE* locus using a green bar. **d**, Fluorescence microscopy-based analysis of the distance between Rok clusters 2 and 3 at stationary growth phase using FROS. 2D interfocal distances of 524 (wild type) and 554 (Δ*rok*) were used for the violin plot in the right panel (unpaired two-tailed *t*-test). The box extends from the 25th to 75th percentile with the line in the box at the median and whiskers from the smallest to largest value. Three biologically independent experiments were performed with similar results. Scale bar, 2 μm.

their interaction in this strain, DnaA was not found to be essential for chromosomal loop formation by Rok (Extended Data Fig. 9).

Interaction of Rok clusters leads to CID boundary formation. Hi-C experiments have shown that bacteria can also form TAD-like structures, which are termed CIDs[1,4,5,20]. CID boundaries are generally established at regions of highly expressed genes in bacteria[1,5,20]. Such domains are physically isolated regions of the genome that are known to co-regulate genes within a CID[4,34]. The long-range interactions we observed between Rok clusters resemble anchor sites for chromosome loops previously observed in eukaryotic Hi-C maps[5] (Fig. 3a). These anchor sites bound by insulator proteins are known to physically isolate chromosomal loops leading to TAD formation in eukaryotes[5,13]. Here, we examined whether interaction between Rok clusters can also influence CID formation in *B. subtilis*. Topological domain boundaries are routinely examined using directionality index (DI) analysis, which quantifies the bias in upstream or downstream interactions for a genomic region and changes abruptly at topological domain boundaries[35]. DI analysis showed that CID boundaries changed from exponential to stationary growth phase (Supplementary Fig. 5). Importantly, several domain boundaries located at Rok cluster sites were disrupted upon Rok deletion in stationary phase. Specifically, preventing the interaction between Rok clusters 2 and 3 (Fig. 4b and Extended Data Fig. 5), by deleting *rok*, resulted in disruption of the domain boundary at Rok cluster 3 (Fig. 6a, left panel). A similar disruption of domain boundaries was observed for Rok cluster 1 within the origin supercluster (Extended Data Fig. 10a). Furthermore, in mutants where Rok cluster 8 (3.4-kb-long region) was relocated ~330 kb away from the original location (in both sense and reverse orientation), new CID boundaries were formed at the new insertion site (Fig. 6b and Extended Data Figs. 7d and 10b). Similarly, interaction between Rok clusters within or near a CID may also increase the intra-CID interaction frequency by formation of self-interacting anchored loops, as shown for interaction between Rok clusters 6, 7 and 8 (Fig. 6a, right panel and Supplementary Fig. 6). This increase in interactions can be seen throughout the origin supercluster (Rok clusters 1, 6, 7 and 8) using the scalogram tool[4], which plots the cumulative contact frequency for each region as a function of genomic distance and hence reflects its relative tightness (Fig. 6a). Thus, interaction between Rok clusters can lead to the formation of new CIDs or modulate interactions within the existing ones.

Interaction of Rok clusters alter short-range interactions. We noticed that the frequency of short-range (~50-kb region) chromosome contacts, which are observed as the primary

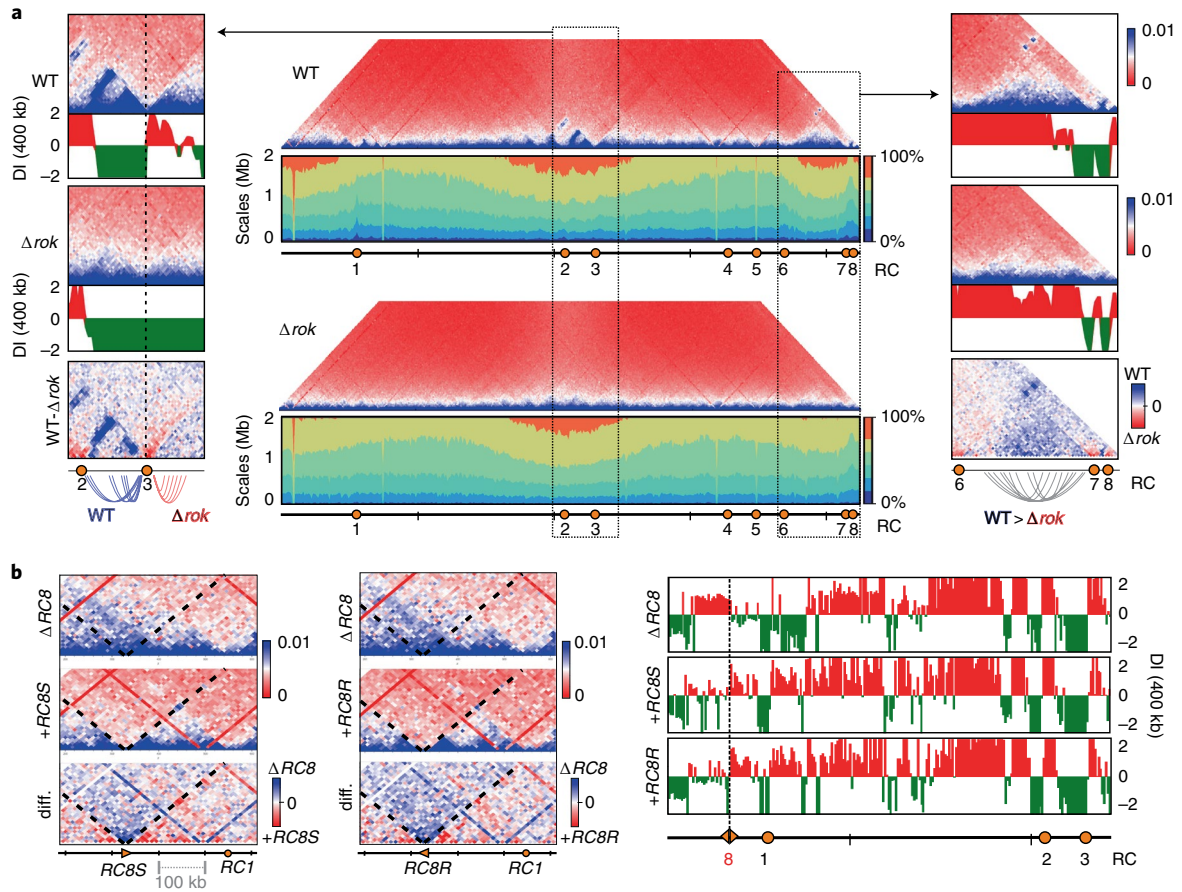


Fig. 6 | Interaction between Rok clusters impacts global chromosomal architecture. a, Normalized Hi-C contact maps of wild type and Δ*rok* strains at stationary phase along with scalogram representations[4]. Rok clusters (orange dots) are marked along the genome below (middle). Normalized Hi-C contact maps along with DI analysis shows higher order domain boundary at Rok cluster 3 (left) and increased intra-CID contact frequency between Rok clusters 6 and 7/8 (right). **b**, Normalized Hi-C contact maps and difference plots (diff.) of Rok cluster 8 (RC8) deletion strain together with its complementation at the ectopic *amyE* locus (left). The direction of complementation (sense + RC8S and reverse + RC8R orientation) is depicted by the orange arrowhead. DI analysis shows formation of new CID boundary at the relocated Rok cluster 8 site in both the complemented strains (right panel).

(horizontal) diagonal in the Hi-C maps, increased at the Rok cluster sites upon *rok* deletion (Fig. 7a,b). Moreover, we also observed an increase in short-range interactions in the Δ*rok* mutant at the three other Rok-binding sites (Fig. 7a and Extended Data Fig. 6, green dots), which were shown to interact with the origin supercluster. Interestingly, no changes were observed at Rok cluster 4, which does not interact with the other Rok clusters (Fig. 7a and Supplementary Fig. 3). These data indicate that long-range interactions between Rok clusters restrict their local interaction with the neighboring regions (Fig. 7b). These changes in short-range chromosome contacts are known to strongly correlate with gene expression in bacteria[4].

To test whether the genomic organization by Rok influenced transcription, we performed RNA-seq analysis at exponential and stationary growth phase of the wild type and Δ*rok* strain, and measured the effect on transcription of hundreds of genes around Rok clusters (Fig. 7c). In the absence of Rok, a modest (less than twofold) but significant upregulation in expression of genes around Rok clusters 1 and 3 was observed in the stationary growth phase when Rok clusters interact most strongly, whereas there was no clear difference in localized expression of these adjacent genes in exponential phase (Fig. 7c).

Discussion

Rok was first identified as a global transcription repressor involved in the development of genetic competence, expression of cell surface genes, biofilm and mobile genetic elements in *B. subtilis*[27–29,36]. Here, we have shown that Rok also mediates long-range chromosomal interactions leading to formation of anchored chromosomal loops and CID boundaries. Such protein-anchored Mb-range interactions have not been demonstrated in bacteria and raise important questions regarding their evolution and functionality.

The formation of large anchored loops by Rok resembles the looping activity of insulator proteins found in eukaryotes. In mammals, CCCTC-binding factor (CTCF) is the main insulator protein and has been extensively studied for its role in regulation of chromatin architecture[37,38]. CTCF interacts with specific regions of the genome and forms chromatin loops by obstructing the cohesin complex as it travels bidirectionally along the DNA, thereby extruding and segregating large sections of chromatin[37,38]. Insulator proteins including CTCF also play important roles in preventing enhancer–promoter interactions and heterochromatin spread in eukaryotes[38]. Although Rok shares some characteristics of insulator proteins, as it also spatially isolates large regions of the chromosome, the other roles of an insulator protein, in the context of eukaryotic enhancer–promoter interactions and heterochromatin spreading, are not translatable to bacteria.

Insulator proteins including the CTCF/cohesin complex interact with other proteins to regulate their activities[39–43]. Rok may function in concert with other cofactors, but this remains to be investigated, although we did observe abrogation of interactions between Rok clusters in a *B. subtilis* strain lacking a functional SMC complex.

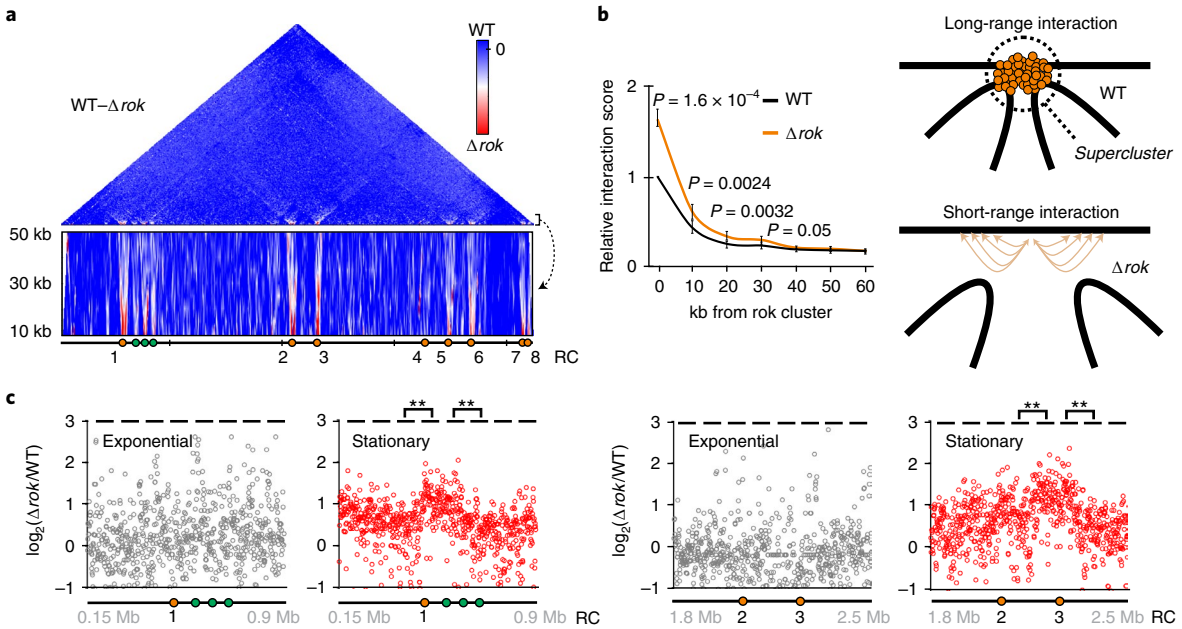


Fig. 7 | Interaction between Rok clusters alters local chromosomal architecture and transcription. a, Hi-C difference plot of wild type and Δ*rok* strains at stationary phase, the color scheme of the plot is shifted to highlight regions of increased interactions (white and red) upon *rok* deletion. The magnified view of short-range contacts between 10 kb and 50 kb along with the Rok clusters (orange dots) and other interacting Rok-binding sites (green dots) are shown below. **b**, Average relative short-range interaction of the eight Rok clusters using normalized Hi-C data from wild type and Δ*rok* strains in Fig. 3a (mean values, s.e., paired two-tailed *t*-test). Hi-C was performed twice with similar results. The illustration at the right shows how elimination of long-range interactions between Rok clusters upon *rok* deletion increases the local short-range interaction. **c**, Changes in expression of genes around Rok clusters 1 and 2/3 upon *rok* deletion is visualized by plotting the changes in mRNA abundance (Δ*rok*/WT) at exponential (gray) and stationary (red) phase. For statistical analysis, transcription changes in groups of 100 genes (black bar) were compared to the adjacent groups to determine the effect of genome position on transcription (unpaired two-tailed *t*-test, ** $P < 0.0001$).

However, this is different from the CTCF/cohesin complex, as Rok does not directly impede the movement of the SMC complexes. It is more likely that the juxtaposition of chromosomal arms by the SMC complex assists in interaction between Rok clusters by promoting physical proximity between the clusters. In a previous study, Rok was also found to interact with DnaA[26]; however, we show here that DnaA is not essential for Rok-mediated loop formation.

In bacteria, CID boundaries are typically found at regions of highly expressed genes[1,4,14,20]. In *C. crescentus*, the relocation of a highly expressed gene (*rsaA*) present at a CID boundary was found to create a sharp new CID boundary at a new location[1]. Our results demonstrate that chromosomal loop formation by Rok and subsequent CID boundary formation can also be relocated to other regions of the genome, and, importantly, this process appears to be independent of local transcription within the Rok cluster. Although topological domains in both eukaryotes and prokaryotes contribute to transcriptional changes[5,34], TAD formation by CTCF/cohesin complex can in some cases be dispensable for developmental gene regulation while contributing to response to external stimuli[44]. The direct role of Rok-mediated loop formation in transcription remains unclear, as Rok also acts as a transcription repressor for several genes in *B subtilis*[28]. In this regard it would be interesting to see if Rok-mediated chromosome remodeling may also contribute to mitigate environmental stressors by regulating transcriptional changes. Finally, insertion of a 2-kb boundary element containing a CTCF-binding site was shown to generate a new TAD boundary in human cells[45]. Likewise, it can be envisaged that incorporation of Rok clusters together with Rok protein may potentially be used to create topological domains in heterologous systems.

Bacteria are thought to lack the long-range promoter regulation that is observed in eukaryotes. A well-known DNA loop-forming protein in Gram-negative bacteria is integration host factor, but this protein facilitates the formation of small DNA loops (sub kb) by directly bending DNA to regulate expression[46]. Active loop extrusion by SMC complex and the related homologs do form large chromosomal loops in bacteria; however, these loops are dynamic and hence not anchored at specific regions of the genome[1,4,47]. ParB is known to deform the local DNA structure by locally spreading to the adjacent DNA (~15 kb) near the *parS* sites[48,49], but loop formation of this large nucleoprotein complex remains undefined[50]. Relatively large artificially anchored DNA loops of 10–20 kb, using DNA-binding proteins such as LacI and Cas9, have been shown to mediate promoter regulation over a distance[51,52]. However, long-range (0.2–2 Mb) anchored loop formation, as shown for Rok, has not been observed before in bacteria. Rok is only found within *Bacillus* species. Nevertheless, several other nucleoid-associated proteins, including H-NS and Lsr2, are also known to possess in vitro DNA-bridging activity[53–55], and such proteins may serve as potential candidates for chromosomal loop formation in bacteria. However, unlike Rok, these proteins have relatively low sequence specificity, which may hamper the formation of anchored chromosomal loops at specific sites[4,30]. Other bacterial transcriptional factors, such as LacI and GalR, that oligomerize and show high specificity for DNA sequences are more likely candidates for anchored loop formation in bacteria[56,57]. In this regard, our sedimentation approach, together with chromosome conformation capture methods such as Hi-C, may help unravel the diversity and functionality of chromosomal loop anchors in prokaryotes.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary

information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-021-00988-8.

Received: 21 May 2021; Accepted: 19 November 2021;
Published online: 24 January 2022

References

1. Le, T. B. K., Imakaev, M. V., Mirny, L. A. & Laub, M. T. High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* **342**, 731–734 (2013).
2. Toro, E. & Shapiro, L. Bacterial chromosome organization and segregation. *Cold Spring Harb. Perspect. Biol.* **2**, a000349 (2010).
3. Wang, X., Llopis, P. M. & Rudner, D. Z. Organization and segregation of bacterial chromosomes. *Nat. Rev. Genet.* **14**, 191–203 (2013).
4. Lioy, V. S. et al. Multiscale structuring of the *E. coli* chromosome by nucleoid-associated and condensin proteins. *Cell* **172**, 771–783 (2018).
5. Szabo, Q., Bantignies, F. & Cavalli, G. Principles of genome folding into topologically associating domains. *Sci. Adv.* **5**, eaaw1668 (2019).
6. Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).
7. Sexton, T. et al. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**, 458–472 (2012).
8. Le Gall, A., Valeri, A. & Nollmann, M. Roles of chromatin insulators in the formation of long-range contacts. *Nucleus* **6**, 118–122 (2015).
9. Ong, C. T. & Corces, V. G. CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* **15**, 234–246 (2014).
10. Parelho, V. et al. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433 (2008).
11. Brasset, E. & Vaury, C. Insulators are fundamental components of the eukaryotic genomes. *Heredity* **94**, 571–576 (2005).
12. Rubin, A. J. et al. Lineage-specific dynamic and pre-established enhancer-promoter contacts cooperate in terminal differentiation. *Nat. Genet.* **49**, 1522–1528 (2017).
13. Dixon, J. R., Gorkin, D. U. & Ren, B. Chromatin domains: the unit of chromosome organization. *Mol. Cell* **62**, 668–680 (2016).
14. Le, T. B. & Laub, M. T. Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries. *EMBO J.* **35**, 1582–1595 (2016).
15. Dame, R. T., Rashid, F. Z. M. & Grainger, D. C. Chromosome organization in bacteria: mechanistic insights into genome structure and function. *Nat. Rev. Genet.* **21**, 227–242 (2019).
16. Hirano, T. Condensin-based chromosome organization from bacteria to vertebrates. *Cell* **164**, 847–857 (2016).
17. Wang, X., Brandão, H. B., Le, T. B. K., Laub, M. T. & Rudner, D. Z. *Bacillus subtilis* SMC complexes juxtapose chromosome arms as they travel from origin to terminus. *Science* **355**, 524–527 (2017).
18. Lioy, V. S., Junier, I., Lagage, V., Vallet, I. & Boccard, F. Distinct activities of bacterial condensins for chromosome management in *Pseudomonas aeruginosa*. *Cell Rep.* **33**, 108344 (2020).
19. Takemata, N., Samson, R. Y. & Bell, S. D. Physical and functional compartmentalization of archaeal chromosomes. *Cell* **179**, 165–179 (2019).
20. Marbouty, M. et al. Condensin- and replication-mediated bacterial chromosome folding and origin condensation revealed by Hi-C and super-resolution imaging. *Mol. Cell* **59**, 588–602 (2015).
21. Schoenfelder, S. & Fraser, P. Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* **20**, 437–455 (2019).
22. West, A. G., Gaszner, M. & Felsenfeld, G. Insulators: many functions, many mechanisms. *Genes Dev.* **16**, 271–288 (2002).
23. Brunwasser-Meirom, M. et al. Using synthetic bacterial enhancers to reveal a looping-based mechanism for quenching-like repression. *Nat. Commun.* **7**, 10407 (2016).
24. Smirnov, A. et al. Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. *Proc. Natl Acad. Sci. USA* **113**, 11591–11596 (2016).
25. Fernandez-Martinez, J., LaCava, J. & Rout, M. P. Density gradient ultracentrifugation to isolate endogenous protein complexes after affinity capture. *Cold Spring Harb. Protoc.* https://doi.org/10.1101/pdb.prot087957 (2016).
26. Seid, C. A., Smith, J. L. & Grossman, A. D. Genetic and biochemical interactions between the bacterial replication initiator DnaA and the nucleoid-associated protein Rok in *Bacillus subtilis*. *Mol. Microbiol.* **103**, 798–817 (2017).
27. Hoa, T. T., Tortosa, P., Albano, M. & Dubnau, D. Rok (YkuW) regulates genetic competence in *Bacillus subtilis* by directly repressing comK. *Mol. Microbiol.* **43**, 15–26 (2002).
28. Albano, M. et al. The rok protein of *Bacillus subtilis* represses genes for cell surface and extracellular functions. *J. Bacteriol.* **187**, 2010–2019 (2005).
29. Smits, W. K. & Grossman, A. D. The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in *Bacillus subtilis*. *PLoS Genet.* **6**, e1001207 (2010).
30. Duan, B. et al. How bacterial xenogeneic silencer rok distinguishes foreign from self DNA in its resident genome. *Nucleic Acids Res.* **46**, 10514–10529 (2018).
31. Slager, J., Kjos, M., Attaiech, L. & Veening, J. W. Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. *Cell* **157**, 395–406 (2014).
32. Karaboja, X. et al. XerD unloads bacterial SMC complexes at the replication terminus. *Mol. Cell* **81**, 756–766 (2021).
33. Murray, H. & Koh, A. Multiple regulatory systems coordinate DNA replication with cell growth in *Bacillus subtilis*. *PLoS Genet.* **10**, e1004731 (2014).
34. Trussart, M. et al. Defined chromosome structure in the genome-reduced bacterium *Mycoplasma pneumoniae*. *Nat. Commun.* **8**, 14665 (2017).
35. Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).
36. Kovács, Á. T. & Kuipers, O. P. Rok regulates yuaB expression during architecturally complex colony development of *Bacillus subtilis* 168. *J. Bacteriol.* **193**, 998–1002 (2011).
37. Hadjur, S. et al. Cohesins form chromosomal *cis*-interactions at the developmentally regulated *IFNG* locus. *Nature* **460**, 410–413 (2009).
38. Ghirlando, R. & Felsenfeld, G. CTCF: making the right connections. *Genes Dev.* **30**, 881–891 (2016).
39. Hu, G. et al. Systematic screening of CTCF binding partners identifies that BHLHE40 regulates CTCF genome-wide distribution and long-range chromatin interactions. *Nucleic Acids Res.* **48**, 9606–9620 (2020).
40. Liu, Z., Scannell, D. R., Eisen, M. B. & Tjian, R. Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* **146**, 720–731 (2011).
41. Qiu, Z. et al. Functional interactions between NURF and Ctcf regulate gene expression. *Mol. Cell. Biol.* **35**, 224–237 (2015).
42. Donohoe, M. E., Zhang, L. F., Xu, N., Shi, Y. & Lee, J. T. Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. *Mol. Cell* **25**, 43–56 (2007).
43. Guastafierro, T. et al. CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery. *J. Biol. Chem.* **283**, 21873–21880 (2008).
44. Stik, G. et al. CTCF is dispensable for immune cell transdifferentiation but facilitates an acute inflammatory response. *Nat. Genet.* **52**, 655–661 (2020).
45. Zhang, D. et al. Alteration of genome folding via contact domain boundary insertion. *Nat. Genet.* **52**, 1076–1087 (2020).
46. Goosen, N. & van de Putte, P. The regulation of transcription initiation by integration host factor. *Mol. Microbiol.* **16**, 1–7 (1995).
47. Wang, X. et al. Condensin promotes the juxtaposition of DNA flanking its loading site in *Bacillus subtilis*. *Genes Dev.* **29**, 1661–75 (2015).
48. Soh, Y. M. et al. Self-organization of parS centromeres by the ParB CTP hydrolase. *Science* **366**, 1129–1133 (2019).
49. Graham, T. G. W. et al. ParB spreading requires DNA bridging. *Genes Dev.* **28**, 1228–38 (2014).
50. Funnell, B. E. ParB partition proteins: complex formation and spreading at bacterial and plasmid centromeres. *Front. Mol. Biosci.* **3**, 44 (2016).
51. Hao, N., Shearwin, K. E. & Dodd, I. B. Programmable DNA looping using engineered bivalent dCas9 complexes. *Nat. Commun.* **8**, 1628 (2017).
52. Hao, N., Shearwin, K. E. & Dodd, I. B. Positive and negative control of enhancer-promoter interactions by other DNA loops generates specificity and tunability. *Cell Rep.* **26**, 2419–2433 (2019).
53. Qin, L., Erkelens, A. M., Markus, D. & Dame, R. T. The *B. subtilis* Rok protein compacts and organizes DNA by bridging. Preprint at *bioRxiv* https://doi.org/10.1101/769117 (2019).
54. van der Valk, R. A. et al. Mechanism of environmentally driven conformational changes that modulate H-NS DNA-bridging activity. *eLife* **6**, e27369 (2017).
55. Chen, J. M. et al. Lsr2 of *Mycobacterium tuberculosis* is a DNA-bridging protein. *Nucleic Acids Res.* **36**, 2123–2135 (2008).
56. Cournac, A. & Plumbridge, J. DNA looping in prokaryotes: experimental and theoretical approaches. *J. Bacteriol.* **195**, 1109–1119 (2013).
57. Semsey, S., Tolstorukov, M. Y., Virnik, K., Zhurkin, V. B. & Adhya, S. DNA trajectory in the Gal repressosome. *Genes Dev.* **18**, 1898–1907 (2004).

Methods

***B. subtilis* strains and growth.** *B. subtilis* strains were grown on LB agar plates supplemented with appropriate antibiotics: spectinomycin (150 μg ml^{-1}), erythromycin (2 μg ml^{-1}) or chloramphenicol (5 μg ml^{-1}). For liquid culture, *B. subtilis* strains were inoculated at optical density (OD_{600}) of 0.005 from an overnight culture and grown at 37 °C in LB medium or Spizizen minimal medium (SMM)[58].

Strain construction. All strains were constructed using transformation and homologous recombination of an overlap PCR product or plasmid. The overlap PCR product or plasmid contained the antibiotic resistance gene and the given insertion or mutation between ~1,000 bp of homologous region on either side. pHJS105 plasmid was used as the base plasmid for insertion/complementation at the *amyE* locus[59]. All strains constructed in this study were verified by whole-genome sequencing or colony PCR. The strains used in this study are listed in Supplementary Table 1. Primers used for the construction of strains are listed in Supplementary Table 4.

Sedimentation-based DNA coverage assay. Top fraction, *B. subtilis* cultures were grown (final OD_{600} of ~0.6) in 200 ml of LB medium. All the cells were collected by centrifugation at 10,000*g* for 5 min and the pellet was immediately frozen in liquid nitrogen. The pellets were stored at −80 °C until use. The pellet was transferred to a 20- or 50-ml stainless steel canister (Retsch) precooled in liquid nitrogen and containing 1 ml 1× PBS (with protease inhibitor). The pellet was cryogenically broken using five rounds of disruption in a TissueLyser II (20 Hz for 2 min, each round). The canister was cooled in liquid nitrogen after each round. The pulverized sample was retained from the canister and 1 ml ice-cold 1× PBS (with protease inhibitor) was added to the sample. Portions (5 μl each) of 10× Fragmentase Reaction Buffer v2 and dsDNA fragmentase (NEB, catalog no. M0348S) were added to the samples and mixed by short vortexing. The samples were incubated at 30 °C for 15 min to mildly fragment the DNA (>200 bp fragments). The sample was then added to the top of two-step sucrose density layers (20% and 60%) and ultracentrifuged for 2 h at 30,000 r.p.m. (SW41 rotor, Beckman) and 4 °C. The top fraction (100 μl) was collected for DNA extraction.

Dense fraction, *B. subtilis* cultures were grown (final OD_{600} of 0.6) in 200 ml LB medium. The samples were treated with rifampicin (100 μM final) for 10 min while shaking to abort transcription before collection. Treatment with rifampicin eliminates transcription-dependent protein–DNA complexes. Cells were then fixed using 0.1% formaldehyde (final) for 10 min at room temperature to stabilize complexes, followed by quenching using glycine. Cell pellets were stored and processed as before for 'top fraction' but this time the fraction (750 μl) at the interphase of 20% and 60% sucrose density layers was collected using a syringe introduced by puncturing the side of the tube. This fraction was diluted to 10% sucrose density using ice-cold 1× PBS (with protease inhibitor) and again loaded on top of two-step sucrose density layers (20% and 60%) and ultracentrifuged for 2 h at 30,000 r.p.m. (SW41 rotor, Beckman) and 4 °C. The dense fraction (500 μl) at the interphase of the 20% and 60% sucrose density layers was collected by puncturing the side of the tube using a syringe.

Nucleic acids were extracted from the fractions collected above (top and dense) using phenol/chloroform/isoamyl alcohol (PCI; Carl Roth, catalog no. A156.3) and the nucleic acid was concentrated using ethanol precipitation. Nucleic acid was resuspended in water and the RNA was removed using RNase A. The DNA was again extracted and concentrated using PCI (Carl Roth, catalog no. A156.3) and ethanol, respectively. The partially fragmented and purified DNA was directly used for library preparation using a NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB, catalog no. E7645S) as per manufacturer's instructions, and subsequently subjected to sequencing using an Illumina NextSeq 550. The details of libraries produced are listed in Supplementary Table 2.

RNA isolation and RNA-seq. Total RNA was extracted from *B. subtilis* cells using the hot-phenol method, as previously described[60], except 5 mg ml^{-1} lysozyme was used to disrupt the *B. subtilis* cell wall. Residual DNA was removed from the total RNA using DNase I (NEB, catalog no. M0303S) digestion as per manufacturer's instructions. Total RNA was checked on agarose gel to verify the integrity of the RNA samples by visualizing the ribosomal RNA bands. rRNA was then removed from 10 μg of total RNA using MICROBExpress Bacterial mRNA Enrichment Kit (Thermo Fisher, catalog no. AM1905) or from 1 μg of total RNA using NEBNext rRNA depletion kit for bacteria (NEB, catalog no. E7850) as per manufacturer's instructions. The RNA was then extracted using PCI for RNA (Carl Roth, catalog no. X985.3) and precipitated using ethanol. A 20–100-ng portion of the rRNA-depleted RNA was used for library preparation using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, catalog no. E7760S), which was subsequently sequenced using an Illumina NextSeq 550. The details of libraries produced, along with the rRNA removal kit used, are listed in Supplementary Table 2.

Mapping and visualization of DNA coverage maps, RNA-seq and ChIP–seq datasets. DNA and RNA sequencing data used in this study were generated using the experiments described above. Raw Rok ChIP data generated by Seid et al.[26] were obtained from NCBI (accession number PRJNA272948). All the sequencing data were processed using the open source web-based platform Galaxy[61]. The quality of the FASTQ files was first assessed using fastQC and checked for basic statistics and per base/sequence quality scores before further downstream processing. All files were then trimmed using Trimmomatic (Galaxy v.0.36.5 without illuminaclip step, sliding window trimming with number of bases to average across (4) and average quality required (20))[62] before mapping them to the *B. subtilis* ssp. *subtilis* str. 168 reference genome (NC_000964.3) using Bowtie2 (Galaxy v.2.3.4.2 with default setting only)[63]. The BAM files were converted to bigwig using the tool bamCoverage (Galaxy v.3.0.2.0, bin size 5 and normalization to reads per kilobase per million)[64] and the coverage maps were visualized using Integrated Genome Browser (IGB v.9.0.2). The RNA-seq BAM files were also used as the input for featureCounts (Galaxy v.1.6.3)[65] along with the *B. subtilis* 168 general feature format (gff3, accession no. ASM904v1) file to quantify and compare gene expression. Normalized coverage files were used to generate DNA coverage and ChIP ratio plots directly using IGB (v.9.0.2). The details of libraries produced are listed in Supplementary Table 2.

Motif density analysis. The 10-kb sequence around each Rok cluster (5 kb upstream and downstream of local minima from the top fraction of sedimentation-based DNA coverage assay, Supplementary Table 3) along with eight highly enriched Rok ChIP sites (not involved in cluster formation, 5 kb from each side of Rok ChIP[26] peak maxima at *metS, rpoB, citA, yjcN, yosW, yggA, braB* and *sboA* genes) were extracted for analysis. The extracted sequences were quantified for presence of high-affinity Rok-binding motifs: 'WWACTAW'. 'WACTA' was identified previously as a high-affinity motif of Rok using protein-binding microarray[30]. The distribution of the motif along the genome was directly visualized in IGB (v.9.0.2). *B. subtilis* ssp. *subtilis* str. 168 reference genome (NC_000964.3) was used to quantify the average abundance of the motif.

Chromosome capture by Hi-C. Hi-C was carried out essentially as previously described[66] with minor modifications. Cultures were grown as described above and samples were collected at different growth phases (exponential, 5 h; late exponential, 10 h and stationary, 22 h) in LB or SMM media. For replication arrest, culture at exponential phase was treated with hydroxyurea (1 mg ml^{-1}) for 75 min. Briefly, 2–5 ml culture was sequentially fixed using 80% methanol and 3% formaldehyde. Cells were washed with ice-cold 1× PBS after each step. The cells were then collected by centrifugation and the pellets were flash frozen in liquid nitrogen. Cells were lysed using Ready-Lyse Lysozyme (Epicentre, catalog no. R1802M) in 1× TE buffer followed by 0.5% SDS treatment. The chromosomal DNA in the cell lysate was then digested using HindIII for 3 h at 37 °C. The restriction ends were filled with biotin-14-dCTP, dGTP, dATP, dTTP using DNA polymerase I, large (Klenow) fragment (NEB, catalog no. M0210S). The chromatin was fractionated by centrifugation and subsequently the pellet was ligated using T4 DNA ligase (NEB, catalog no. M0202M) overnight at 16 °C. Samples were then treated with RNase A and de-cross-linked at 65 °C for 6 h in the presence of proteinase K. The DNA was then extracted using PCI (Carl Roth, catalog no. A156.3) and precipitated using ethanol. Biotin was removed from the nonligated ends using T4 DNA polymerase (M0203S) in the presence of dATP and dGTP. Then, the DNA was again extracted and precipitated as before, and then fragmented using dsDNA fragmentase (NEB, catalog no. M0348S) treatment for 15 min at 37 °C. The fragmented DNA was used for library preparation using the NEBNext Ultra II DNA Library Prep Kit for Illumina (NEB catalog no.E7645S) as per manufacturer's instructions until adapter ligation and purification using AMPure XP beads (Beckman Coulter, catalog no. A63881). Biotinylated library fragments were extracted from the sample using 20 μl of Dynabeads MyOne Streptavidin T1 beads (Thermo Fisher, catalog no. 65601) as per manufacturer's instructions. The washed beads (with biotinylated DNA) were used for PCR library amplification (8–10 cycles) using NEBNext Ultra II Q5 Master Mix (NEB, catalog no. M0544S). The amplified library was purified using AMPure XP beads followed by paired-end sequencing using an Illumina NextSeq 550. The details of libraries produced in this study are listed in Supplementary Table 2.

Hi-C data mapping and contact matrix. Hi-C matrices were constructed using the Galaxy HiCExplorer webserver: https://hicexplorer.usegalaxy.eu[67]. Briefly, paired-end reads were mapped separately to *B. subtilis* genome (NCBI Reference Sequence NC_000964.3) using very sensitive local setting mode in Bowtie2 (Galaxy v.2.3.4.2). The mapped files were used to build the contact matrix using the tool hicBuildMatrix (Galaxy v.2.1.2.0) using a bin size of 10 kb, HindIII restriction site (AAGCTT) and AGCT as the dangling sequence. The contact matrix (.cool format) was then used for further analysis and visualization as described below.

Hi-C data visualization. *Contact maps.* Hi-C contact maps were assessed, compared and prepared for the illustration using the interactive browser-based visualization tool 'Bekvaem'[68] whose source code is published and made available to the public (https://doi.org/10.11588/data/KGYOS6).

Comparison of contact maps. First, Hi-C contact frequency matrices were normalized using the Sinkhorn–Knopp balancing algorithm[69]. In a subsequent step, the normalized contact probability matrices were compared via their difference, that is, two contact probability matrices $A \equiv (a_{ij})$ and $B \equiv (b_{ij})$ were

compared via their difference $D \equiv (d_{ij})$ by (element-wise) matrix subtraction $d_{ij} = a_{ij} - b_{ij}$. To be able to detect only very small local differences also, the (element-wise) logarithmic ratio was computed as $R \equiv r_{ij} = \log_2|a_{ij}/b_{ij}|$.

Domain detection. Domain boundaries within Hi-C contact maps were detected algorithmically using an in-house Python implementation based on the directionality index approach presented by Dixon et al.[35]. The method is motivated by the observation that domains are demarcated by regions that are biased in their interaction probability; the upstream domain boundary is preferentially interacting downstream whilst the downstream boundary is preferentially interacting upstream. This method is based on a two-step strategy. First, the two-dimensional (2D) contact information is translated into the directionality index encoding the ratio of downstream and upstream interactions. Next, downstream interactions are compared to upstream interactions to derive whether the strength of interactions are significantly stronger in one direction compared to the other. Domain boundaries are characterized in that this preferred direction of interactions changes abruptly[1].

Scalograms. Scalograms visualizing the dispersion of the contact probability signal along the spatial scales were implemented in Python by following the description of Lioy et al.[4] on GitHub (https://github.com/koszullab/E_coli_analysis/). Scalograms visualize the extent to which genomic regions are affected by neighboring sequences. For each genomic position, the number of contacts between its position and an increasing number of neighboring bins is cumulated. Subsequently, the resulting heat map is depicted using a contour line function. In this illustration, the number of cumulated contacts is divided into five color-coded areas. This representation helps to understand the intrachromosomal contact probability $P(s)$ as a function of the genomic distance, which can be modeled using a power law. Therefore, it helps in understanding the polymeric nature of chromosomes on a local scale.

For virtual 4C analysis, the bait region (10-kb bin) containing the Rok cluster was used as input for the tool hicPlotDistVsCounts (Galaxy v.2.1.4.0) in the Galaxy HiCExplorer web server. The total contacts of the input bait region with all bins were quantified for a Hi-C matrix and normalized for comparison between different conditions (genotype and growth phase). Average short-range interaction around Rok clusters was also quantified similarly for Fig. 7a.

FROS and image analysis. LacO48 (kan) and TetO48 (cat) arrays were amplified from the strain BWX1200 (ref. [70]) and integrated near Rok clusters 2 and 3, respectively. *tetR-CFP* and *lacI-mYPet* (from BWX1200) were then introduced in the above strain to visualize the integrated arrays (Strain GD201, Supplementary Table 1). For imaging, cells from stationary growth phase were immobilized on 1% agarose (in 1× PBS) and imaged with a Hamamatsu ORCA-Flash-4.0LT CMOS camera mounted on an Olympus BX-60 fluorescence microscope. The spots (fluorescence maxima) of the arrays within each cell were determined using Coli-Inspector[71] by using manually configured fluorescence thresholds. The distance between CFP and mYPet (Yfp derivative) spots was then determined and analyzed for pairs of spots that were present in the same cell and were less than 0.5 µm apart.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All raw and processed sequencing datasets generated during this study can be accessed at Gene Expression Omnibus repository under accession number GSE144475.

Code availability

All source code used in this study has been published before and is referenced in the Methods section.

References

58. Spizizen, J. Transformation of biochemically deficient strains of *Bacillus subtilis* by deoxyribonucleate. *Proc. Natl Acad. Sci. USA* **44**, 1072–1078 (1958).
59. Jahn, N., Brantl, S. & Strahl, H. Against the mainstream: the membrane-associated type I toxin BsrG from *Bacillus subtilis* interferes with cell envelope biosynthesis without increasing membrane permeability. *Mol. Microbiol.* **98**, 651–66 (2015).
60. Dugar, G. et al. High-resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. *PLoS Genet.* **9**, e1003495 (2013).
61. Afgan, E. et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46**, W537–W544 (2018).
62. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
63. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
64. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–165 (2016).
65. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
66. Crémazy, F. G. et al. Determination of the 3D genome organization of bacteria using Hi-C. *Methods Mol. Biol.* **1837**, 3–18 (2018).
67. Wolff, J. et al. Galaxy HiCExplorer: a web server for reproducible Hi-C data analysis, quality control and visualization. *Nucleic Acids Res.* **46**, W11–W16 (2018).
68. Hofmann, A., Müggenburg, J., Crémazy, F. & Heermann, D. W. Bekvaem: integrative data explorer for Hi-C data. *J. Bioinform. Genom.* https://doi.org/10.18454/jbg.2019.2.11.1 (2019).
69. Sinkhorn, R. & Knopp, P. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific J. Math.* **21**, 343–348 (1967).
70. Wang, X., Llopis, P. M. & Rudner, D. Z. *Bacillus subtilis* chromosome organization oscillates between two distinct patterns. *Proc. Natl Acad. Sci. USA* **111**, 12877–12882 (2014).
71. Vischer, N. O. E. et al. Cell age dependent concentration of *Escherichia coli* divisome proteins analyzed with ImageJ and ObjectJ. *Front. Microbiol.* **6**, 586 (2015).

Acknowledgements

We thank R. Dame and F. Z. Rashid for providing critical input on Hi-C protocol, S. van Leeuwen (MAD, University of Amsterdam) for providing excellent sequencing services and N. Vischer for help with FROS image analysis. This work was funded by EMBO, ALTF 936–2016 (G.D.), European Commission MCSA-IF grant no. 749510 (G.D.), Netherlands Organization for Scientific Research (NWO) VENI grant-VI.VENI.192.103 (G.D.) and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC-2181/1–390900948, the Heidelberg STRUCTURES Cluster of Excellence (D.W.H.).

Author contributions

G.D. conceived the project, designed and performed all experiments (including Hi-C library preparation, analysis and visualization), analyzed data and wrote the manuscript. A.H. and D.W.H. performed Hi-C data analysis and visualization, and provided input on manuscript preparation. L.W.H. conceived the project, analyzed data and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

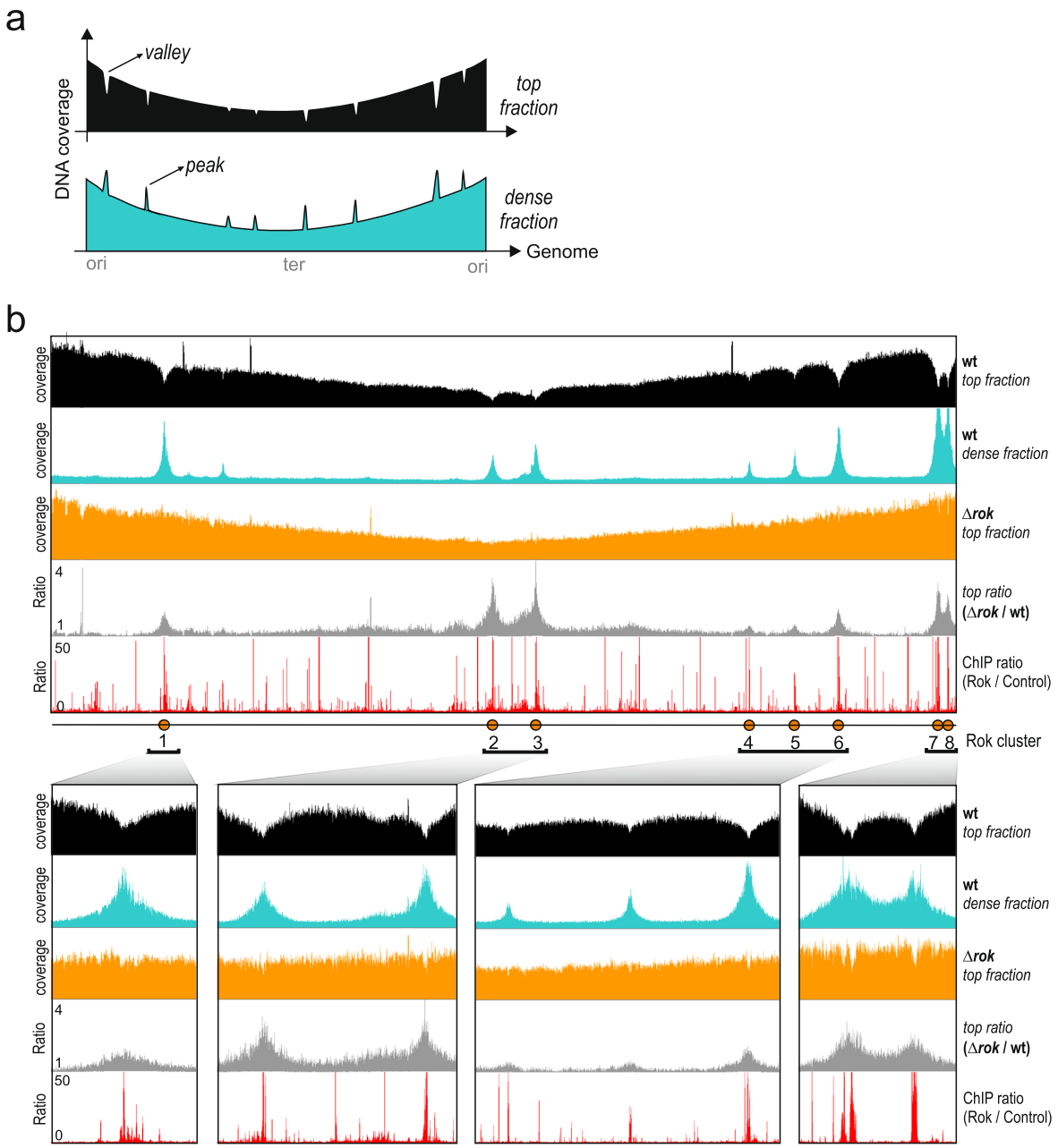
Extended data is available for this paper at https://doi.org/10.1038/s41588-021-00988-8.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41588-021-00988-8.

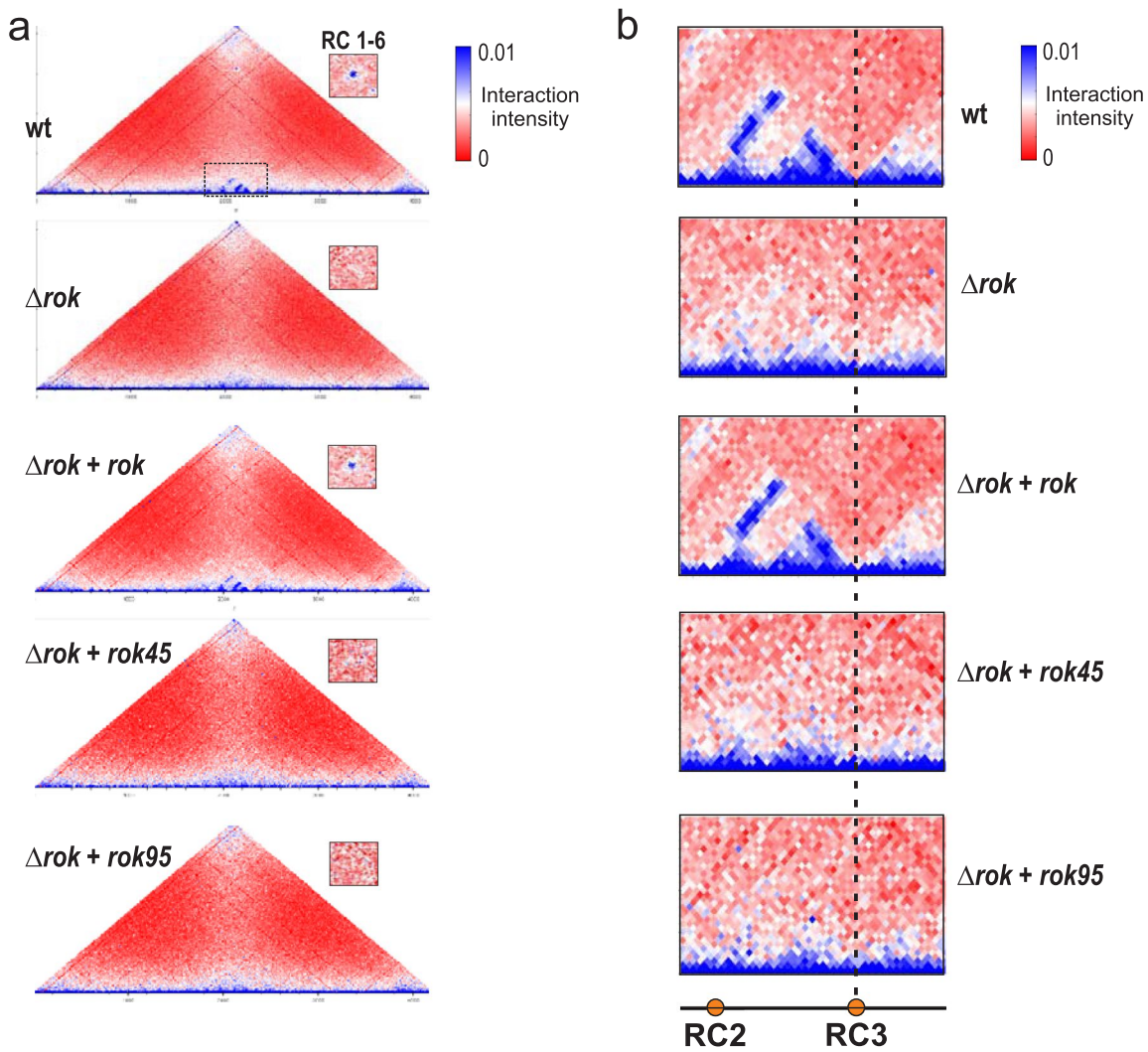
Correspondence and requests for materials should be addressed to Gaurav Dugar or Leendert W. Hamoen.

Peer review information *Nature Genetics* thanks Marcelo Nollmann, Jesse Dixon and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

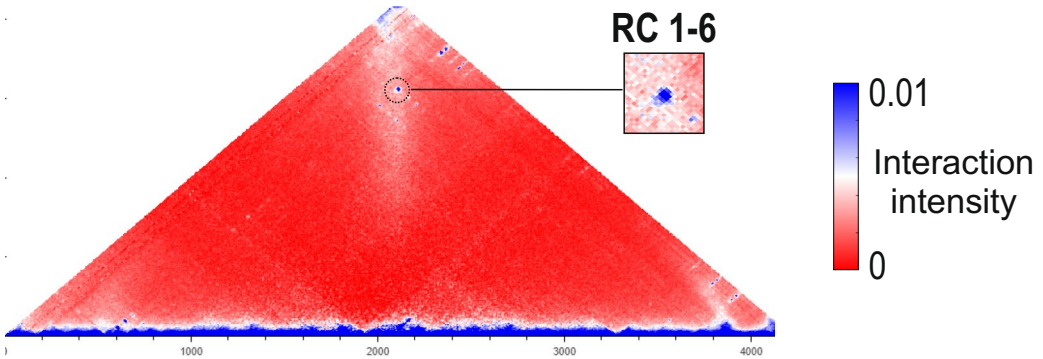
Reprints and permissions information is available at www.nature.com/reprints.



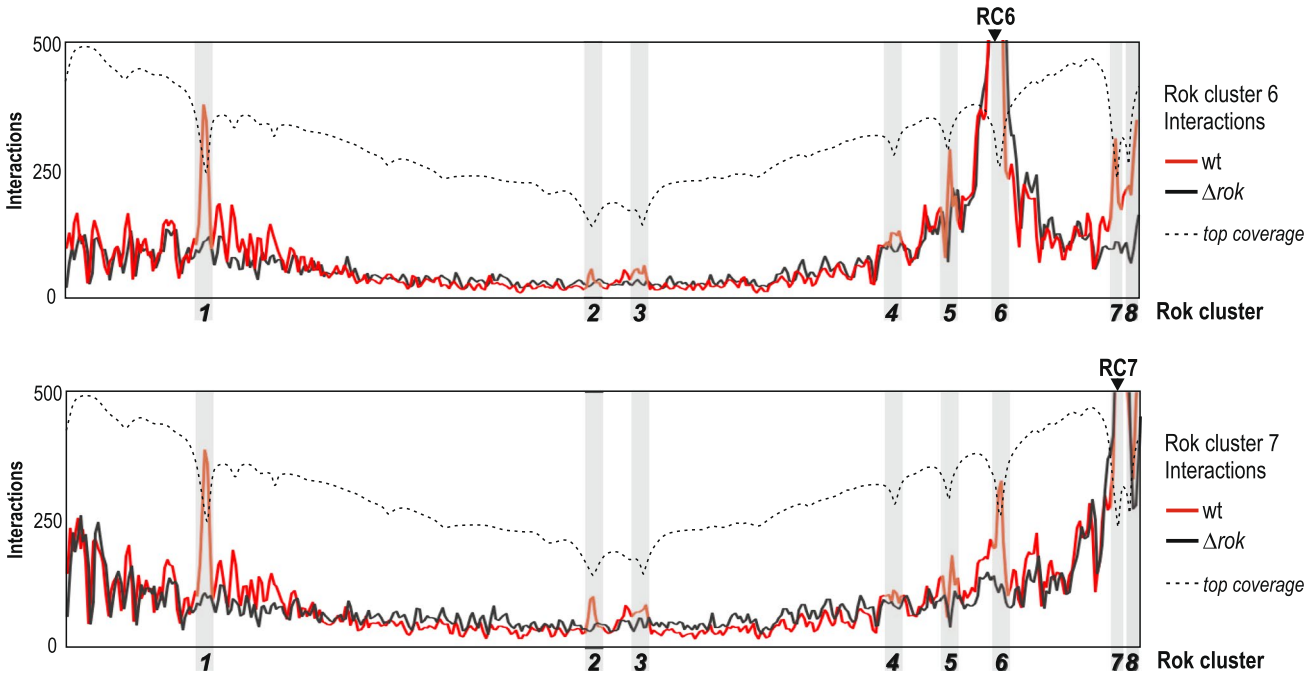
Extended Data Fig. 1 | Sedimentation based DNA coverage maps and identification of Rok associated large DNA complexes. a) Illustration of the expected sedimentation based DNA coverage maps. **b)** DNA coverage maps from the top fraction of wild type (black) and Δ*rok* (orange) strains. The ratio of top fraction coverage plots (Δ*rok*/wt) shows the Rok clusters as peaks (grey). DNA coverage of DNA obtained from the dense fraction is shown in cyan. Rok clusters are defined at sites where both local minima and local maxima are observed in DNA coverage of top fraction and dense fraction, respectively (Supplementary Table 3). Rok ChIP data in red is also shown along the coverage files.



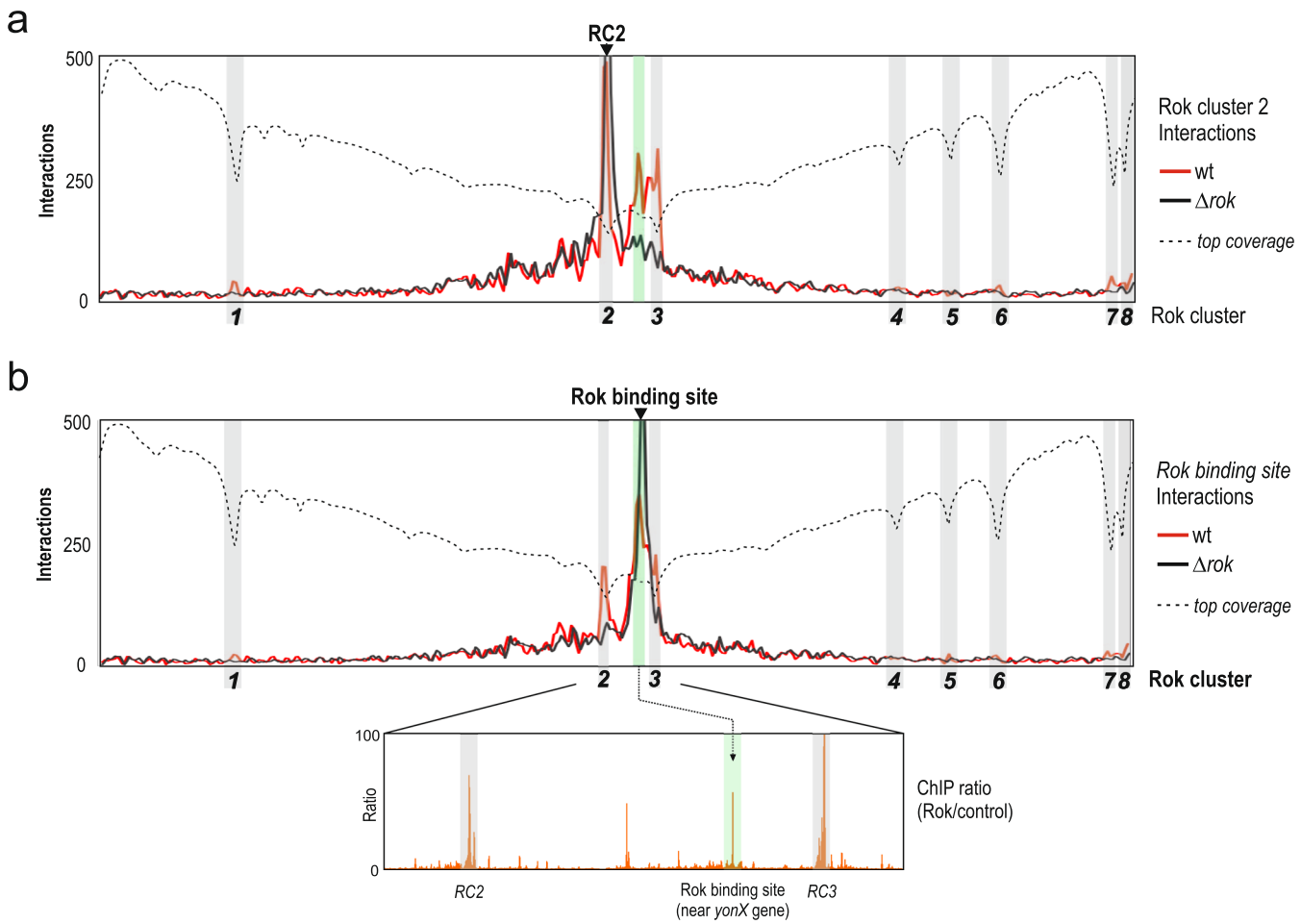
Extended Data Fig. 2 | Rok multimerization is essential for interaction between Rok clusters. a) Normalized Hi-C contact maps of wild type, *rok* deletion (Δ*rok*), *rok* complementation (Δ*rok* + *rok*), complementation with *rok* mutant with 45 aa truncation from the N-terminus (Δ*rok* + *rok45*) and complementation with *rok* mutant with 95 aa truncation from the N-terminus (Δ*rok* + *rok95*). The interaction between Rok clusters 1 and 6 is shown in the inset for each strain. **b)** Interaction between Rok terminus supercluster in the wild type and the different *rok* mutant strains. The region shown in b) is marked in the Hi-C map of wild type strain in a).



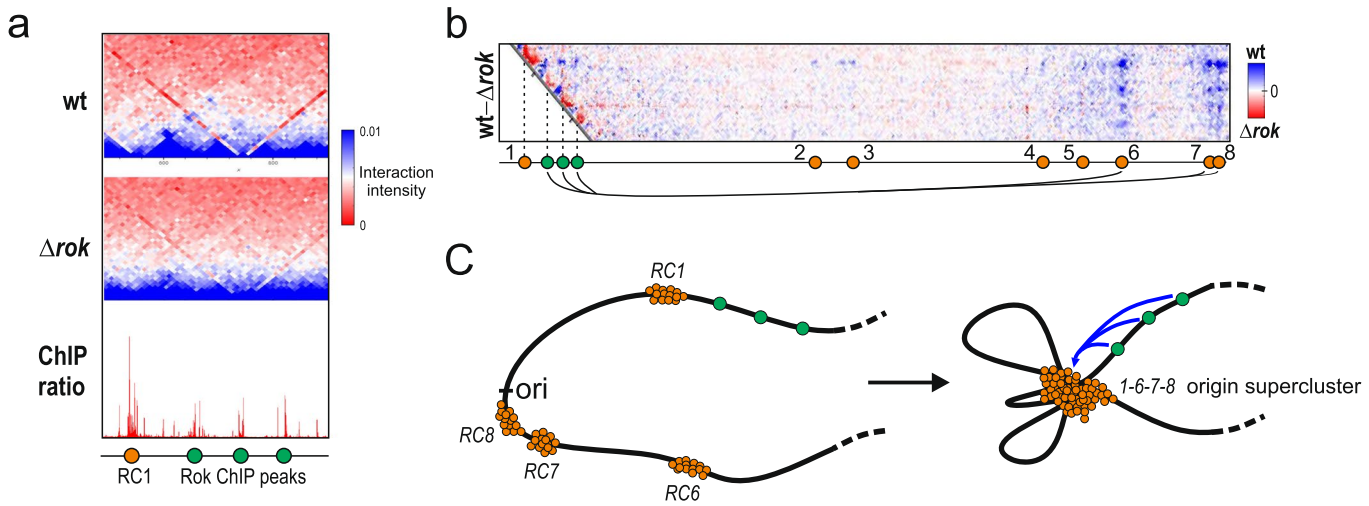
Extended Data Fig. 3 | Inhibition of replication promotes Rok mediated chromosomal loop formation in exponential phase. Normalized Hi-C contact maps of wild type strain at exponential phase after treatment with the replication inhibitor hydroxyurea (1 mg/ml). The interaction between Rok clusters 1 and 6 is shown in the inset. Hydoxyurea shows some inhibition of the SMC complex (reduced contacts in the secondary vertical diagonal showing juxtaposition of the two chromosome arms), presumably since this complex traverses from the origin to the terminus and encounters arrested replisomes.



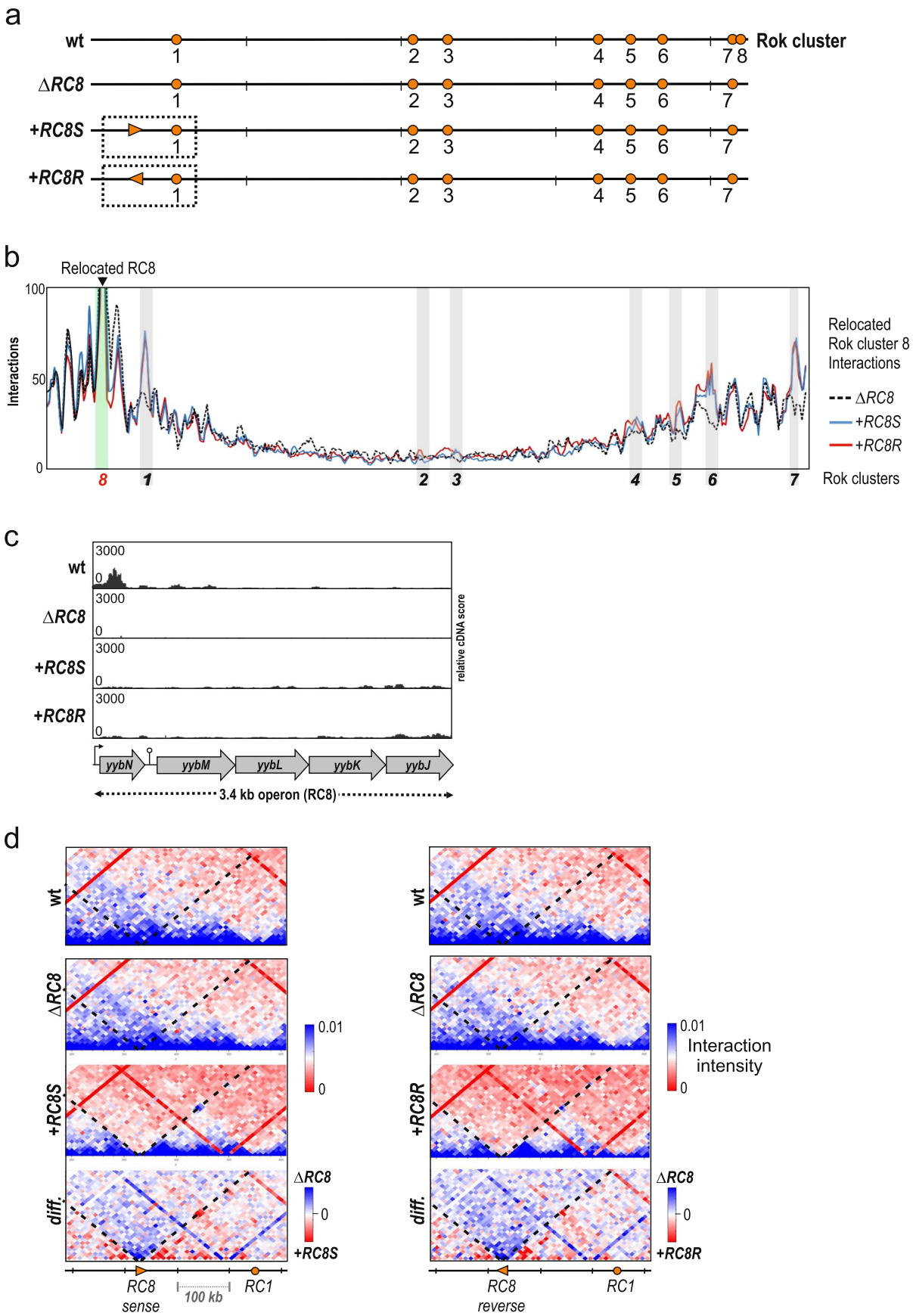
Extended Data Fig. 4 | Virtual 4 C analysis of Rok clusters at the origin supercluster. Interactions of Rok clusters 6 and 7 with the whole genome during stationary phase in wild type (wt) and Δ*rok* strains. Rok clusters are marked using grey bars. DNA coverage of top fraction obtained from the wild type strain is shown as dotted line.



Extended Data Fig. 5 | Virtual 4 C analysis of Rok clusters at the terminus supercluster. a) Interactions of Rok cluster 2 with the whole genome during stationary phase in wild type and Δ*rok* strains. Rok clusters are marked using grey bars. **b)** Interactions of a Rok binding site (located between Rok cluster 2 and 3, near *yonX* gene) with the whole genome during stationary phase in wild type and Δ*rok* strains. This site was found to interact with both Rok cluster 2 and 3 (see Fig. 4b) and is recruited to the Rok terminus supercluster during stationary phase. Rok ChIP data[1] (orange) at the terminus supercluster is shown below. The Rok binding site is marked using green bar. DNA coverage of top fraction obtained from the wild type strain is shown as dotted line.

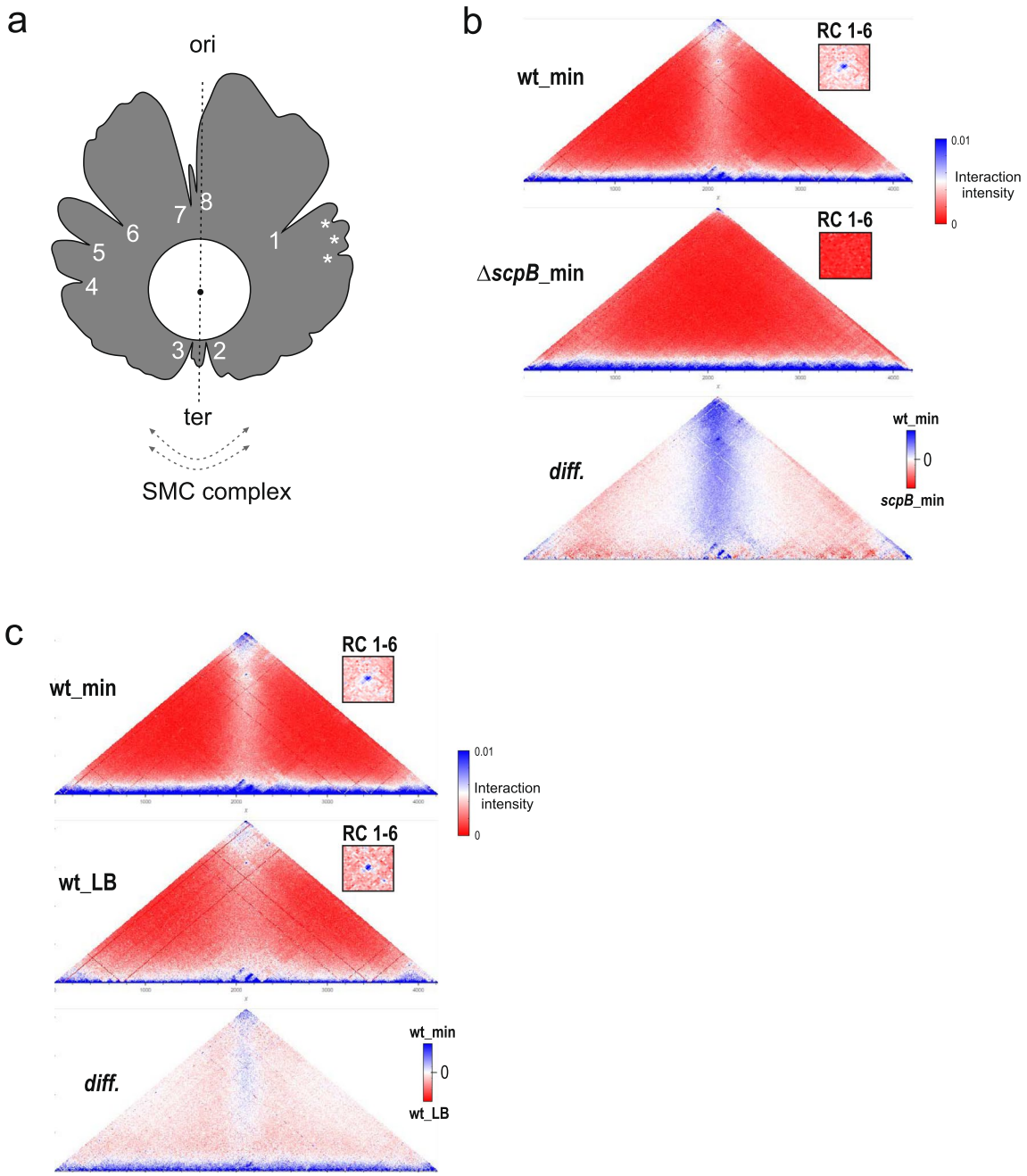


Extended Data Fig. 6 | Three other Rok ChIP sites interact with the Rok origin supercluster. a) Normalized Hi-C contact maps of wild type and Δ*rok* strains near Rok cluster 1 at stationary phase. Rok ChIP data (red) is also shown along the genome below highlighting the other Rok binding sites as green dots. **b)** Difference plot shows Rok dependent interaction of three other Rok binding sites (green dots) with Rok clusters (1, 6,7 and 8) from the origin supercluster. **c)** Illustration shows association of Rok clusters 1,6,7 and 8 to form the origin supercluster and their interaction with the three nearby Rok-binding sites recruited to the origin supercluster at stationary phase.

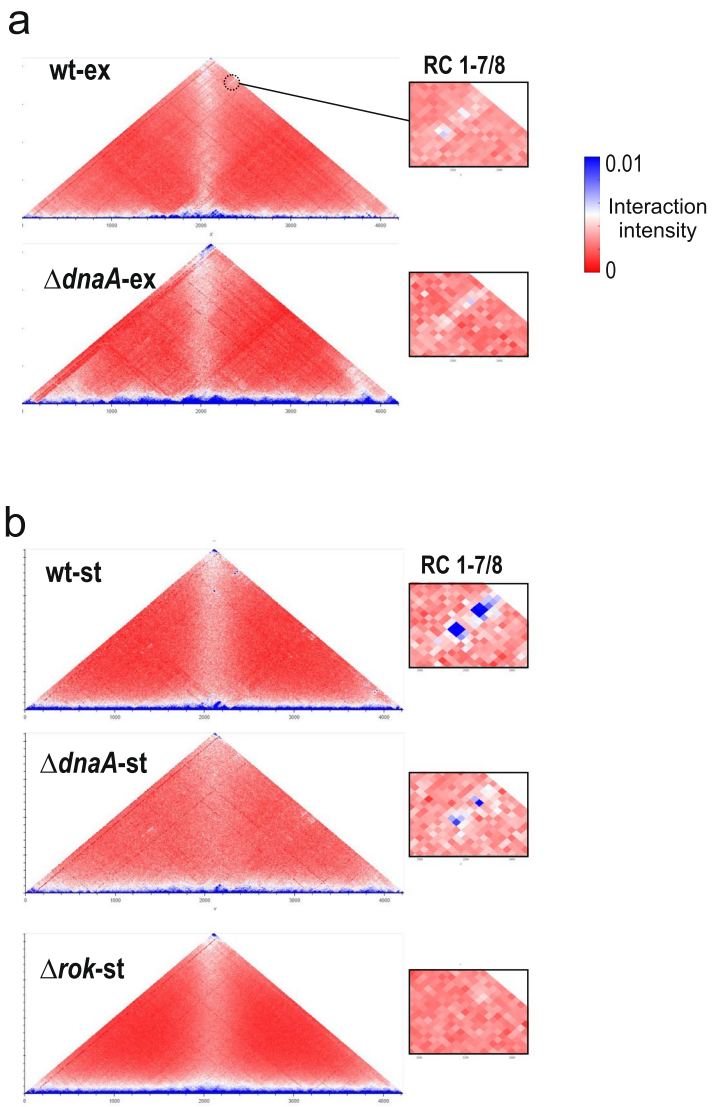


Extended Data Fig. 7 | See next page for caption.

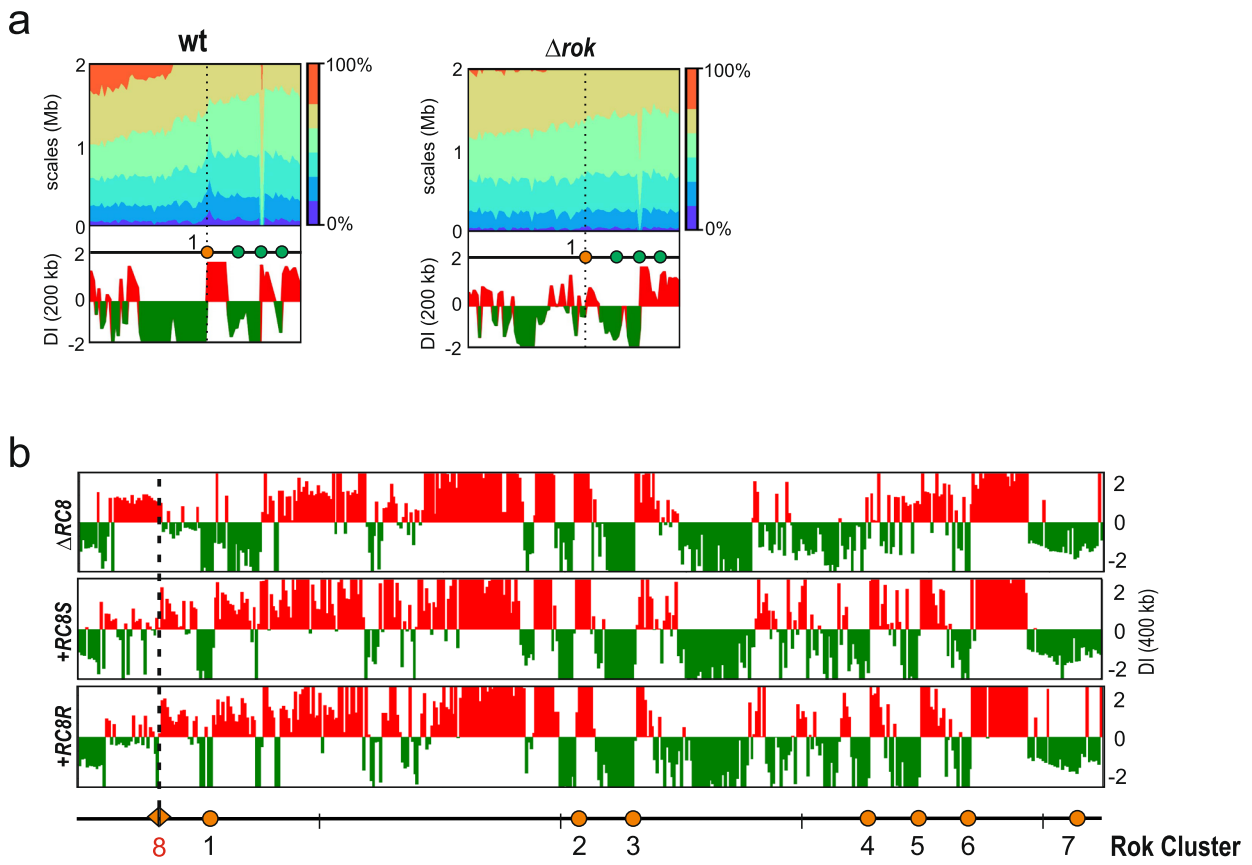
Extended Data Fig. 7 | Relocated Rok cluster 8 (RC8) interacts with other Rok clusters. a) Location of Rok clusters in the wild type and modified genomes. RC8 was deleted and complemented at an ectopic locus (*amyE*) within the origin supercluster in both sense (+RC8S) and reverse (+RC8R) orientation. **b)** Virtual 4 C analysis to determine the interactions of the *amyE* locus (containing the complemented RC8) with the whole genome during stationary phase in Δ*RC8* and the complemented strains. Rok clusters are marked using grey bars and RC8 complementation at the *amyE* locus is marked using a green bar. **c)** Screenshot showing relative cDNA reads (RNA-seq data) mapped to RC8 in the wild type, RC8 deletion and the RC8 complementation strains (without *yybN* promoter). The transcription start site and the terminator around the *yybN* gene obtained from *Subti*Wiki is marked in the annotation below. **d)** Normalized Hi-C contact maps and difference plots of wild type and the RC8 mutant strains near the *amyE* complementation locus at stationary phase. The genomic regions shown is boxed in a).



Extended Data Fig. 8 | SMC complex mediates Rok cluster interactions. a) Circular representation of DNA coverage from top fraction of differential sedimentation assay (wild type coverage from Fig. 1) after smoothing. The other three Rok binding sites which were found to interact with the origin supercluster (see Extended Data Fig. 6) are marked using an asterisk. **b)** Normalized Hi-C contact maps of wild type (wt) and Δ*scpB* strains along with the difference plot at stationary growth phase in minimal media (SMM). The SMC protein forms a homodimer, and together with the kleisen protein ScpA and the kite protein ScpB it forms the SMC complex. Deletion of *scpB* inactivates the SMC-complex. The interaction between Rok clusters 1 and 6 is shown in the inset for both strain. The Δ*scpB* strain is only viable when grown in minimal medium. However, Rok cluster are also formed in wild type strain during stationary growth phase in minimal medium **(c)**.



Extended Data Fig. 9 | Role of DnaA in Rok mediated chromosomal loop formation. a) Normalized Hi-C contact maps of wild type and *dnaA* deletion strains at exponential phase. **b)** Normalized Hi-C contact maps of wild type, *dnaA* deletion and *rok* deletion strains at stationary phase. The interaction between Rok clusters 1 and 7/8 is shown in the inset.



Extended Data Fig. 10 | Topological domains boundaries are formed by interaction between Rok clusters. a) Scalogram and DI analysis (200 Kb scale) of wild type and Δ*rok* strains near Rok cluster 1 and 3 other Rok binding sites (green dots) recruited to the origin supercluster at stationary growth phase (see Extended Data Fig. 6). **b)** DI analysis (400 Kb scale) of RC8 mutant strains (see Extended Data Fig. 7) at stationary growth phase. The RC8 complementation locus is marked using a dotted line.

nature research

Corresponding author(s): Gaurav Dugar, Leendert W. Hamoen

Last updated by author(s): Nov 14, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection
Data analysis	All data analysis steps are described in detail in the Methods section. The following were used for sequencing data analysis and visualization- Galaxy webserver - https://usegalaxy.org/ -Trimmomatic (Galaxy Version 0.36.5) -Bowtie2 (Galaxy version 2.3.4.2) -BamCoverage (Galaxy version 3.0.2.0) -FeatureCounts (Galaxy Version 1.6.3) HiCExplorer webserver - https://hicexplorer.usegalaxy.eu -Bowtie2 (Galaxy version 2.3.4.2) -hicBuildMatrix (Galaxy version 2.1.2.0) -hicPlotDistVsCounts (Galaxy version 2.1.4.0) Bekvaem - https://doi.org/10.11588/data/KGYOS6 Scalogram Analysis- https://github.com/koszullab/E_coli_analysis/ Integrated Genome Browser (IGB version 9.0.2) Python (3.8) For microscopy image analysis Coli-Inspector (https://sils.fnwi.uva.nl/bcb/objectj/examples/Coli-Inspector/Coli-Inspector-MD/coli-inspector.html) was used.

GraphPad Prism 8.3.0 (538) was used for statistical analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

-Bacillus subtilis subsp. subtilis str. 168 reference genome file rom NCBI - NC_000964.3
-Bacillus subtilis subsp. subtilis str. 168 annotation file (gff3) from NCBI - accession no. ASM904v1
-Rok ChIP data generated by Seid et. al. (Molecular Microbiology 2017) was obtained from NCBI (accession number PRJNA272948)
-All raw and processed sequencing datasets generated during this study can be accessed at Gene Expression Omnibus (GEO) repository under the accession number GSE144475

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical method was used to determine sample size. Sample size was determined based on general variability of the datasets in published literature for chromosome capture in bacteria (like Lioy et. al., Cell 2018).
Data exclusions	No data were excluded from the analyses.
Replication	Two replicates were performed to generate Hi-C data on wild type and Rok mutant in Bacillus subtilis at stationary growth phase to define the interaction between Rok clusters (both replicates show similar results). All other sequencing data on mutants was performed once as only qualitative inference was made (presence of interaction between Rok clusters or not). Mapping statistics of all sequencing libraries are shown in Supplementary Table 2. All microscopy experiments were conducted atleast twice with similar results.
Randomization	Randomization of chromosome capture analysis is not standard in the field
Blinding	The investigators were not blinded during data collection and analysis. Knowledge of labels was necessary for experiment design and analysis.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging