Euryarchaeal genomes are folded into SMC-dependent loops and domains, but lack transcription-mediated compartmentalization

Charlotte Cockram, Agnès Thierry, Aurore Gorlas, Roxane Lestini, Romain Koszul

Version of Record: https://www.sciencedirect.com/science/article/pii/S1097276520309059
Manuscript_5653ca1173c538b6e81d37d825ddeb65

Gene resolution HiC reveals Euryarchaeal genomes are folded into SMC-dependent loops and domains, but lack transcription-mediated compartmentalization

Charlotte Cockram[1,] Agnès Thierry[1], Aurore Gorlas[3], Roxane Lestini[2] and Romain Koszul[1,4*]

Affiliations :

1 Institut Pasteur, Unité Régulation Spatiale des Génomes, CNRS UMR 3525, F-75015 Paris, France

2 Laboratoire d'Optique et Biosciences, École Polytechnique, CNRS UMR7645 – INSERM U1182, IP Paris, 91128 Palaiseau Cedex, France

3 Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

[4] Lead Contact

* Corresponding author: romain.koszul@pasteur.fr

Summary

Hi-C has become a routine method for probing the 3D organization of genomes. However, when applied to prokaryotes and archaea, the current protocols are expensive and limited in their resolution. We develop a cost-effective Hi-C protocol to explore chromosome conformations of these two kingdoms at the gene or operon level. We first validate it on *E. coli* and *V. cholera,* generating sub-kilobase resolution contact maps, and then apply it to the Euryarchaeota *H. volcanii*, *Hbt. salinarum* and *T. kodakaraensis.* With a resolution of up to 1 kb, we explore the diversity of chromosome folding in this phylum. In contrast to Crenarchaeota, these Euryarchaeota lack (active/inactive) compartment-like structures, and their genomes are instead composed of self-interacting domains and chromatin loops. In *H. volcanii*, these structures are regulated by both transcription and the archaeal SMC protein, further supporting the ubiquitous role of these processes in shaping the higher-order organization of genomes.

Introduction

Due to their small size and dynamic architecture, the three-dimensional (3D) organization of microbial genomes have historically proved challenging to investigate. Archaeal genomes are similar to bacteria in terms of size and overall organization, with both possessing circular genomes that are unconfined by a membrane or nucleus (Ausiannikava et al., 2018; Koonin and Wolf, 2008). Yet in terms of DNA replication, genome segregation, and cell division, these two kingdoms are quite different. Bacterial genomes typically consist of a single chromosome with a unique origin of replication (*oriC)*, although a few species also contain large secondary replicons comparable in size to the main chromosome (Harrison et al., 2010), but relying on plasmid-based mechanisms for their replication (Egan et al., 2005). In contrast, archaea chromosomes commonly have multiple origins of replication and utilize the Orc1/Cdc6 initiator proteins, which are homologous to eukaryotic systems (Makarova and Koonin, 2013). They also often have large secondary replicons, but unlike bacteria, these mini-chromosomes rely on the same replication machinery as the main chromosome (Ng et al., 1998).

The application of chromosome conformation capture (3C/Hi-C) (Lieberman-Aiden et al., 2009) approaches has provided the opportunity to further highlight the myriad of different processes responsible for chromosome organization within these two domains. Notably, all bacterial chromosomes display self-interacting chromosomal interaction domains (CIDs), ranging in size from 30 – 420 kb, whose precise nature and function remain elusive (Le et al., 2013; Lioy et al., 2018; Marbouty et al., 2015; Wang et al., 2015). In addition, molecular complexes called condensins, belonging to the structural maintenance of chromosomes (SMC) proteins family, bridge the chromosome arms of *Bacillus subtilis* and many other species (Böhm et al., 2020; Le et al., 2013; Marbouty et al., 2015, 2017; Wang et al., 2015), while promoting long-range DNA loops (Marbouty et al., 2015). Despite the presence of the SMC homolog MukBEF, these phenomena are absent in *E. coli* (Lioy et al., 2018). More recently, the genomes of the hyperthermophile Crenarchaeota, *Sulfolobus acidocaldarius,* and *Sulfolobus islandicus* were investigated by Hi-C, resulting in the first archaeal contact maps (Takemata et al., 2019). These maps show that the *Sulfolobus* species display a bipartite chromosomal conformation distinct to that of bacteria, with the presence of two compartments that correlate with their gene expression activity. These Crenarchaeota also appear to have evolved a condensin-independent mechanism of chromosome organization in which a distant SMC homolog protein, called coalescin (ClsN), maintains this bi-compartmentalization. Whether these results are common to all archaea or if, as in bacteria, different species display significantly different genome organization remains to be seen.

Taken together, the variability already observed across these two kingdoms demonstrates the need for more in-depth investigations of chromosome architecture in the different clades of bacteria and archaea. However, this is not necessarily easy, as many bacterial and archaeal species exist in extreme environments which are difficult to manipulate in a laboratory setting. In this respect, Hi-C has already proven to be a readily available tool for gaining a glimpse at the genome folding of those species that are less amenable than the standard laboratory models. However, despite key improvements to eukaryotic Hi-C protocols leading to the development of commercial kits, the prokaryotic assay remains highly inefficient and costly. Early experiments performed in the bacteria *Caulobacter crescentus* and *B. subtilis* used restriction enzymes that recognize six base-pair sequences to digest the

genome into fragments of several kilobases in size (Le et al., 2013; Wang et al., 2015). Although these contact maps were informative, their resolution was limited to ~10 kb (**Table S1**). A slightly better resolution was reached using four base-pair cutting enzymes, which digest the genomes into smaller fragments, but at a much higher cost (**Table S1**) (Marbouty et al., 2014, 2015). Despite years of collective efforts, the highest resolution of published prokaryotic contact maps remains limited to ~4-5 kb for bacteria and ~30 kb for archaea, well above the average size of a gene or operon.

In the present work, we optimized the protocol to bypass the current limitations in bacterial and archaeal Hi-C assays, significantly improving the resolution of contact maps at a fraction ($1/5^{th}$) of the sequencing cost. Not only is this approach more cost and time-effective compared to existing protocols, but the resulting contact matrices display up to a 30-fold improvement in resolution. We validate this protocol on two model bacteria, *E. coli* and *V. cholera*, utilizing the 500 bp resolution to reveal features previously unseen in published contact maps, and then apply it to *Haloferax volcanii*, a member of the Euryarchaeota phylum that is distinct from the previously studied *Sulfolobus* species. This halophilic archaeon, first described in 1975 following isolation from the bottom sediment of the Dead Sea (Mullakhanbhai and Larsen, 1975), is one of the most widely studied archaeal models, and as a polyploid is representative of the Euryarchaeota phylum as a whole (Barillà, 2016). By applying the optimized Hi-C protocol to *H. volcanii* and two additional Euryarchaeota, the mesophilic halophile *Halobacterium salinarum* and the hyperthermophile *Thermococcus kodakaraensis,* we show that Euryarchaeota have a chromosome conformation that is distinct to that of *S. acidocaldarius* and *S. islandicus*, and that is more similar to bacteria. We then explore the processes responsible for this conformation, identifying a role for the *H. volcanii* SMC protein, and then propose how these results may be reflective of the Euryarchaeota phylum as a whole.

Results

Achieving gene-level resolution in bacterial contact maps

We developed a Hi-C protocol more suited to bacteria but that could potentially be applied to archaea. Our starting point was the previously published bacterial Hi-C (Le et al., 2013) and 3C-seq

protocols (Marbouty et al., 2014), which were combined and optimized through a process of trial and error, with some of the changes being directly inspired by the evolution of the eukaryotic Hi-C protocol(s), while others were identified for the purpose of this work (see **STAR Methods**). We then benchmarked this protocol with the published contact maps of *E. coli* and *V. cholera* (Lioy et al., 2018; Val et al., 2016) to determine the extent to which it facilitated higher-resolution studies of prokaryotic chromosomes. The level of improvement was measured in two ways; firstly, by comparing the number of reads in the final contact map with the total number of reads sequenced **(Table S1)**, and secondly by determining the maximum resolution (defined here as bin size) of the final contact map.

Briefly, we tested two different four-base cutting restriction enzymes with biotinylation to generate Hi-C contact maps of *E. coli* **(Figure 1A, B** and **Figure S1**; **STAR Methods)**, with HpaII (C^CGG) and MluCI (^AATT) both performing well **(Table S1)**. The HpaII digestion gave a better signal-to-noise ratio for *E. coli* **(Figure S1A-C)** and the highest increase in informative reads in the contact map **(Table S1** and **Figure S1D)** and so was retained to perform all subsequent experiments. MluCI could nevertheless be of interest if a genome is particularly AT-rich, or if a specific AT-rich region is under investigation. We also added modifications to the blunt-end ligation step **(Figure 1A, Figure S1E,** and **F**; **STAR Methods)** to obtain a ~3-fold increase in informative reads compared to that previously reported (Lioy et al., 2018), allowing the generation of higher-resolution matrices with fewer empty bins **(Figure S1E)**. Finally, we included a streptavidin-mediated pull-down of biotinylated interactions immediately after sonication and size-selection of the library **(Figure 1A)**. The preparation of the sequencing library was then done on the streptavidin beads carrying biotinylated chimeric DNA fragments (Orlando et al., 2018; Rao et al., 2014), facilitating a further 3.5-fold enrichment of informative reads and a ~10-fold increase overall (**Figure S1F**).

To evaluate the quality of the data generated through this approach, the resulting bacterial contact maps **(Figure 1B** and **C)** were compared to published datasets from different laboratories, processed using the same computational pipeline (**Table S1**). Firstly, we compared the total number of reads sequenced with the number of reads retained in the final contact maps after alignment and filtering (Cournac et al., 2012). Starting with a tenth of material, the protocol retains between five and twenty-fold more reads compared to former experiments using the same strains and restriction enzymes **(Table**

S1) (Lioy et al., 2018; Marbouty et al., 2015; Val et al., 2016). The reduction in starting material makes this protocol accessible for precious, limited samples, whilst the increase in informative reads means that for less than a fifth of the sequencing cost, one can obtain Hi-C matrices as good as those published, making prokaryotic Hi-C experiments much more cost-effective. This increase in informative reads also increases the coverage of Hi-C maps, so that they are more "filled-in" with informative contacts. This, in turn, results in an improved resolution of up to 500 bp, a 10-fold increase in resolution compared to published bacterial contact maps **(Figure 1D-G, Figure S1E** and **Table S1)** (Lioy et al., 2018; Val et al., 2016)**.**

Although the overall global chromosome organization observed in the bacterial Hi-C maps was similar to those previously reported, the high-resolution matrices were much crisper with more contrasting features. In *E. coli* (binned at 5 kb and 500 bp), the thickness of the main-diagonal was much more heterogeneous compared to previous maps **(Figure 1B)**. This improved visualization is particularly beneficial in highlighting local changes in chromosome organization. For example, we observed that the 800 kb region surrounding *dif* and constituting the Ter domain (Boccard et al., 2005) exhibited an enrichment in short to mid-range contacts compared with the rest of the chromosome **(Figure 1B, D,** and **E).** The 500 bp resolution contact maps also allowed us to further disclose the sub-structures of the Ter MD. The domain appears more homogenous than the rest of the chromosome, the notable exception being the splitting of the region at *dif* **(Figure 1D** and **E)**. The sub-kilobase resolution also paves the way to study bacterial chromosomes at the gene or operon level. For instance, the 500 bp binned *E. coli* maps facilitated the identification of previously uncharacterized chromosomal borders at the *fliF-R* and *flgB-J* flagella operons **(Figure 1F** and **G)**. These operons are 7-10 kb in size and display similar structures, i.e., dense signals along the diagonal that insulate the neighboring chromosome regions from each other. Furthermore, the 500 bp resolution detected similar dense contacts at smaller operons and at individual, highly-expressed genes located on either side of the flagella genes **(Figure 1F** and **G)**, highlighting the sensitivity of this improved resolution, and suggesting the presence of gene domains along the chromosome. This increased resolution was also beneficial in studying *V. cholera,* facilitating an increased definition of structures throughout the genome, such as the superintegron on chromosome 2 (Mazel et al., 1998) (**Figure 1H**). Altogether, this

approach paves the way to further in-depth analysis of the influence of transcription on the local folding of the DNA fiber in bacteria.

Finally, to determine whether the method introduced biases regarding the type of reads that were enriched, we compared the intra vs. inter-chromosomal contact ratio of the two chromosomes of *V. cholera*. Both inter- and intra-chromosomal contacts show a ~10-fold increase compared to published data (**Figure 1I**), maintaining a similar intra/inter ratio and indicating that the protocol does not bias the type of contacts enriched. The increase in inter-chromosomal contacts, as seen for *V. cholera* **(Figure 1J),** will be particularly important for studying species with multiple chromosomes or extra-chromosomal elements.

Features of chromosome organization in *H. volcanii*

The optimized Hi-C protocol was next applied to exponentially growing cells of the archaeal species, *H. volcanii.* The *H. volcanii* genome is composed of a 2.85 Mb main chromosome, three smaller chromosomes ((pHV4 (636 kb), pHV3 (438 kb), pHV1 (85.1kb)), and the pHV2 plasmid (6.4kb) (Hartman et al., 2010). The main chromosome has three origins of replication (*oriC*1–3) and the smaller chromosomes are distinct from the plasmid because they each carry a unique *oriC* (Ausiannikava et al., 2018), whereas the plasmid has none. Despite the polyploid nature of these cells, highly reproducible contact matrices with a resolution of up to 1 kb were generated **(Figure 2A** and **Figure S2A-E)**. As expected, the contact map of each chromosome displays a single, strong diagonal signal resulting from the enrichment of contacts between neighboring loci. In agreement with previous reports (Hawkins et al., 2013), the H26 strain no longer contains the pHV2 plasmid (Wendoloski et al., 2001), while the pHV4 chromosome has integrated into the main chromosome (coordinates: 249,185bp - 884,970 bp). The 30-fold increase in resolution, compared to previously published archaeal Hi-C maps (Takemata et al., 2019), now facilitates the visualization of more detailed structures **(Figure 2A** and **Figure S2A**), notably:

Self-interacting chromosome domains

The Hi-C matrix displays self-interacting regions, that appear as squares along the main diagonal, reminiscent of bacterial CIDs (Le et al., 2013). A directional index (DI) analysis performed at 100 kb and 300 kb identified ~10 larger domains ranging from ~70 kb - 570 kb in size (**Figure 2B**) and ~23 domains ranging from ~25 kb to 200 kb in size **(Figure 2C)**, scales that are very similar to those observed previously in various bacteria (Le et al., 2013; Lioy et al., 2018; Marbouty et al., 2015; Wang et al., 2015) and to those recently identified in Crenarchaeota (Takemata and Bell, 2020). A comparison between the DI analysis and the transcription profile of the cells shows a good correlation between the CID boundaries (100 kb scale) and gene expression **(**Pearson coefficient: 0.41, **Figure 2C-E)**. Correlations between domain boundaries and gene expression have previously been shown in both prokaryotes and eukaryotes and are in good agreement with the idea that transcription creates plectonemic-free regions that insulate adjacent regions of the chromosome (Hsieh et al., 2015; Le and Laub, 2016; Sexton et al., 2012)

Plaid-like contact pattern

The *H. volcanii* contact matrix displays a plaid pattern involving very short (~2-15 kb) DNA segments throughout the genome **(Figure 2A).** This particular *Haloferax* genome has an average GC content of ~65%, but there is extensive variation throughout all of the chromosomes **(Figure 2F)** with 102 AT-rich regions, corresponding to insertion sequence (IS) elements, scattered throughout the genome (Hartman et al., 2010). A comparison between the GC content and the Hi-C contacts **(Figure 2A** and **F)** shows a strong correlation between AT-rich regions of the genome and the plaid patterns. This correlation is even more apparent for the pHV4 chromosome and the surrounding main chromosome **(Figure 3A).** Indeed, the pHV4 chromosome, which recently inserted into the main chromosome, is much more AT-rich than the surrounding sequence and contains nearly half of the 102 IS elements found in the genome. In fact, the genomic rearrangement that led to the integration of the pHV4 chromosome occurred between two identical ISH18 sequences (Hawkins et al., 2013) located in either chromosome. Similar correlations were observed for the pHV1 and pHV3 chromosomes. The 438 kb pHV3, which is GC-rich and has fewer IS elements than the other chromosomes, has much fewer instances of the plaid pattern **(Figure 3B)**. The pHV1 chromosome, on the other hand, is AT-rich and contains 16 IS elements over 85 kb, resulting in a dense plaid pattern and many empty bins within

the Hi-C matrix **(Figure 3C)**. Neither the density of the GC-rich HpaII sites (C^CGG), conserved for instance in the pHV4 region, the restriction enzyme used, or the length of sequencing reads (**Figure S2F** and **G, STAR Methods**), affects the presence of the plaid pattern in the contact map, suggesting that the pattern results most likely from the nature of the DNA within these regions (see below) rather than from intrinsic experimental distortions.

Discrete DNA loops scattered throughout the chromosomes

Finally, the increased resolution facilitated the visualization and characterization of finer structures such as intra-chromosomal boundaries **(Figure 3D)** and DNA loops **(Figure 3E)**. To detect and quantify these structures we utilized Chromosight, a computer vision-based program recently developed to call *de novo* DNA motifs in Hi-C maps (Matthey-Doret et al., 2020). Chromosight detected an average of 26 borders in the Hi-C contact maps of WT *H. volcanii* cells across three biological replicates **(Figure 3F),** in good agreement with the 23 detected by the DI analysis **(Figure 2C)**. Furthermore, and more intriguingly, Chromosight detected an average of 64 loop-like structures in the *H. volcanii* genome **(Figure 3F)**, most of which are on the main chromosome.

Euryarchaea genomes are not organized into active and inactive compartments

The AT-rich, IS element-correlated plaid pattern observed in the *H. volcanii* contact map **(Figure 2A** and **Figure 3A-C)** is reminiscent of the bipartite organization found in the *Sulfolobus* species and some eukaryotes (Lieberman-Aiden et al., 2009; Takemata et al., 2019). In these species, this bipartite organization typically involves much larger regions of DNA and is associated with the compartmentalization of DNA into transcriptionally active (A) and inactive (B) compartments. We, therefore, sought to determine whether the same is true for *H. volcanii* and if compartment organization is a general feature of Archaea

We generated contact maps of exponentially growing wild-type cells grown in the presence/absence of the transcriptional inhibitor, actinomycin D **(Figure 4Ai,** and **Bi)** and concomitantly performed RNA-seq analysis on the same cells **(Figure 4Aii** and **Bii)**. Since actinomycin D works by intercalating with DNA, giving rise to the potential for pleiotropic effects, we limited the

treatment time to 30 min (Delmas et al., 2013; Hundt et al., 2007; Jäger et al., 2003). Pearson correlation matrices, computed from the resulting Hi-C maps **(Figure S3)** were then used to perform principal component analysis (PCA; **Figure 4Aiii, Biii,** and **Figure S3**) for the identification of compartment-like structures (Lieberman-Aiden et al., 2009; Takemata et al., 2019). We also generated the differential contact map corresponding to the log2 ratio of WT and actinomycin D treated cells Hi-C maps **(Figure 4C**) and performed a differential expression analysis between the two conditions **(Figure 4D)**.

At first glance, the PCA analysis of exponentially growing, untreated-WT cells unveiled two types of loci behavior in the contact map, corresponding to either the GC-rich, active, main chromosome or to the more AT-rich regions, including the smaller pHV1, pHV3, and integrated pHV4 chromosomes **(Figure 4Ai-iii).** When compared to the untreated control, actinomycin D treated cells exhibited a decrease in short- to mid-range contacts and a global softening of all chromosomal structures **(Figure 4Ai, Bi** and **C),** which correlates with the results from the *Sulfolobus* species (Takemata et al., 2019) and is similar to that seen in bacteria treated with a transcriptional inhibitor (rifampicin; (Le et al., 2013). However, the similarity with the chromosome organization of Crenarchaeota stops there. Indeed, and in contrast to what was observed in *Sulfolobus*, the plaid pattern remained clearly visible in the Hi-C contact map of actinomycin D treated cells **(Figure 4Ai** and **Bi**; see also white stripes on the ratio map of **Figure 4C** pointing at few changes between the two conditions). Furthermore, despite the transcriptome of actinomycin D treated cells being significantly altered, with > 30% of the genome showing a > 2-fold decrease in gene expression ($p<0.05$; **Figure 3D**), the compartmentalization pattern throughout the genome remained unchanged in these cells compared to the untreated control **(Figure 4Aiii, Biii, E** and **G)**. This suggests that the observed splitting of the *H. volcanii* chromosome, revealed by PCA analysis, doesn't reflect the organization of the genome into spatially distinct, transcriptionally active, and inactive compartment-like structures.

This result was further supported by comparing exponential growing cultures to cells growing in stationary phase (**Figure 4A, Figure S4A** and **B)**. Here too, although RNA-seq analyses showed a significant remodeling of the transcriptome between exponentially growing and stationary phase cells, that also resulted in a reduction in short- to mid-range contacts (**Figure S4Ai, Aii** and **B)**, the plaid pattern persisted at a similar level to other chromosomal contacts (e.g. white stripes on the ratio map of

Figure S4B). Furthermore, despite 55% of genes showing a >2-fold change in gene expression ($p > 0.05$; **Figure S4C**), these changes didn't associate with corresponding changes in the compartment-like structures, as determined by principal component analysis **(Figure S4Aiii** and **D-F).** This is also in contrast to the *Sulfolobus* species, where compartmentalization signals were weakened in stationary phase (Takemata et al., 2019). These observations strongly support the idea that the plaid-like pattern seen in the *H. volcanii* contact maps isn't equivalent to that seen in the *Sulfolobus* species and in some eukaryotes. Finally, Hi-C on a strain lacking the four origins of replication on the main chromosome (*OriC1-3* and *ori-pHV4)* (Hawkins et al., 2013) displays very little overall difference in the contact patterns compared to WT **(Figure S4G** and **H)**. Although the polyploid nature of the *H.* volcanii genome may prevent the detection of smaller, localized structures, it suggests that replication is not strongly implicated in the chromosome structuring of *H. volcanii.* Instead, a possibility is that the plaid pattern corresponds to a side-effect of the AT-rich nature and an increased density of IS-elements in these regions of the genome (Discussion).

To test this hypothesis, and broaden the exploration of Euryarchaeota genomes, we performed Hi-C on two additional species, *Halobacterium salinarum* and *Thermococcus kodakaraensis* (**Figure 4Hi** and **Figure S4Ii)**. *Hbt. salinarum* is an extreme halophile that is similar to *H. volcanii* because its ~2.57 Mb genome is spread across a main chromosome and two smaller mini-chromosomes (Ng et al., 2000). The genome is also GC-rich but contains AT-rich islands **(Figure 4Hii)**. Unlike in *H. volcanii,* where the AT-rich regions are more dispersed throughout the genome, in *Hbt. salinarum* these regions that are mostly constrained to a 61 kb region on the main chromosome and the two smaller chromosomes **(Figure 4Hii)**. These AT-islands contain most of the 91 IS elements found in the *Hbt salinarum* genome (Pfeiffer et al., 2020) **(Figure 4Hiii)**. In contrast, the hyperthermophile species *T. kodakaraensis* has a much more stable GC-content and contains only 7 IS elements scattered throughout its single chromosome **(Figure S4Iii** and **Iiii).** This species is difficult to grow because it requires anaerobic growth conditions and high temperatures (85°C) (**STAR Methods**), making it more challenging to work with. Nevertheless, we obtained good quality Hi-C maps for both species, with the protocol working well with a limited amount of material. The *Hbt. salinarum* genome revealed a similar structure to *H. volcanii* (**Figure 4A and Hi**), with a directional index (DI) analysis, performed at 100

kb, identifying ~11 domains ranging from ~20 kb to 140 kb in size **(Figure 4Hiv)**. Similarly to *H. volcanii,* and distinct to the *Sulfolobus* species, the Hi-C analysis identified discrete DNA loops located throughout the genome **(Figure 4Hi)**. We also observed a strong correlation between the plaid pattern of small DNA segments and the location of AT-rich regions IS-dense regions of the genome **(Figure 4Hi-iii)**, a phenomenon that is completely absent in *T. kodakaraensis* **(Figure S4Ii-iii).** Furthermore, PCA analysis revealed that neither the *Hbt. salinarum* or *T. kodakaraensis* genomes are organized into compartment-like structures **(Figure 4Hv and Figure S4Iiv)**. Taken together these results suggest that, as in prokaryotes and eukaryotes, different archaea possess different types of genome organization. The absence of a plaid pattern in *T. kodakaraensis* also reveals that compartmentalization is not a feature associated with thermophilic archaea and that perhaps this type of genome organization is either limited to the *Sulfolobus* species studied (Takemata et al., 2019) or to the Crenarchaeota in general. Taken together, these results strongly support that the Euryarchaeota don't partition their genomes into active and inactive compartment-like structures, but instead organize them into long discrete DNA loops and large self-interacting domains.

DNA loops correlate with transcriptional activity in *H. volcanii*

Hi-C of stationary phase cultures revealed an altered chromosome structure, with a reduction in short- to mid-range contacts throughout the genome **(Figure S4Ai** and **B)**. In addition, a loss of chromosomal loops was observed throughout the main chromosome and the smaller chromosomes **(Figure S4Ai, Di, and Ei),** suggesting that these structures may relate to gene expression through the loading of a protein involved in the maintenance of the contacts between pairs of loci**.** To alleviate the potential indirect effects of physiological changes induced upon entering stationary phase, we investigated transcription and DNA loops in cells treated with actinomycin D **(Figure4, Bi, Ei, and Fi)**. The expression profile of the actinomycin D treated cells was much more homogeneous across the genome, with the Hi-C contact map showing a decrease in short-mid range contacts, chromosomal borders, and a disappearance of DNA loops when compared to the untreated control **(Figure 4Bi, Bii, E, and F)**. We applied Chromosight (Matthey-Doret et al., 2020) to the different Hi-C matrices to quantify the loss of these structures **(Figure 4I-K)**, showing that stationary phase and actinomycin-D

treated cells displayed a ~90% decrease in the number of chromosomal loops and a significant reduction in loop score (strength) for the remaining 10% **(Figure 4L)**. For both conditions, the loss of loops correlated with areas of reduced transcription (**Figure 4I-K**). Chromosomal borders, on the other hand, were less affected with a 32% and 65% reduction in the number of borders for actinomycin D-treated and stationary phase cells respectively. No decrease in border score was seen for those that remained **(Figure 4M)** Altogether, these results show that, contrary to all bacterial genomes investigated to date, the presence of DNA loops in the archaea *H. volcanii* is associated with transcriptional activity.

Role of SMC in chromosome organization

SMC-dependent DNA loops have been identified in the contact maps of most eukaryotic genomes investigated using Hi-C, as well as in bacteria that possess a canonical member of the SMC family, such as SMC-ScpAB in *B. subtilis* (Marbouty et al., 2015). Recently, SMC proteins have also been identified in archaea, with condensin shown to be highly conserved throughout archaeal species (Barillà, 2016), the notable exception being Crenarchaea, which instead encode a distant homolog known as coalescin (ClsN; **Figure 5A)** (Takemata et al., 2019).

H. volcanii encodes three SMC encoding genes; SMC (HVO_0689) and Rad50 (HVO_0854), as well as the SMC-like Sph4 (HVO_B0173) protein. Sequence analysis reveals that the SMC protein contains both the characteristic ATP-binding N-terminal and flexible hinge domains, and is highly conserved across all members of the Halobacteria class (>65% identity) and among other species of the Euryarchaeota phylum, whilst also possessing some (25-30%) sequence identity with members of the Asgard and TACK superphyla **(Figure 5A)**. We hypothesized that the link between transcription and DNA-loops in *H. volcanii* may be mediated by the SMC protein and so generated a strain in which the *smc* gene (HVO_0689) had been deleted. Interestingly, and contrary to many species, the *H. volcanii smc* mutant was viable and grew similar to WT cells **(Figure 5B)**. It also exhibited normal cellular morphology and DNA content by microscopy **(Figure 5C)**, yet Hi-C experiments revealed a complex pattern of changes across the genome of *smc* mutant cells **(Figure 5D)**. First, there was an overall decrease in short-range contacts in Δ*smc* cells compared to wild-type. We also observed a reduction in

both the number and strength of some DNA loops across the genome, although the effect was less than that seen in actinomycin D-treated or stationary phase cultures. However, the most striking observation was the complete loss of ~9 topological borders throughout the main chromosome (**Figure 5E**). The transcription profile of *smc* mutant cells was generated to test whether these 3D changes were associated with alterations in gene expression at these sites **(Figure 5F)**. To our surprise, the transcriptome of the Δ*smc* mutant was highly similar to wild-type, with less than 5% of genes showing > 2-fold change in gene expression between the two conditions. Furthermore, only one of ~9 modified topological borders correlated with a site of > 2-fold change in gene expression (**Figure 5G-I**). Despite close inspection of the sequences and gene annotations at these borders **(Figure 5E** and **Table S2)**, no DNA pattern was found to suggest why these specific sites were affected in *smc* mutant cells. To determine whether this pattern was conserved across the SMC family members present in *Haloferax*, we performed Hi-C on Δ*rad50* and Δ*sph4* mutants **(Figure S5B-E)**. Like the SMC protein, Rad50 is conserved across most Euryarchaea but is less-well conserved in the Asgard and TACK groups, whereas the Sph4 protein seems to be limited to the Haloferacaceae. Both mutants grew similar to WT and Δ*smc* cells **(Figure 5B and C**) and exhibited the same loss of short-mid range contacts as the Δ*smc* mutant **(Figure 5D** and **E**; **Figure S5B-E)**. However, we did not observe the loss of boundaries or loops as seen in the Δ*smc* mutant, indicating that this phenotype was specific to the Δ*smc* cells **(Figure 5J** and **K)**. Application of Chromosight (Matthey-Doret et al., 2020) to these Hi-C maps **(Figure 5J, L** and **M)** confirmed a 50% reduction of both chromosomal borders and DNA loops in Δ*smc* cells, whereas the Δ*rad50* and Δ*sph4* mutants were very similar to WT cells **(Figure 5J, L and M)**. Taken together these results highlight a role for SMC, along with transcription, in shaping Euryarchaea genomes.

Discussion

The simple, cost-effective Hi-C approach described above allows to explore bacterial and archaeal chromosome conformations at the gene /operon level. Applied to the well-studied *E. coli* and *V. cholera* species, it allowed us to generate sub-kilobase resolution contact maps of these species, unveiling previously undetected gene-level domains (Lioy et al., 2018; Val et al., 2016). This is an

important advancement in understanding the 3D genome organization of prokaryotes, as it is a bottleneck that 3C/Hi-C and microscopy techniques have long been trying to break. We then applied this approach to the lesser-studied Euryarchaeota, *H. volcanii*, *Hbt. salinarum* and *T. kodakaraensis,* generating Hi-C matrices at a resolution of up to 1kb. This increase in resolution, compared to the currently available archaeal Hi-C maps of the *Sulfolobus* species (Takemata et al., 2019), facilitated the identification of chromosomal architectures in all three species. These observations show that, as in bacteria and eukaryotes, different types of chromosome structuring mechanisms coexist in the archaeal kingdom.

Euryarchaea genomes are not segregated into active and inactive compartments

In contrast to *Sulfolobus,* the Hi-C analysis of the genomes of *H. volcanii* didn't reveal the genome was organized into active and inactive spatial compartments. Indeed, when compared to wild-type cells, the transcriptional changes observed following treatment with actinomycin D or by entering into stationary phase were not associated with changes in the corresponding Hi-C contact maps. The plaid pattern observed in the contact maps of this species and for two other Euryarchaeota instead correlates with the presence of AT-rich regions of these chromosomes that are dense with IS-elements. Notably, the hyperthermophile, *T. kodakaraensis*, which contains just seven IS elements compared to the ~100 identified in the genomes of *Hbt. salinarum* and *H. volcanii* didn't display any plaid patterning. Why AT-rich regions in the *Halobacterium* and *Haloferax* genomes result in this plaid-pattern is unclear. These regions have previously been shown in *Hbt. salinarum* to be hypermutable due to random transposition and recombination events between the IS elements contained within these regions. Although this hypermutable phenotype is yet to be demonstrated for *H. volcanii* (López-García et al., 1992), similar mechanisms may exist. Many of the IS elements between the two species contain similar inverted repeats (Hofman et al., 1986) and we know that the pHV4 chromosome integrated into *H. volcanii* main chromosome via a recombination event between two ISH18 elements (Hawkins et al., 2013). If frequent enough, these events could result in a heterogeneous genome population in the cell culture, and read misalignments resulting in a plaid contact. Alternatively, these patterns may simply be the result of indirect biases of the experimental protocol on these low-complexity regions. Whatever

their origin, these contacts are unlikely to correspond to the segregation of functional DNA into transcriptionally active and inactive domains.

These results suggest that the segregation of the *Sulfolobus* genomes into different chromatin domains could be a trait specific to this lineage. We can also conclude that this segregation is not necessarily linked to the extreme temperature of its environment since *T. kodakaraensis* doesn't display such an organization. Unlike many archaeal species, including the three species tested in this study, members of the Crenarchaeota phylum do not contain histones (Barillà, 2016). To test whether this absence correlates with the presence of compartments will be facilitated by the availability of an efficient Hi-C protocol to rapidly sample multiple archaeal species.

Transcription and SMC organize the Haloferax genome into DNA loops and domains

Discrete DNA loop structures were identified throughout the genome, whose strength and visibility strongly correlate with the transcriptional status of the cells. We hypothesized that the link between transcription and DNA-loops in *H. volcanii* may be mediated by an SMC protein. In most species, disruption of the SMC proteins results in lethality or a severely impaired growth phenotype. A *Haloferax* strain carrying a deletion of *smc* displays no obvious defect, a phenotype previously described in several bacteria species. Indeed, in *M. tuberculosis, M. smegmatis, C. glutamicum,* and *D. radiodurans* Δ*smc* mutations have also been shown to have very little or no effect on cell growth or the frequency of anucleate cells (Böhm et al., 2020; Bouthier de la Tour et al., 2009; Güthlein et al., 2008). *D. radiodurans* is of particular interest as, like *H. volcanii,* it is an extremophile with a multipartite genome present in several copies (Cox and Battista, 2005). *D. radiodurans* Δ*smc* mutants are similar to WT cells in every way, except that they are hypersensitive to DNA gyrase inhibitors, suggesting SMC is required to maintain the supercoiling equilibrium in these cells (Bouthier de la Tour et al., 2009). Furthermore, it was shown that SMC is located at several distinct foci on the nucleoid of wild-type cells and that these foci are unevenly distributed throughout the cell body, signifying that SMC may bind to a few distinct DNA regions throughout the *D. radiodurans* genome (Bouthier de la Tour et al., 2009). Although the effect of Δ*smc* on the topology of the *H. volcanii* genome remains to be characterized, the chromosomal borders that are most severely disrupted in *smc* mutant cells may correspond to specific

SMC binding sites. This would also explain why we don't see the same phenotype in *rad50* and *sph4* mutant cells. The peculiarities of these sites remain an open question, as no conserved sequences could be found at the boundaries. A possibility could be that these sites are implicated in chromosome segregation or cohesion between homologous chromosomes, but this is difficult to test by Hi-C as the technique is not able to discern between homologous chromosomes. The SMC protein is well conserved throughout the Euryarchaeota phylum, which consists mainly of polyploids. Here too, further studies will determine whether these structures are conserved in these species.

This study pinpoints that understanding the chromosomal architecture of archaea is in its early stages. The stark contrast between these Euryarchaeota and the *Sulfolobus* species may be pondered now that higher-resolution maps of the latter are available (Takemata and Bell, 2020). These 3C-seq maps reveal that the chromosome of these Crenarchaeota do in fact possess some of the features identified in Euryarchaeota but that many differences still exist. However, the extreme and diverse environments that many archaeal species inhabit, may as well prelude to the identification of even more discrepancies between the different lineages. Since Hi-C can be applied to cells directly crosslinked in their natural environment (Marbouty et al., 2014), there should be no limitation to lab-based model organisms, and it would be of interest to generate high-resolution contact maps of species spanning the archaeal tree to reach a broad overview of their differences and similarities. The availability of high-resolution metagenomic Hi-C assays will further allow the sampling of genome folding of hundreds of species from the wild, including archaea. This technique will further boost the exploration of the evolutionary role of genome folding in the microbial world.

Limitations

This work details an improved Hi-C protocol that provides insights into the chromosome organisation of two bacterial and three archaeal species at the gene/operon level. However, it is worth noting that the study suffers from limitations. Firstly, the Hi-C method relies on a population of cells and as a result the contact maps reflect an average of the population which may mask the presence of cell-cell heterogeneity in different conditions and mutant backgrounds. Second, the Euryarchaeota are

polyploid, so inter-chromosomal contacts between homologous chromosomes cannot be determined using the Hi-C technique. The archaea themselves impose a limit on this study, they are very slow growing, making genetic manipulations more difficult and time-consuming. They are also resistant to most antibiotics. Furthermore, many of the *H. volcanii* genes are yet to be assigned functions and this currently prevents a detailed analysis of the changes observed in the transcriptome of treated and mutant cells.

Acknowledgments

We thank Olivier Espéli, Frédéric Boccard, and Vicky Lioy for the careful reading of the manuscript. We thank Jacques Oberto for his help in running the *T. kodakaraensis* experiment with support from Project EVOMOBIL (ERC Grant Agreement no. 340440). We also thank the members of the RSG lab and especially Martial Marbouty for insightful discussions about experiments, as well as Théo Foutel-Rodier for computational guidance, and Cyril Matthey-Doret and Axel Cournac for providing access to the Chromosight program. This research was supported by funding to R.K from the European Research Council under the Horizon 2020 Program (ERC grant agreement 771813).

Author Contributions

C.C, R.L, and R.K conceived of the study and designed the experiments. C.C. and A.T adapted the Hi-C protocol. C.C and R.L performed experiments, with assistance from A.G for the T. k. cultures. C.C performed the analysis. C.C and R.K interpreted the data and wrote the manuscript.

Declaration of Interests:

The authors declare no competing financial interests.

Figure Legends

Figure 1: Improved bacterial contact maps using an optimized Hi-C assay

(A) Overview of the Hi-C protocol, highlighting the steps that have been optimized. Normalized Hi-C contact maps of asynchronously growing populations of WT bacterial cells. The X and Y axes represent the coordinates of the chromosome and the colorscale reflects the frequency of contacts between two regions of the genome (arbitrary units), from white (rare contacts) to dark red (frequent contacts). Features of interest are indicated along the top axis.

(B) Hi-C contact matrix of WT *E. coli* binned at 5 kb.

(C) Hi-C contact matrix of WT *V. cholera binned* at 5 kb.

(D and E) Side-by-side comparison of the 800 kb region surrounding the terminus of replication *dif,* from both the improved and originally published contact maps (Lioy *et al*., Cell, 2018). Matrices are binned at either 5 kb or 0.5 kb showing the 10-fold improvement in resolution afforded by the new Hi-C protocol.

(F) Side-by-side comparison of a 50 kb region surrounding the *fliF-R* operon from both the improved and originally published contact maps (Lioy et al 2018), with 0.5 kb binning revealing the 10-fold improved Hi-C resolution reveals gene-level chromosomal structures.

(G) Side-by-side comparison of a 50 kb region surrounding the *flgB-J* operon from both the improved and originally published contact maps (Lioy et al 2018), with 0.5 kb binning.

(H) Side-by-side comparison of 250 kb region surrounding the superintegron located on chromosome 2 of *V. cholera* from both the improved and originally published contact maps (Val *et al.,* Science Advances, 2016).

(I) Percentage of 3D intra- and 3D inter-chromosomal events obtained from *V. cholera* obtained through the improved and published approaches.

(J) Side-by-side comparison of inter-chromosomal contact maps (bin: 5 kb) centered on *dif* (top panels) or *ori* (bottom panels) of chromosomes 1 and 2 obtained from data obtained using the improved and published approaches (Val *et al.,* Science Advances, 2016). The crtS site is indicated.

Figure 2: Overall chromosomal organization of the archaea *H. volcanii*

Normalized Hi-C contact maps of asynchronously growing populations of WT *H. volcanii* cells. The X and Y axes represent the coordinates of the chromosome and the colorscale reflects the frequency of

contacts between two regions of the genome (arbitrary units), from white (rare contacts) to dark red (frequent contacts). Features of interest are indicated along the top axis.

(A) Hi-C contact matrix of WT *H. volcanii* binned at 5 kb.

(B) Directionality index analysis (300 kb scale), downstream (yellow), and upstream (grey) biases are indicated.

(C) Directionality index analysis similar to that shown in (B) but at the scale of 100 kb.

(D) RNA-seq depicting the transcription levels across the genome (data is an average of n=3 biological replicates).

(E) Correlation between transcription (red) and short-range chromosome contacts captured by Hi-C (blue). The Pearson correlation between these two signals and the corresponding p-value computed over the entire genome are indicated.

(F) GC content (50 kb sliding window) of the genome.

Figure 3: Finer levels of organization of the *H. volcanii* genome

(A, B, C) Magnifications of the 85 kb pHV1 chromosome, 437.9 kb pHV3 chromosome, and a 1 Mb region (coordinates: 65 kb – 1,065 kb) surrounding the pHV4 chromosome which has integrated into the main chromosome. For each magnification, the location of insertion (IS) elements, GC percentage (1 kb binning), and RNA-seq (average of n=3 biological replicates) are shown.

(D) Magnification of a 100 kb region of the main chromosome (coordinates 1474 kb – 1574 kb; bin = 1 kb).

(E) Magnification of 1 Mb region of the main chromosome (coordinates 800 kb – 1800 kb). The position of chromosomal loops and borders that have been detected with Chromosight are indicated in black and blue respectively.

(F) Bar plot showing the average number of borders and loops detected in the entire genome of an exponentially growing culture of *H. volcanii* (detected using Chromosight, n=3 Hi-C replicates). Associated pileup plots of windows centered on the detected loops in each condition are shown in the panels above.

Figure 4: Effect of transcription on the chromosome organization of *H. volcanii*

Normalized Hi-C contact maps of asynchronously growing populations of WT and mutant *H. volcanii* cells (5 kb bin). The X and Y axes represent the coordinates of the chromosome and the colorscale reflects the frequency of contacts between two regions of the genome (arbitrary units), from white (rare contacts) to dark red (frequent contacts). Features of interest are indicated along the top axis.

(A, B) Hi-C contact matrix of WT *H. volcanii* growing the presence/absence (exponential phase) of actinomycin D (i). For both conditions, the RNA-seq (average of n=3 biological replicates) (ii) and principal component analysis (PCA, iii) are shown.

(C) Differential contact map corresponding to the log2 ratio of Hi-C interactions between exponential and stationary phase cultures. Blue to red colorscale reflects the enrichment in contacts in one population with respect to the other, white indicates no difference.

(D) Differential expression analysis comparing actinomycin D-treated and untreated cells n=3 biological replicates).

(E, F) Magnifications of the 437.9 kb pHV3 chromosome in actinomycin D-treated and untreated cells (i). The RNA-seq analysis (average of n=3 biological replicates) (ii) and PCA (iii) are also shown.

(G) Differential expression analysis for the 437.9 kb pHV3 chromosome comparing actinomycin D-treated and untreated control cells (n=3 biological replicates).

(H) Normalized Hi-C contact map and PCA of an asynchronous population of WT *Hbt.* salinarum cells (i, 5kb bin). The GC content (ii, 50 kb sliding window), the position of IS elements (iii), DI analysis (iv, 100 kb scale), and PCA analysis (v) are also shown. Empty bins represent a ~145kb sequence common to the PNRC100 and PNRC200 chromosomes (Ng et al., 2000).

(I, J, K) Magnifications of the 437.9 kb pHV3 chromosome for exponential and stationary phase cultures, and cells treated with actinomycin D. The RNA seq (average of n=3 biological replicates) is shown for each condition and the position of chromosomal loops and borders that have been detected with Chromosight are indicated in black and blue respectively.

(L, M) Comparison of loop and border score distributions obtained using Chromosight for the Hi-C matrices shown in panels I-K. The number of loops detected and the mean of the Pearson coefficient for each condition are shown above (n=2 Hi-C replicates).

Figure 5: Role of SMCs in *H. volcanii* chromosome organization.

(A) Phylogenetic tree showing the conservation of canonical SMCs, Rad50, and the recently identified Coalescin (ClsN) proteins between different archaeal lineages. Adapted from (Spang et al., 2017; Takemata et al., 2019) and explained in the **STAR Methods**.

(B) Generation times of wild-type (WT), Δ*smc,* Δ*rad50,* and Δ*sph4* mutants. Error bars represent the standard deviation of the mean where n >3. p values are based on t-test comparisons between each mutant and the WT strain.

(C) Images of DIC, Hoechst 33342 signal, and the merged image (DIC signal in grey and the DNA signal in cyan) of WT, Δ*smc,* Δ*rad50*, and Δ*sph4* cells. Scale bar = 5 µm.

(D) Normalized Hi-C contact maps of asynchronously growing populations of WT and Δ*smc* mutant cells (5 kb bin). The X and Y axes represent the coordinates of the chromosome and the colorscale reflects the frequency of contacts between two regions of the genome (arbitrary units), from white (rare contacts) to dark red (frequent contacts). Features of interest are indicated along the top axis.

(E) Differential contact map corresponding to the log2 ratio of Hi-C interactions between WT and Δ*smc* mutant cells. Blue to red colorscale reflects the enrichment in contacts in one population with respect to the other, white indicates no difference. Chromosomal borders that have been lost in Δ*smc* mutant are highlighted by dashed lines. The exact coordinates and the genes found at the site of these borders are detailed in **Table S2**.

(F) Differential expression analysis of WT and Δ*smc* cells. Red and green represent >2 and <-2 log2-fold-change respectively (n=3 biological replicates).

(G) Magnification of 1 Mb region of the main chromosome (coordinates 340 kb – 1340 kb) for WT (top) and Δ*smc* mutant cells (bottom).

(H) Differential contact map corresponding to the log2 ratio of Hi-C interactions between WT and Δ*smc* mutant cells (coordinates 340 kb – 1340 kb).

(I) Differential expression analysis of Δ*smc* mutant cells compared to WT (coordinates 340 kb – 1340 kb, n=2 biological replicates).

(J) Magnification of 1 Mb region of the main chromosome (coordinates 1620 kb – 2620 kb) for WT, Δ*smc,* Δ*rad50, and* Δ*sph4* mutant cells. The position of chromosomal loops and borders that have been detected with Chromosight (Matthey-Doret et al., 2020) are indicated in black and blue respectively.

(L) Differential contact maps corresponding to the log2 ratio of Hi-C interactions between WT and mutant cells.

(L and M) Comparison of loop and border score distributions obtained using Chromosight for the Hi-C matrices shown in panels (D) and (F). The number of loops detected and the mean of the Pearson coefficient for each condition are shown above (n=2 Hi-C replicates).

STAR Methods

LEAD CONTACT

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Romain Koszul (romain.koszul@pasteur.fr).

MATERIALS AVAILABILITY

All strains and plasmids are available, without restriction, from the Lead Contact.

DATA AND CODE AVAILABILITY

The accession number for the sequencing reads reported in this study is PRJNA587586. All accession numbers are listed in the Key Resources Table. Open-access versions of the programs and pipelines used (E. *coli* analysis pipeline and Chromosight) are available online via the lab Github account (https://github.com/koszullab/).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial Strains

The *Escherichia coli* MG1655 K12 (Blattner et al., 1997) and *Vibrio cholera* N16961 (Heidelberg et al., 2000) strains used in this study are listed in the Key Resources Table.

Archaea Strains

The *Haloferax volcanii* H26 (Δ*pyrE*) (Allers et al., 2004), *Thermococcus.kodakaraensis* (Gorlas et al., 2013, 2014) and *Halobacterium salinarum* (Pfeiffer et al., 2020) strains used in this study are listed in the Key Resources Table.

METHOD DETAILS

Growth conditions

E. coli and *V. cholera* were both grown at 37°C in 1 x Minimal Media A supplemented with 0.2 % casamino a and 0.5 % glucose. The archaea, *H. volcanii* H26 strain (Δ*pyrE*) is referred to as "wild-type" during this study. *H. volcanii* strains were grown using enriched Hv-YPC media at 45°C, as described

previously (Allers et al., 2004). To block transcription, cells were exposed for 30 min to Actinomycin D (Sigma Aldrich, 5 µg/ml final concentration, prepared in 100% DMSO). The 30 min time point was chosen based on previous studies in halophilic archaea (Delmas et al., 2013; Hundt et al., 2007; Jäger et al., 2003) and confirmed by transcriptomic analysis. *Escherichia coli* strains XL1-Blue MRF' and GM121 were used for cloning. GM121 was also used to prepare unmethylated plasmid DNA for efficient transformation of *H. volcanii*. *T.kodakaraensis* was grown at 85°C, in anaerobiosis, using Ravot medium supplemented with elemental sulfur (10g/L) as previously described (Gorlas et al., 2013, 2014) and *H. salinarum* (NRC-1) was grown at 37°C in YPC containing 3.5 M NaCl.

Construction of the H. *volcanii Δsmc*, *Δrad50*, and *Δsph4* strains

Detailed information of all strains and plasmids used for constructions can be found in the Key Resources Table and all primers are listed in Table S3.

Δsmc

To generate the Δ*smc* (HVO_0689) construct, the upstream region (US) and downstream region (DS) of *smc* were cloned into pTA131(Allers et al., 2004) to generate the pRL93 plasmid used to construct the deletion strain by a gene knockout system(Bitan-Banin et al., 2003). The US region was generated by PCR on H26 genomic DNA using primers RL292 and RL296. The DS region was generated by PCR on H26 genomic DNA using primers RL294 and RL295. Each PCR product contained 30 bp homology with adjacent fragments for SLIC cloning(Li and Elledge, 2007): (i) with the pTA131 linearized fragments after NotI and EcoRI double-digestion (contained in RL292) in the US fragment, (ii) with the US fragment (contained in RL294) and with the pTA131 linearized fragments after NotI and EcoRI double-digestion (contained in RL295) in the DS fragment. Following the SLIC method, PCR fragments and the linearized plasmid were digested using T4 DNA polymerase for 45 minutes at 22°C to generate 3'-single-stranded extremities, then all amplification products were mixed in a 1:1:1 molar ratio and incubated for 30 minutes at 37°C before transformation into *E. coli* XL1-blue. Transformants were selected on LB plates containing 100 µg/ml ampicillin, 0.5 µM IPTG, and 80 µg/ml X-Gal. The

presence of the correct insert in the plasmid, as determined by white colonies, was tested by colony PCR using primers pBSF2 and pBSR3. The sequence of one selected plasmid, dubbed pRL93, was further confirmed by Sanger sequencing. Then pRL93 was used to transform H26 strain using the pop-in/pop-out method as described previously (Bitan-Banin et al., 2003). Pop-out colonies were plated on Hv-Ca plates containing 5-FOA and thymidine. The absence of the *smc* was tested by colony lift on 100 colonies from the pop-out plates using a probe targeting the *smc* gene. The DIG-labeled probe was generated by PCR on H26 genomic DNA using primers RL305 and RL310 and the PCR DIG (digoxigenin) labeling Mix (Roche). Probe hybridization was detected using the DIG Luminescent Detection Kit (Roche) and a ChemiDoc MP (BioRad). 10 colonies out of 100 were deleted for the *smc* gene. One Δ*smc* construct was conserved for further studies and dubbed HvRL138. The deletion of *smc* was further confirmed by PCR on HvRL138 genomic DNA using primers RL325 and RL326.

Δrad50

The Δ*rad50* construct (HVO_0854) was generated using the same strategy. The upstream region (US) and downstream region (DS) of *rad50* were cloned into pTA131 to generate the pRL96 plasmid used to construct the deletion strain by a gene knockout system. The *rad50* US region was generated by PCR on H26 genomic DNA (Δ*pyrE2*) using primers RL333 and RL334. The *rad50* DS region was generated by PCR on H26 genomic DNA (gDNA) using primers RL335 and RL336. The absence of the *rad50* was tested by colony lift on 100 colonies from the pop-out plates, using a DIG-labeled probe targeting the *rad50* gene generated by PCR on H26 genomic DNA using primers RL329 and RL330. 19 colonies out of 50 were deleted for the *rad50* gene. One Δ*rad50* constructed was conserved for further studies and dubbed HvRL148. The deletion of *rad50* was further confirmed by PCR on HvRL148 genomic DNA using primers RL342 and RL343.

Δsph4

For the Δs*ph4* construct (HVO_B0173). The upstream region (US) and downstream region (DS) of *sph4* were cloned into pTA131 to generate the pRL95 plasmid used to construct the deletion strain by a gene knockout system. The *sph4* US region was generated by PCR on H26 genomic DNA using

primers RL317 and RL318. The *sph4* DS region was generated by PCR on H26 genomic DNA (gDNA) using primers RL319 and RL340. The absence of the *sph4* was tested by colony lift on 100 colonies from the pop-out plates, using a DIG-labeled probe targeting the *sph4* gene generated by PCR on H26 genomic DNA using primers RL323 and RL324. 23 colonies out of 50 were deleted for the *sph4* gene. One Δ*sph4* constructed was conserved for further studies and dubbed HvRL147. The Deletion of *sph4* was further confirmed by PCR on HvRL147 genomic DNA using primers RL344 and RL345.

Optimized microbial Hi-C protocol

Cell fixation

All Hi-C experiments were performed with ~1 x 10^8 cells growing in the exponential growth phase. For *E. coli, V. cholera, H. volcanii,* and *Hbt. salinarum,* protein-DNA interactions were chemically crosslinked by the addition of fresh formaldehyde directly to the cultures (3% final concentration) for 30 min at room temperature with gentle agitation. For *T. kodakaraensis,* protein-DNA interactions were chemically crosslinked by mixing 20 ml of culture with 80 ml of room temperature formaldehyde (diluted in 1X PBS containing 0.8M NaCl, 4% final concentration). Crosslinking was quenched by the addition of glycine (0.5 M final concentration) for 20 min at room temperature. For *H.volcanii,* which is very sensitive to salt concentrations, 2.5M glycine was prepared in 18 % (w/v) saltwater solution (2.5 M NaCl, 90 mM $MgCl_2.6H_2O$, 90 mM $MgSO_4.7H_2O$, 60 mM KCl, 10 mM Tris–HCl pH 7.5) to avoid cell lysis. Cells were then collected by centrifugation (4000 x g, 10 min, room temperature) resuspended in 50 ml of 1 x PBS (18% (w/v) saltwater for *H. volcanii*), and centrifuged again. The pellet was resuspended in 1 ml 1 x PBS and transferred to a 1.5 ml Eppendorf tube before a final centrifugation step (4000 x g, 5 min, room temperature), the supernatant was then removed and the pellet stored at -80°C. In the original bacterial Hi-C study (Le et al., 2013) it was recommended to process pellets within 2 weeks of collection. Here, we checked that cell pellets can be stored for > 1 year without any effect on the quality of the final Hi-C library.

Cell lysis and DNA digestion

Cell pellets were removed from the -80°C freezer and allowed to completely thaw on ice. The pellet was then resuspended in 1 ml of 1x TE + complete protease inhibitor cocktail (EDTA-free, Sigma Aldrich) and transferred to a 2 ml Precellys tube containing 0.5 mm glass beads (VK05, Ozyme). Cells were then subject to mechanical disruption using the Precellys Evolution tissue homogenizer (V7500: 5 x 30s, 20s pause). Mechanical disruption has been shown to perform equally well on both Gram-positive and Gram-negative bacteria, as well as archaea. Please note, that for Precellys machines without the Cryolys cooling attachment (Ozyme), tubes need to be removed every 3 cycles and placed on ice for 5 min to stop sample overheating and subsequent degradation. The lysate (~1ml in volume) was then carefully transferred to a 5 ml Eppendorf tube, avoiding the transfer of any glass beads. Proteins that were not crosslinked to DNA during formaldehyde fixation were then degraded by the addition of SDS (Thermo Fisher, 0.5% final concentration) for 10 min at room temperature. DNA was then prepared for digestion by the addition of 3 ml dH2O, 500 µl 10X Digestion buffer (200 mM Tris-HCl pH 7.5, 100 mM MgCl2, 10 mM DTT, 1 mg/ml BSA) and 500 µl 10 % Triton-X-100 (Thermo-Fisher). After thoroughly mixing the reaction, 400 µl was removed and transferred to a 1.5 ml Eppendorf tube as a non-digested (ND) control. The restriction enzyme, HpaII (New England Biolabs, 1000 U) was then added to the remaining sample and the tube incubated with gentle agitation for 3h at 37°C.

Choice of restriction enzymes for Hi-C

The frequency of restriction sites correlates with the GC-content of a genome and thus the choice of restriction enzymes can introduce significant bias into 3C/Hi-C contact maps (Cournac et al., 2012; Marbouty et al., 2017). For this reason, eukaryotic Hi-C protocols typically use DpnII (^GATC) to offer more uniform coverage, but this is not possible for bacterial species, such as *E. coli* and *V. cholera,* whose genomes are subject to Dam methylation. For this reason, the first bacterial Hi-C studies utilized 6 bp cutters, such as BglII (A^GATCT) and NcoI (C^CAATG) to digest the genome, but a limited number of cut sites in the bacterial genomes resulted in a maximum Hi-C resolution of 10 kb (**Table**

S1) (Le et al., 2013). We have previously utilized both HpaII (GGC^C) and MluCI (A^ATT) for bacterial 3C and metagenomic experiments (Marbouty et al., 2014, 2017) and both these enzymes have between ~7,700 - 31,000 sites in all five of the organisms in this study, generating an average fragment length of 188 bp (**Table S1**). We decided to test these two enzymes in combination with biotinylation to see if we could improve our Hi-C resolution, we also combined both HpaII and MluCI in the same experiment, which would, in theory, reduce the average fragment size to ~82 bp, further increasing the resolution of the Hi-C data and covering both AT and GC-rich regions of the genome in the same experiment. The biotinylation step and subsequent streptavidin-enrichment were done using Biotin-14-dCTP and Biotin-14-dATP for HpaII and MluCI, respectively. For the experiments performed using both enzymes, we used the same quantity of enzyme as the single-enzyme experiments and combined the two biotinylated nucleotides in equal amounts for the enrichment step. Compared to previously published data, the combination of the four base cutters with biotinylation did improve the number of informative reads for each experiment (**Figure S1**), with HpaII and MluCI performing well. As expected, MluCI failed to digest the GC-rich *H. volcanii* genome and in all organisms, the combination of the two enzymes didn't add to the resolution, if anything it made the contact maps nosier than using either enzyme alone (**Figure S1**).

Biotinylation of DNA ends

Following digestion, 400µl of the sample was removed and transferred to a 1.5ml Eppendorf as the digested (D) control. In contrast to previous bacterial Hi-C studies (Le et al., 2013) the remaining lysate was centrifuged (16,000 x g, 20 min, room temperature) to pellet the insoluble fraction containing the protein-DNA complexes of interest. We found that only taking the insoluble fraction significantly improved the quality of the final Hi-C data. After removing the supernatant, the pellet was resuspended in 400 µl dH2O. To this the following was added; 50 µl 10x Ligation Buffer (500 mM Tris-HCl pH 7.5, 100 mM MgCl2, 100 mM DTT), 4.5 µl 10 mM dAGTTp, 37.5 µl Biotin-14-dCTP (Thermo Fisher), 50 Units of DNA Polymerase I - Large Klenow Fragment (New England Biolabs). After briefly mixing, the reaction was incubated with agitation for 45 min at 37°C.

Blunt-end ligation

For further gains, we took inspiration from published human protocols and reduced the volume of the blunt-end ligation step by 20-fold to mimic the conditions of 'in situ Hi-C' protocols (Gavrilov et al., 2013; Rao et al., 2014). Even though previous Hi-C and 3C studies in bacteria had performed ligations under dilute conditions (Le et al., 2013; Lioy et al., 2018; Marbouty et al., 2015), the rationale is that in a smaller volume there would be less spurious events than in dilute conditions and, as a result, less noise in the resulting Hi-C matrices. Following further trial-and-error optimization, we performed the ligation at room temperature and reduced its duration (**Figure 1A** and **Figure S1**). Once the fill-in reaction was complete, the ligation was set up by adding the following; 120µl 10x Ligation Buffer, 12 µl 10 mg/ml BSA, 12 µl 100mM ATP, 540 µl dH2O, 480 U T4 DNA ligase (Thermo Fisher). The reaction was mixed gently and then incubated with gentle agitation for 3h at room temperature.

Reverse crosslinking and DNA purification

Following ligation, proteins were denatured by the addition of 20 µl 500 mM EDTA, 20 µl 10 % SDS, and 100 µl 20 mg/ml proteinase K (EuroBio). The non-digested (ND) and digested (D) controls were also treated using 20 µl 500 mM EDTA, 20 µl 10 % SDS and 10 µl 20mg/ml proteinase K (EuroBio). All samples were incubated at 65°C overnight to reverse formaldehyde-mediated protein-DNA crosslinks.

The following day, DNA was purified by the addition of an equal volume of Phenol:Chloroform: Isoamyl alcohol (25:24:1, Amresco), the mixture was vortexed for 30s and then centrifuged (12,000 x g, 5 min, room temperature). Following centrifugation, the upper aqueous phase was carefully removed and transferred to a new Eppendorf tube. DNA was precipitated by the addition of 2.5 x volume of ice-cold 100 % EtOH and 1/10 volume of 3 M NaOAc (pH 5.0). Samples were then incubated at -80°C for 30 min. Precipitated DNA was then pelleted by centrifugation (12,000 x g, 20 min, 4°C), the supernatant removed and the pellets washed in 500 µl 70% EtOH. After a second centrifugation (12,000 x g, 5 min, 4°C), the EtOH was removed and the pellets dried on a 37°C heat block for 5-10 min. Once the remaining EtOH had evaporated, pellets were resuspended in 140 µl 1X TE buffer + 1mg/ml RNase

(Euromedex) and incubated for 30 min at 37°C with agitation. When the DNA had been completely resuspended, 10 µl of HiC libraries and the ND and D controls were checked on a 1 % agarose gel (**Figure S2E**). This step is included to ensure the DNA isn't degraded, was successfully digested by the restriction enzyme, and subsequently ligated after biotin integration. Once the quality of the HiC library has been confirmed, the ND and D controls can be discarded.

Preparation of sequencing libraries

DNA sonication and size-selection

For efficient sonication, a maximum of 5 µg of DNA should be used, if the HiC library exceeds this, then an aliquot should be taken and the remaining DNA stored at -20°C, as a backup. Once the correct quantity of DNA was obtained, the sample volume was adjusted to 130 µl using 1x TE buffer. HiC samples were transferred to a sonication tube (microTUBE AFA Fiber Pre-Slit Snap-Cap, Covaris). The DNA was then sheared using the Covaris S220 Focused Ultrasonicator to yield an average fragment size of 300 bp (140 W peak incidence power, 10 % duty factor, 200 cycles per burst, 7°C). Following sonication, the DNA was transferred to a 1.5 ml Eppendorf tube and an equal volume of AmPure XP beads (Beckman) added. Although previous bacterial 3C/Hi-C protocols have utilized PCR purification columns for all DNA clean-up/size-selection steps (Le et al., 2013; Lioy et al., 2018; Marbouty et al., 2015), we drew inspiration from eukaryotic protocols (Rao et al., 2014) and switched to AmPure beads for all DNA purification steps. We found that significantly increased our DNA yield and ultimately the quality of the final bacterial Hi-C library. The sample was mixed 10x by gentle pipetting and then incubated for 5 min at room temperature to allow the DNA fragments to bind to the magnetic beads. The tube was then transferred to a magnetic rack for 1 min or until the beads separated to the wall of the tube. The supernatant was then carefully removed and the beads washed 2x using freshly-prepared 70 % EtOH. The pellet was then allowed to airdry for 30s, to remove residual ethanol, before being resuspended in 320 µl of elution buffer (10 mM Tris-HCl, pH 8,5). Size-selection was checked by running 18 µl of the DNA on a 1% agarose gel (**Figure S2E**) and DNA quantity was checked by QuBit analysis, the remaining 300 µl was taken for sequencing library preparation.

Biotin pull-down

The biotin-streptavidin pull-down maximizes the number of "true" HiC contacts in an experiment by selecting for chimeric DNA molecules that have been digested, biotin-filled and blunt-end ligated. To do this, the HiC library was bound to Streptavidin C1 Dynabeads (Thermo Fisher). First, the beads were mixed well by vortexing and then 30 µl of beads were aliquoted to a 1.5 ml Eppendorf. The beads were placed on a magnet and left for 1 min to clear the supernatant, the supernatant was then removed and 500 µl of 1x Tween Wash Buffer (TWB: 5 mM Tris-HCl pH7.5, 0.5 mM EDTA, 1M NaCl, 0.05% Tween 20) was added to rinse the beads. The solution was then separated on a magnet and the supernatant discarded, the beads were then re-suspended in 300 µl of 2x Binding Buffer (BB: 10 mM Tris-HCl pH7.5, 1 mM EDTA, 2 M NaCl), 300 µl of HiC sample was added and the mixture incubated on a tube rotator for 15 min at room temperature. The tube was then placed back on the magnet for 1 min to separate the biotinylated DNA bound to the streptavidin beads from the supernatant containing unbiotinylated DNA. The beads were then washed twice in 500 µl TWB in a thermomixer (2 min, 55°C, 1,000 rpm). Finally, the beads were resuspended in 100 µl of 1x T4 ligase buffer (New England Biolabs), and the mixture transferred to a new 1.5 ml Eppendorf.

End-repair

The end-repair mix was prepared as follows: 85 µl 1x T4 ligase buffer, 5 µl 10mM dNTPs, 50 U T4 PNK, 12 U T4 DNA polymerase, 5 U DNA Polymerase I, Large Klenow Fragment. After the HiC bank was placed on the magnet and the supernatant removed, the beads were resuspended in the end repair mix and the reaction incubated for 30 min at room temperature.

A-tailing

Following end-repair, the beads were washed twice in 500 µl TWB and once in 100 µl 1x NEB2 buffer using a thermomixer (2 min, 55°C, 1000 rpm). The beads containing the end-repaired DNA were then resuspended in 100 µl of A-tailing Mix (90 µl 1xNEB 2 buffer, 0.5 mM dATP, 25 U Klenow Fragment (3'-5'exo-)) and incubated for 30 min at 37°C.

Adapter ligation

After A-tailing, the beads were washed twice in 500 µl TWB and once in 50 µl 1x Quick Ligase Buffer (New England Biolabs) using a thermomixer (2 min, 55°C, 1000 rpm). The beads were then resuspended in the Ligation Mix (50 µl 1x Quick Ligase Buffer, 2 µl Quick DNA Ligase) and 2 µl of sequencing adapter was added (NEXTflex Illumina Barcodes, Bioo Scientific). The reaction was mixed well and then incubated at room temperature for 10 min, the beads were washed twice in 500 µl TWB and once in 100 µl of 10 mM Tris-HCl pH 8 using a thermomixer (2 min, 55°C, 1000 rpm). Finally, the beads were resuspended in 50 µl of 10 mM Tris-HCl pH 8 and transferred to a new tube.

Library amplification by PCR

The PCR reaction was set up as follows: 40 µl of Phusion 2x High Fidelity Master Mix (New England Biolabs), 5 µl of 2 µM Primer Mix (NEXTflex, Bioo Scientific), 32 µl dH2O and 3 µl Streptavidin-beads containing the HiC library. The library was amplified for 12 cycles following the manufacturers' instructions. Reducing the number of PCR cycles, compared to previous bacterial Hi-C studies (Le et al., 2013), significantly reduced (~10-fold) the number of PCR duplicates in the final Hi-C library (**Table S1**). After PCR, the reaction was placed on a magnet for 1 min to separate the streptavidin beads from the PCR reaction containing the amplified library. The supernatant was transferred to a new tube and the beads discarded. The PCR reaction was then purified by adding an equal volume of AmPure XP beads, the reaction was then mixed well and incubated at room temperature for 5 min. Following two washes with 80 % EtOH, the beads were mixed with 50 µl of 10 mM Tris-HCl pH 8 and incubated for 5 min at room temperature. Finally, the reaction was placed on the magnet and the DNA-containing supernatant transferred to a new Eppendorf tube. The DNA was then checked on a 1% agarose gel to determine the size of the final library and to check for the presence of primer dimers. Once this was confirmed, the Hi-C library was ready to prepare for paired-end sequencing using the 75 cycle High-Output Kit v2.5. Sequencing was performed on a Nextseq500 according to Illumina's instructions.

Novaseq sequencing of Hi-C libraries

To determine if the plaid patterning observed in *H. volcanii* Hi-C contact maps was the result of using Illumina short-read sequencing (PE35), we re-sequenced one of the wild-type exponential phase samples (CC402) using Illumina PE150 sequencing, performed by Novogene. The data obtained were processed using the same pipeline as all other libraries.

QUANTIFICATION AND STATISTICAL ANALYSIS

Processing of the reads, computation of contact matrices, and generation of contact maps

All data were processed as previously described (Cournac et al., 2012; Lioy et al., 2018) and statistics relating to each of the processing steps are listed in **Table S1**. Reads were aligned and the contact data processed using a custom-made pipeline from the lab (https://github.com/koszullab/E_coli_analysis). So that our datasets could be compared to those previously published, the raw reads were downloaded from the SRA database (https://www.ncbi.nlm.nih.gov/sra) and re-processed using our pipeline **(See Key Resources Table** and **Table S1)**. Briefly, PCR duplicates were removed and the reads aligned independently to the reference genome (*E. coli*: GCF_000005845.2, *V. cholera*: GCF_000006745.1, *H. volcanii*: GCF_000025685.1, *Hbt salinarum*: GCF_000006805.1, *T. kodakaraensis*: GCF_000009965.1) by Bowtie2 (Langmead and Salzberg, 2012) using a very-sensitive and iterative approach. Each read was then assigned to a restriction fragment, with uninformative events, such as self-circularized and uncut fragments discarded by filtering. The filtered fragments were then binned into 500 bp, 1 kb, or 5 kb segments, and the corresponding contact maps were generated and normalized using the sequential component normalization (SCN) procedure (Cournac et al., 2012). By analyzing and re-analyzing so many 3C/Hi-C datasets from various species, all with very different growth conditions, we were able to make some interesting observations as to what constitutes a good Hi-C dataset **(Table S1)**. We noted that some growth conditions, such as stationary phase cultures or actinomycin-D treated cells, result in libraries that are resistant to restriction digest, which will in turn lead to an increased proportion of 'uncut' DNA fragments and a reduced number of reads in the final Hi-C matrix. However, this doesn't seem to bias the type of reads in the final contact map because the

proportion of inter- and intra-chromosomal contacts is maintained **(Table S1)**. We also observed that even if a high number of reads pass the Hi-C filtering steps, it doesn't guarantee high-resolution in the final contact map. This is due to the restriction enzyme used but also because some random ligation events are capable of passing through the Hi-C filtering steps **(Table 1).**

Calculating the generation time of wild-type and mutant strains

Overnight cultures were diluted to OD_{600nm} ~ 0.05 and incubated at 45°C. Every 2h, aliquots were collected and serial dilutions made in 18% salt water (18% SW) containing 2.5M NaCl, 90 mM $MgCl_2$ $6H_2O$, 90 mM $MgSO_4$ $7H_2O$, 60 mM KCl, 10 mM Tris HCl pH 7.5. 20 µl cell aliquots were then spotted on Hv-YPC plates and individual colonies were counted after 4 days of incubation at 45 °C. The curve CFU/ml = f(time) was drawn. One early time-point and one late time point were then selected to calculate generation time (G) using the formula: $G = [(t_{Late} - t_{Early})] / [(Log_{10}(N_{Late}) - Log_{10}(N_{Early}) / Log_{10}(2)]$, t designating time and N the number of colonies per ml. Generation time measurements were performed at three independent times for each strain, to provide associated standard errors.

Wide-field microscopy and DNA staining of *H. volcanii* cells

DNA was labeled using the fluorescent Hoechst 33342 dye (excitation 361nm / emission 486 nm). Exponentially growing cells were centrifuged and resuspended in an equal volume of 18% SW containing 5 µg *per* ml of Hoechst 33342. Cells were incubated for 10 min in the dark and then mounted onto glass slides covered with a thin layer of 1% agarose (suspended in 18% SW solution). DIC and fluorescent images were obtained at room temperature using a ZEISS Axio Observer microscope equipped with a 40x, 1.4 NA oil immersion objective. 365 nm excitation at the maximum available intensity (2 $W.cm^{-2}$) and filter set 49 (EX G 365, BS FT 395, EM BP 445/50) were used for fluorescence imaging of Hoechst signal. A Z-stack (30 slices) centered on DIC focus was performed and data was collected sequentially for DIC and Hoechst signals at each plane. The optical slice with the maximum number of cells at the focus was chosen for further study.

Generation of ratio maps

We compared the contact maps of different conditions by computing their ratio. To do this the maps were first SCN normalized (Cournac et al., 2012) and then the ratio for each bin was calculated by dividing the number of contacts in one condition by the number of contacts in the other condition. The Log2 ratio of this was then plotted with a Gaussian filter. The blue to red colorscale reflects the enrichment in contacts in one condition with respect to the other, white indicates no difference.

Correlation of Hi-C contact maps and transcription

To investigate the correlation between transcription and chromosome contacts in WT *H. volcanii*, we took the corresponding RNA-seq dataset and mapped the reads against the reference genome. Only reads with a mapping quality > 30 were conserved. The signal was then binned to match the binning of the corresponding scn normalized (Cournac et al., 2012) contact maps (5kb) and the two datasets plotted according to their genomic coordinates (Figure 2E). Both the Hi-C contact and transcription signals were smoothed with a Savitzky-Golay filter and the Pearson coefficient was determined.

RNA-seq

Total RNA was extracted from *H. volcanii* using the Nucleospin RNA Extraction Kit (Macherey-Nagel) according to manufacturers' instructions. After an additional DNase treatment with Turbo DNase (Thermo Fisher), the RNA was purified by phenol extraction (pH 4.5, Amresco) and ethanol precipitation. The RNA was then resuspended in DEPC-treated water. Ribosomal RNA depletion, cDNA library preparation, and Illumina PE150 sequencing were performed by Novogene. Three biological replicates were generated for each condition and on average ~7million reads were generated per sample. Clean reads were aligned to the reference genome with Bowtie2 using default parameters (Kim et al., 2019). Reads were then processed using SAMtools (Li et al., 2009), to convert SAM files to BAM files, to sort BAM files by read name (-n option), and then to index them. For comparisons within the same sample, bamCoverage (implemented in deepTools2; (Ramírez et al., 2016)) was used to calculate Reads per Kilobase region per Million mapped reads (RPKM).

Differential expression analysis

The sorted BAM files and a GFF annotation file of the reference genome were used as inputs for HtSeq-count (Anders et al., 2014) to obtain read counts per transcript. HTSeq-count was run with default parameters, except that the strand option was set to suit dUTP libraries (-s reverse) and the mapping quality score was increased (-a 30). The analysis was performed using RStudio (RStudio Team, 2020) (version: 1.3.959) and R (version: 4.0.1). DeSeq2 (Love et al., 2014) (version: 1.28.1) was used to import read counts for the investigation of transcriptomic differences between different strains/treatments. Differentially expressed genes were defined as those with an adjusted p-value < 0.05 and log2 fold change in expression of >2 or < -2. The full list of differentially expressed genes for each condition has been uploaded to the SRA database **(Key Resources Table)**. For the Differential expression analysis of actinomycin D compared to the untreated control, 26% of genes were upregulated and 31% of genes were downregulated. We mostly focused on regions of the genome that showed a decrease in gene expression, but one would assume that the upregulated genes are involved in the stress response. However, because many of the genes in the *H. volcanii* genome are yet to be assigned a function, we were unable to determine if this was indeed the case. When stationary phase cells were compared to exponentially growing cells, 54% of genes showed >2-fold change in gene expression. In *smc* mutant cells < 5 % of genes with a >2-fold change in gene expression compared to WT cells in the same growth phase. Full details of all R scripts are available upon request.

Directional index analysis

The directional index (DI) is a statistical parameter that quantifies the degree of upstream and downstream contact biases for a genomic region(Dixon et al., 2012). It is based on a t-test between vectors to the left and right of each bin, up to a certain scale. We utilized a workflow that has been described previously (Lioy et al., 2018) to perform a DI analysis for domains consistent with the scale of bacterial (100 kb) and larger structures such as macrodomains (300 kb).

Identification of active / inactive compartments and compartment-like structures

The compartment index was calculated as described previously (Lieberman-Aiden et al., 2009). To facilitate the comparison of Euryarchaeal compartments with those previously published in Crenarchaea (Takemata et al 2019), the raw reads from the latter were downloaded from the SRA database (https://www.ncbi.nlm.nih.gov/sra) and re-processed using our pipeline **(See Key Resources Table** and **Table S1)**. For each Hi-C matrix, a distance-normalized matrix was generated. These matrices were then converted into Pearson correlation matrices and from this, the eigenvalues and first eigenvector were calculated. The eigenvector allows the A and B compartments to be attributed. Positive and negative eigenvalues were arbitrarily chosen to correspond to the A and B compartments respectively.

Calling chromosomal loops and borders using Chromosight

Chromosight 1.1.2 (Matthey-Doret et al., 2020) is an algorithm based on computer vision approaches that are capable of automatically detecting common patterns seen in Hi-C contact maps. These patterns are then associated with a score (Pearson coefficient) which allows for the interpretation of results and the comparison of different mutants/treatments. We applied the Chromosight "*detect"* function to each of our *H. volcanii* matrices (balanced contact maps, 2 kb binning) and searched for loops (Pearson: 0.38, minimum loop length: 10000, maximum loop length: 250000), borders (Pearson: 0.2), and hairpins (default parameters). We detected both loops and borders in our Hi-C matrices but failed to detect any hairpin structures. We then used the Chromosight *"quantify"* function, which quantifies the similarity of the loop and border scores between wild-type and mutant/treated cells to determine the effect on these finer chromosome structures. The quantify mode was used with default parameters, except that all of the contact maps were subsampled to contain the same number of reads.

Phylogenetic tree of SMCs, Rad50, and ClsN in different archaeal lineages

The tree shown in Figure 4 has been adapted from two previous studies which identified (Spang et al., 2017; Takemata et al., 2019) condensin-type SMCs using a blastp search, with *Pyrococcus furiosus* SMC as the query and searching against each archaea species listed. The addition of details about the

conservation of Rad50 in the different archaeal lineages. To identify probable homologs of Rad50, we used the sequence of human Rad50 as a query (accession number: AAB07119.1) to perform a BLAST search against new archaeal genomes (Hopfner et al., 2000). We then classified the hits into one of four groups: i) present, ii) absent, iii) distant homologs present in all species, iv) distant homologs present in some species. From this, we conclude that, like canonical SMCs, Rad50 is present in many euryarchaea and although Rad50 has been observed in the Asgard and Tack groups, it seems to be less conserved for the species within these groups.

References:

Allers, T., Ngo, H.-P., Mevarech, M., and Lloyd, R.G. (2004). Development of Additional Selectable Markers for the Halophilic Archaeon *Haloferax volcanii* Based on the leuB and trpA Genes. Appl. Environ. Microbiol. *70*, 943–953.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Ausiannikava, D., Mitchell, L., Marriott, H., Smith, V., Hawkins, M., Makarova, K.S., Koonin, E.V., Nieduszynski, C.A., and Allers, T. (2018). Evolution of Genome Architecture in Archaea: Spontaneous Generation of a New Chromosome in Haloferax volcanii. Mol. Biol. Evol. *35*, 1855–1868.

Barillà, D. (2016). Driving Apart and Segregating Genomes in Archaea. Trends Microbiol. *24*, 957–967.

Bitan-Banin, G., Ortenberg, R., and Mevarech, M. (2003). Development of a Gene Knockout System for the Halophilic Archaeon Haloferax volcanii by Use of the pyrE Gene. J. Bacteriol. *185*, 772–778.

Blattner, F.R., Plunkett, G., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., et al. (1997). The Complete Genome Sequence of *Escherichia coli* K-12. Science *277*, 1453–1462.

Boccard, F., Esnault, E., and Valens, M. (2005). Spatial arrangement and macrodomain organization of bacterial chromosomes. Mol. Microbiol. *57*, 9–16.

Böhm, K., Giacomelli, G., Schmidt, A., Imhof, A., Koszul, R., Marbouty, M., and Bramkamp, M. *(2020). Chromosome organization by a conserved condensin-ParB system in the actinobacterium Corynebacterium glutamicum.* Nat. Commun. *11*, 1–17.

Bouthier de la Tour, C., Toueille, M., Jolivet, E., Nguyen, H.-H., Servant, P., Vannier, F., and Sommer, S. (2009). The *Deinococcus radiodurans* SMC protein is dispensable for cell viability yet plays a role in DNA folding. Extrem. Life Extreme Cond. *13*, 827–837.

Cournac, A., Marie-Nelly, H., Marbouty, M., Koszul, R., and Mozziconacci, J. (2012). Normalization of a chromosomal contact map. BMC Genomics *13*, 436.

Cox, M.M., and Battista, J.R. (2005). *Deinococcus radiodurans* — the consummate survivor. Nat. Rev. Microbiol. *3*, 882–892.

Delmas, S., Duggin, I.G., and Allers, T. (2013). DNA damage induces nucleoid compaction via the Mre11-Rad50 complex in the archaeon *Haloferax volcanii*. Mol. Microbiol. *87*, 168–179.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Egan, E.S., Fogel, M.A., and Waldor, M.K. (2005). Divided genomes: negotiating the cell cycle in prokaryotes with multiple chromosomes. Mol. Microbiol. *56*, 1129–1138.

Gavrilov, A.A., Gushchanskaya, E.S., Strelkova, O., Zhironkina, O., Kireev, I.I., Iarovaia, O.V., and Razin, S.V. (2013). Disclosure of a structural milieu for the proximity ligation reveals the elusive nature of an active chromatin hub. Nucleic Acids Res. *41*, 3563–3575.

Gorlas, A., Alain, K., Bienvenu, N., and Geslin, C. (2013). *Thermococcusprieurii* sp. nov., a hyperthermophilic archaeon isolated from a deep-sea hydrothermal vent (Microbiology Society,).

Gorlas, A., Croce, O., Oberto, J., Gauliard, E., Forterre, P., and Marguet, E. (2014). *Thermococcusnautili* sp. nov., a hyperthermophilic archaeon isolated from a hydrothermal deep-sea vent (Microbiology Society,).

Güthlein, C., Wanner, R.M., Sander, P., Böttger, E.C., and Springer, B. (2008). A Mycobacterial *smc* Null Mutant Is Proficient in DNA Repair and Long-Term Survival. J. Bacteriol. *190*, 452–456.

Harrison, P.W., Lower, R.P.J., Kim, N.K.D., and Young, J.P.W. (2010). Introducing the bacterial "chromid": not a chromosome, not a plasmid. Trends Microbiol. *18*, 141–148.

Hartman, A.L., Norais, C., Badger, J.H., Delmas, S., Haldenby, S., Madupu, R., Robinson, J., Khouri, H., Ren, Q., Lowe, T.M., et al. (2010). The Complete Genome Sequence of *Haloferax volcanii* DS2, a Model Archaeon. PLOS ONE *5*, e9605.

Hawkins, M., Malla, S., Blythe, M.J., Nieduszynski, C.A., and Allers, T. (2013). Accelerated growth in the absence of DNA replication origins. Nature *503*, 544–547.

Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L., Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature *406*, 477–483.

Hofman, J.D., Schalkwyk, L.C., and Doolittle, W.F. (1986). ISH51: a large, degenerate family of insertion sequence-like elements in the genome of the archaehacterium, *Halobacterium volcanii*. Nucleic Acids Res. *14*, 6983–7000.

Hopfner, K.-P., Karcher, A., Shin, D., Fairley, C., Tainer, J.A., and Carney, J.P. (2000). Mre11 and Rad50 from Pyrococcus furiosus: Cloning and Biochemical Characterization Reveal an Evolutionarily Conserved Multiprotein Machine. J. Bacteriol. *182*, 6036–6041.

Hsieh, T.-H.S., Weiner, A., Lajoie, B., Dekker, J., Friedman, N., and Rando, O.J. (2015). Mapping Nucleosome Resolution Chromosome Folding in Yeast by Micro-C. Cell *162*, 108–119.

Hundt, S., Zaigler, A., Lange, C., Soppa, J., and Klug, G. (2007). Global Analysis of mRNA Decay in *Halobacterium salinarum* NRC-1 at Single-Gene Resolution Using DNA Microarrays. J. Bacteriol. *189*, 6936.

Jäger, A., Samorski, R., Pfeifer, F., and Klug, G. (2003). Individual gvp transcript segments in *Haloferax mediterranei* exhibit varying half-lives, which are differentially affected by salt concentration and growth phase. Nucleic Acids Res. *30*, 5436–5443.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915.

Koonin, E.V., and Wolf, Y.I. (2008). Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. Nucleic Acids Res. *36*, 6688–6719.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Le, T.B., and Laub, M.T. (2016). Transcription rate and transcript length drive formation of chromosomal interaction domain boundaries. EMBO J. e201593561.

Le, T.B.K., Imakaev, M.V., Mirny, L.A., and Laub, M.T. (2013). High-Resolution Mapping of the Spatial Organization of a Bacterial Chromosome. Science *342*, 731–734.

Li, M.Z., and Elledge, S.J. (2007). Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. Nat. Methods *4*, 251–256.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinforma. Oxf. Engl. *25*, 2078–2079.

Lieberman-Aiden, E., Berkum, N.L. van, Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. Science *326*, 289–293.

Lioy, V.S., Cournac, A., Marbouty, M., Duigou, S., Mozziconacci, J., Espéli, O., Boccard, F., and Koszul, R. (2018). Multiscale Structuring of the *E. coli* Chromosome by Nucleoid-Associated and Condensin Proteins. Cell *172*, 771-783.e18.

López-García, P., Abad, J.P., Smith, C., and Amils, R. (1992). Genomic organization of the halophilic archaeon *Haloferax mediterranei*: physical map of the chromosome. Nucleic Acids Res. *20*, 2459–2464.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Makarova, K.S., and Koonin, E.V. (2013). Archaeology of eukaryotic DNA replication. Cold Spring Harb. Perspect. Biol. *5*, a012963.

Marbouty, M., Cournac, A., Flot, J.-F., Marie-Nelly, H., Mozziconacci, J., and Koszul, R. (2014). Metagenomic chromosome conformation capture (meta3C) unveils the diversity of chromosome organization in microorganisms.

Marbouty, M., Le Gall, A., Cattoni, D.I., Cournac, A., Koh, A., Fiche, J.-B., Mozziconacci, J., Murray, H., Koszul, R., and Nollmann, M. (2015). Condensin- and Replication-Mediated Bacterial

Chromosome Folding and Origin Condensation Revealed by Hi-C and Super-resolution Imaging. Mol. Cell *59*, 588–602.

Marbouty, M., Baudry, L., Cournac, A., and Koszul, R. (2017). Scaffolding bacterial genomes and probing host-virus interactions in gut microbiome by proximity ligation (chromosome capture) assay. Sci. Adv. *3*, e1602105.

Matthey-Doret, C., Baudry, L., Breuer, A., Montagne, R., Guiglielmoni, N., Scolari, V., Jean, E., Campeas, A., Chanut, P.H., Oriol, E., et al. (2020). Computer vision for pattern detection in chromosome contact maps. Nat. Commun. *11*, 5795.

Mazel, D., Dychinco, B., Webb, V.A., and Davies, J. (1998). A Distinctive Class of Integron in the *Vibrio cholerae* Genome. Science *280*, 605.

Mullakhanbhai, M.F., and Larsen, H. (1975). Halobacterium volcanii spec. nov., a Dead Sea *halobacterium* with a moderate salt requirement. Arch. Microbiol. *104*, 207–214.

Ng, W.V., Ciufo, S.A., Smith, T.M., Bumgarner, R.E., Baskin, D., Faust, J., Hall, B., Loretz, C., Seto, J., Slagel, J., et al. (1998). Snapshot of a large dynamic replicon in a halophilic archaeon: megaplasmid or minichromosome? Genome Res. *8*, 1131–1141.

Ng, W.V., Kennedy, S.P., Mahairas, G.G., Berquist, B., Pan, M., Shukla, H.D., Lasky, S.R., Baliga, N.S., Thorsson, V., Sbrogna, J., et al. (2000). Genome sequence of *Halobacterium* species NRC-1. Proc. Natl. Acad. Sci. *97*, 12176.

Orlando, G., Kinnersley, B., and Houlston, R.S. (2018). Capture Hi-C Library Generation and Analysis to Detect Chromatin Interactions. Curr. Protoc. Hum. Genet. *98*, e63.

Pfeiffer, F., Losensky, G., Marchfelder, A., Habermann, B., and Dyall-Smith, M. (2020). Whole-genome comparison between the type strain of *Halobacterium salinarum* (DSM 3754T) and the laboratory strains R1 and NRC-1. MicrobiologyOpen *9*, e974.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. Cell *159*, 1665–1680.

RStudio Team (2020). RStudio: Integrated Development Environment for R (Boston, MA: RStudio, PBC.).

Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A., and Cavalli, G. (2012). Three-Dimensional Folding and Functional Organization Principles of the Drosophila Genome. Cell *148*, 458–472.

Spang, A., Caceres, E.F., and Ettema, T.J.G. (2017). Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. Science *357*, eaaf3883.

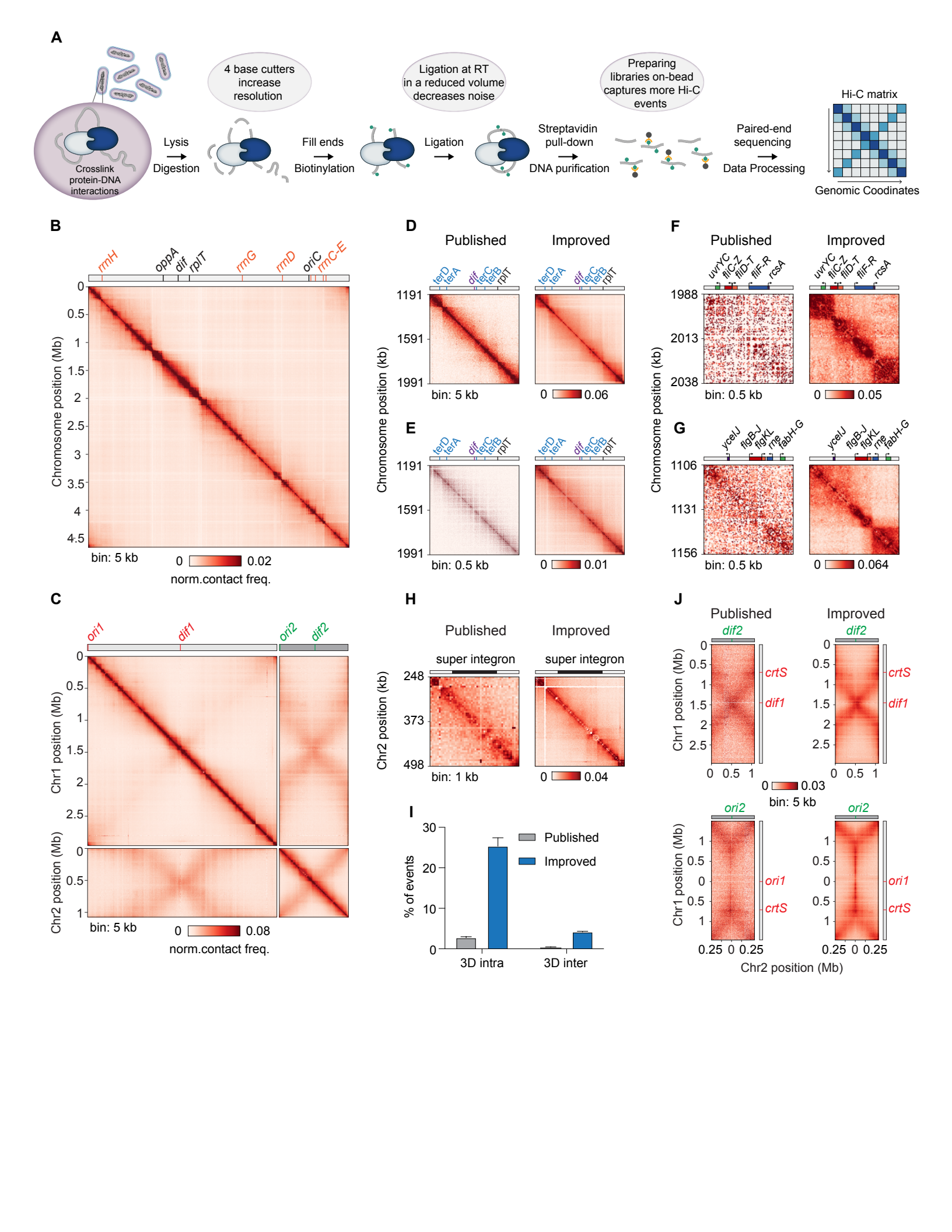
Takemata, N., and Bell, S.D. (2020). Multi-scale architecture of archaeal chromosomes. Mol. Cell.

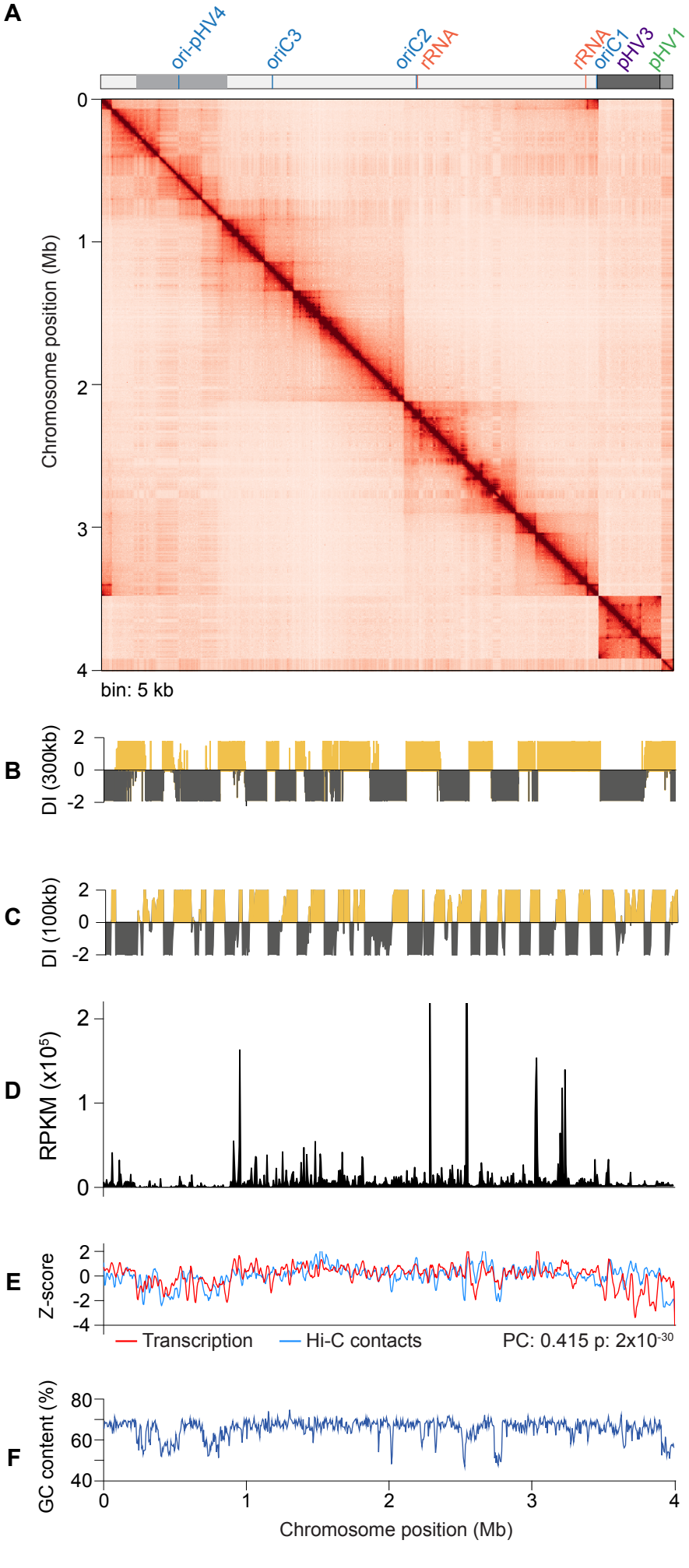
Takemata, N., Samson, R.Y., and Bell, S.D. (2019). Physical and Functional Compartmentalization of Archaeal Chromosomes. Cell *179*, 165-179.e18.

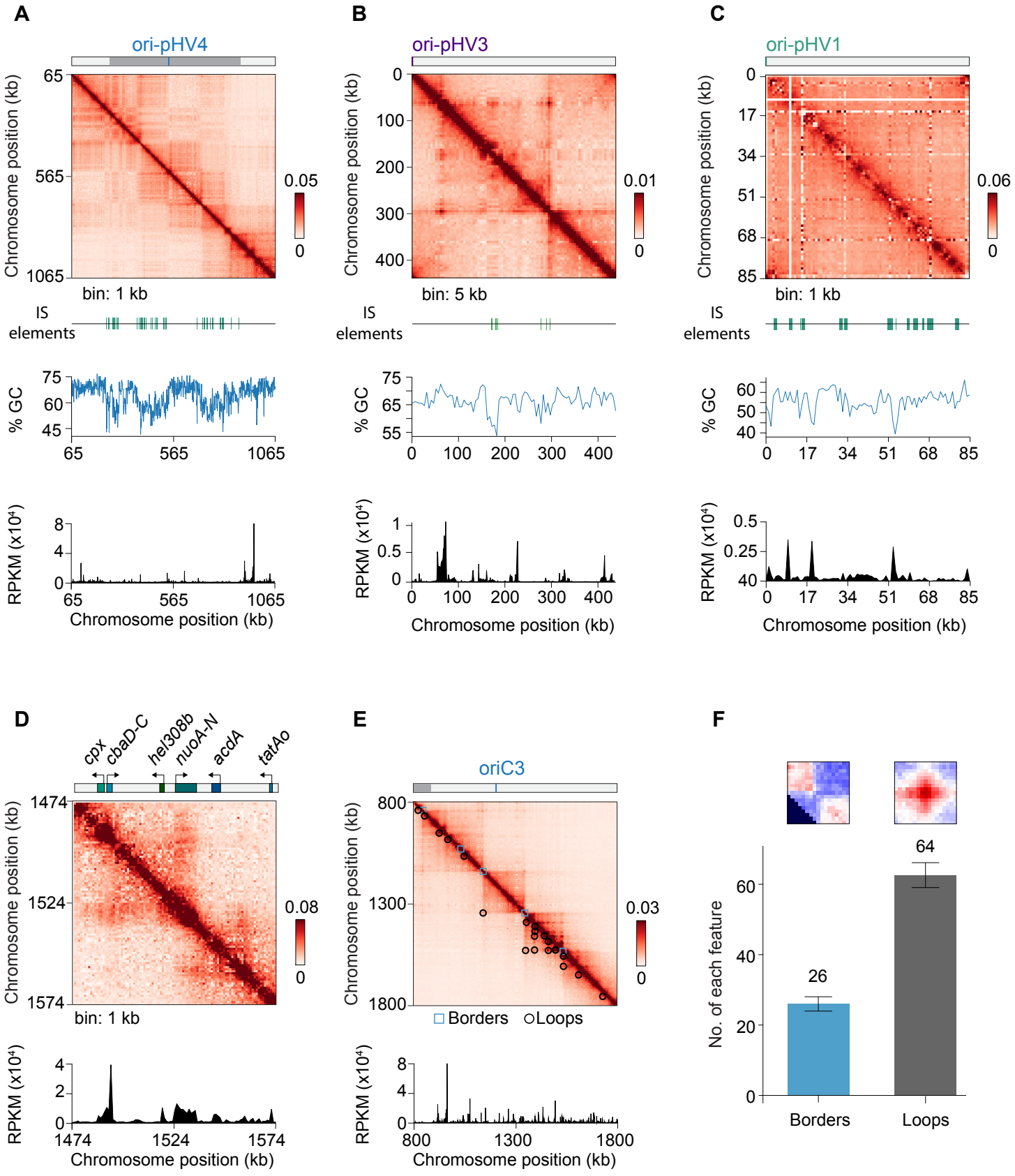
Val, M.-E., Marbouty, M., Martins, F. de L., Kennedy, S.P., Kemble, H., Bland, M.J., Possoz, C., Koszul, R., Skovgaard, O., and Mazel, D. (2016). A checkpoint control orchestrates the replication of the two chromosomes of *Vibrio cholerae*. Sci. Adv. *2*, e1501914.

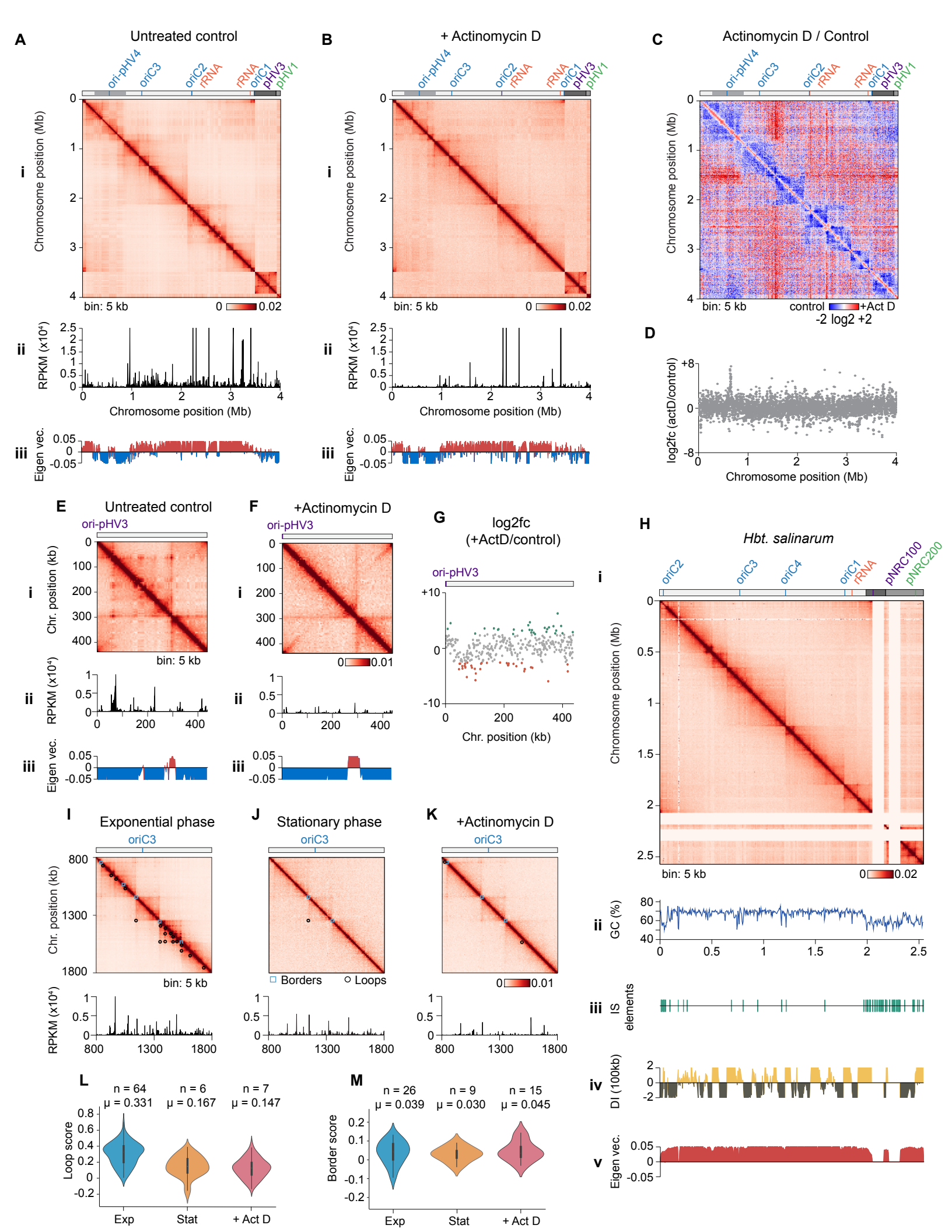
Wang, X., Le, T.B.K., Lajoie, B.R., Dekker, J., Laub, M.T., and Rudner, D.Z. (2015). Condensin promotes the juxtaposition of DNA flanking its loading site in Bacillus subtilis. Genes Dev. *29*, 1661–1675.

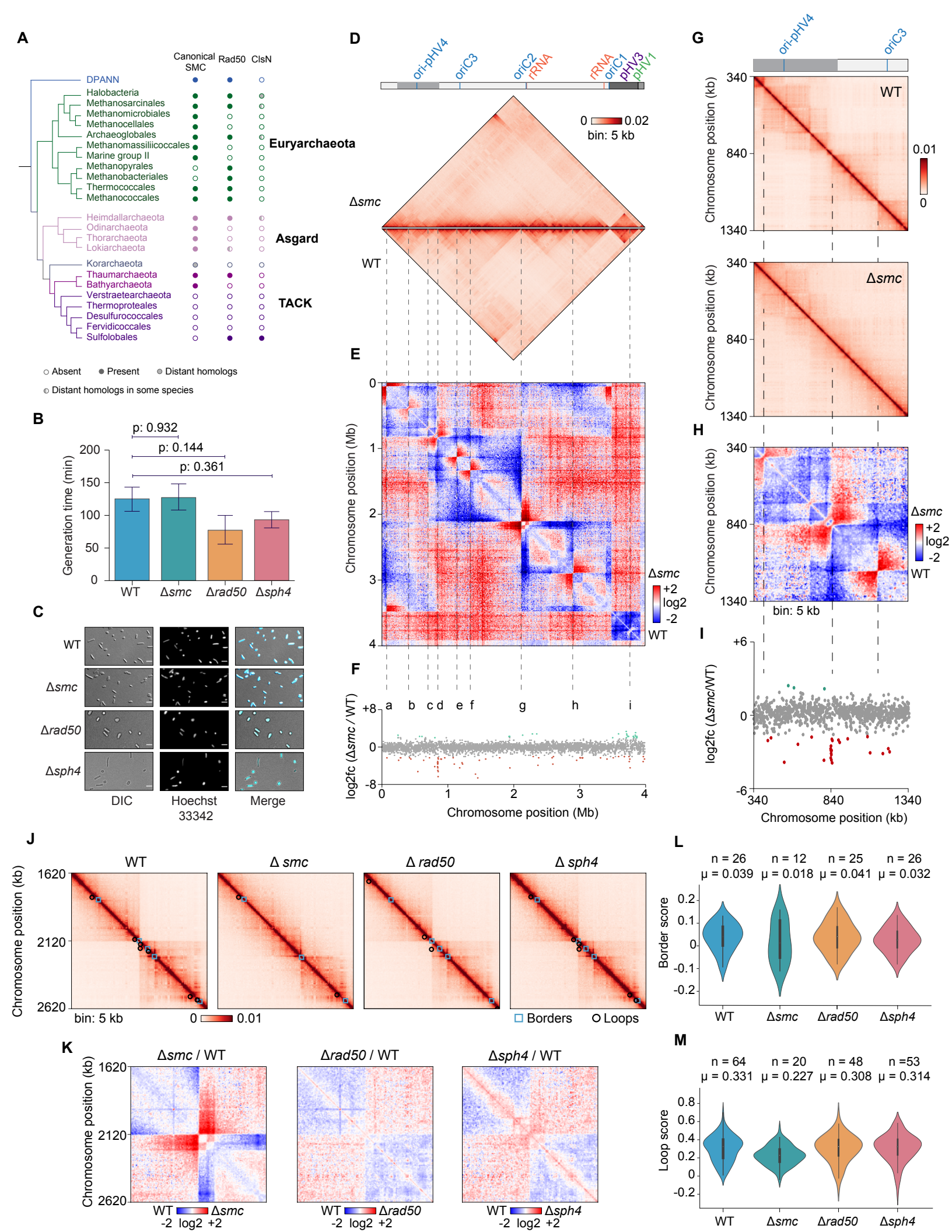
Wendoloski, D., Ferrer, C., and Dyall-Smith, M.L. (2001). A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2The GenBank accession number for the sequence reported in this paper is AF123438. (Microbiology Society,).


A
4 base cutters increase resolution
Ligation at RT in a reduced volume decreases noise
Preparing libraries on-bead captures more Hi-C events
Crosslink protein-DNA interactions
Lysis
Digestion
Fill ends
Biotinylation
Ligation
Streptavidin pull-down
DNA purification
Paired-end sequencing
Data Processing
Hi-C matrix
Genomic Coodinates
B
rrnH
oppA
dif
rplT
rrnG
rrnD
oriC
rrnC-E
Chromosome position (Mb)
bin: 5 kb
0
0.02
norm.contact freq.
C
ori1
dif1
ori2
dif2
Chr1 position (Mb)
Chr2 position (Mb)
bin: 5 kb
0
0.08
norm.contact freq.
D
Published
Improved
terD
terA
dif
terC
terB
rplT
Chromosome position (kb)
1191
1591
1991
bin: 5 kb
0
0.06
E
1191
1591
1991
bin: 0.5 kb
0
0.01
F
Published
Improved
uvrYC
fliC-Z
fliD-T
fliF-R
rcsA
Chromosome position (kb)
1988
2013
2038
bin: 0.5 kb
0
0.05
G
yceIJ
figB-J
figKL
rne
fabH-G
1106
1131
1156
bin: 0.5 kb
0
0.064
H
Published
Improved
super integron
Chr2 position (kb)
248
373
498
bin: 1 kb
0
0.04
I
% of events
30
20
10
0
Published
Improved
3D intra
3D inter
J
Published
Improved
dif2
crtS
dif1
Chr1 position (Mb)
0
0.03
bin: 5 kb
ori2
ori1
crtS
0.25 0 0.25
Chr2 position (Mb)

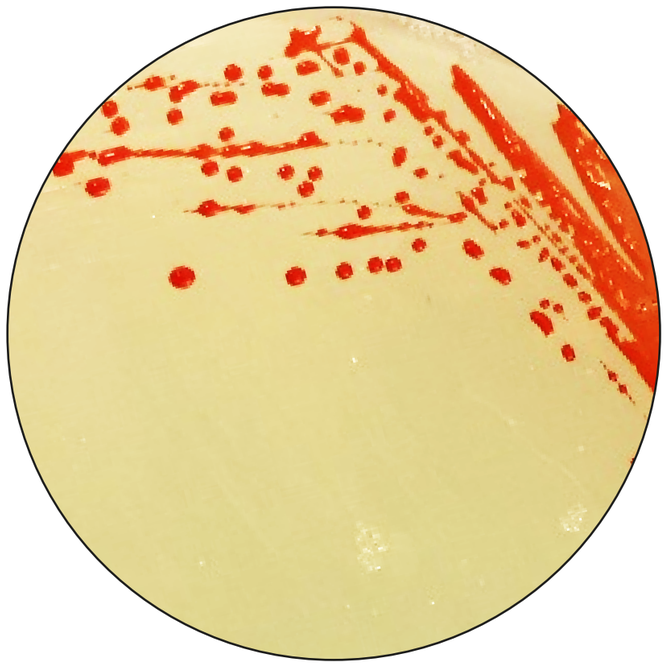

A
ori-pHV4
oriC3
oriC2
rRNA
rRNA
oriC1
pHV3
pHV1
Chromosome position (Mb)
bin: 5 kb
B
DI (300kb)
C
DI (100kb)
D
RPKM (x10^5)
E
Z-score
Transcription
Hi-C contacts
PC: 0.415 p: 2x10^-30
F
GC content (%)
Chromosome position (Mb)


A
ori-pHV4
Chromosome position (kb)
65
565
1065
0.05
0
bin: 1 kb
IS elements
% GC
75
60
45
RPKM (x10⁴)
8
4
0
Chromosome position (kb)
B
ori-pHV3
0
100
200
300
400
0.01
bin: 5 kb
55
65
C
ori-pHV1
0
17
34
51
68
85
0.06
bin: 1 kb
40
50
60
0.5
0.25
D
cpx
cbaD-C
hel308b
nuoA-N
acdA
tatAo
1474
1524
1574
0.08
bin: 1 kb
E
oriC3
800
1300
1800
0.03
Borders
Loops
F
No. of each feature
64
26
Borders
Loops

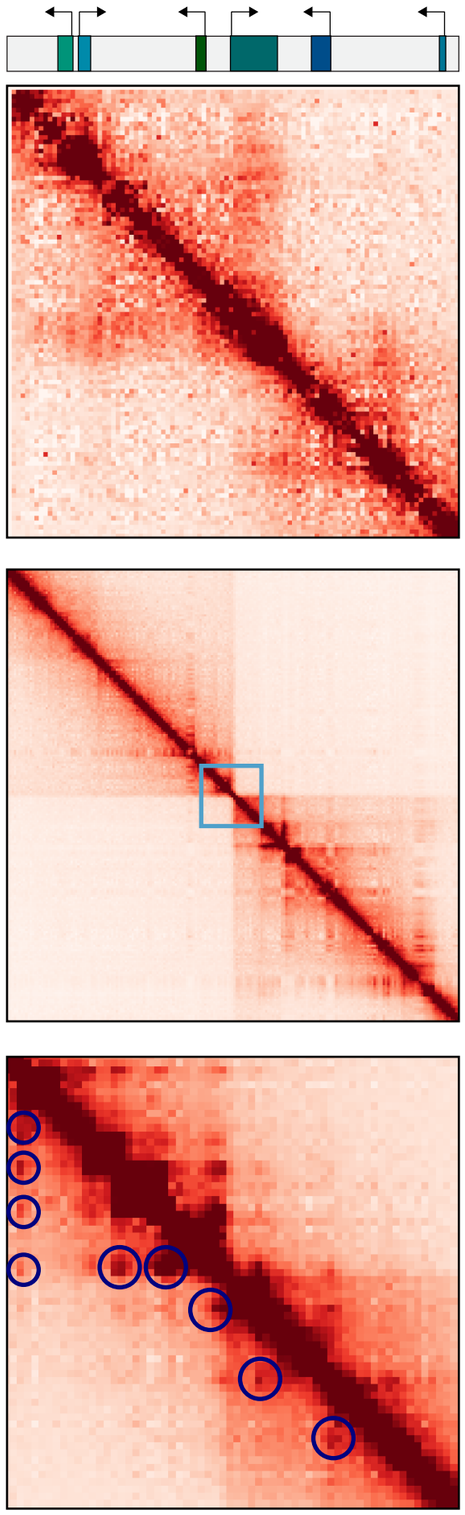

A
Untreated control
B
+ Actinomycin D
C
Actinomycin D / Control
ori-pHV4
oriC3
oriC2
rRNA
rRNA
oriC1
pHV3
pHV1
Chromosome position (Mb)
bin: 5 kb
0
0.02
control
+Act D
-2 log2 +2
i
ii
iii
RPKM (x10⁴)
Chromosome position (Mb)
Eigen vec.
D
log2fc (actD/control)
E
Untreated control
F
+Actinomycin D
G
log2fc
(+ActD/control)
ori-pHV3
Chr. position (kb)
0.01
Chr. position (kb)
H
Hbt. salinarum
oriC2
oriC3
oriC4
oriC1
rRNA
pNRC100
pNRC200
GC (%)
IS elements
DI (100kb)
v
I
Exponential phase
J
Stationary phase
K
+Actinomycin D
oriC3
Borders
Loops
L
n = 64
μ = 0.331
n = 6
μ = 0.167
n = 7
μ = 0.147
Loop score
Exp
Stat
+ Act D
M
n = 26
μ = 0.039
n = 9
μ = 0.030
n = 15
μ = 0.045
Border score


A
Canonical SMC
Rad50
ClsN
DPANN
Halobacteria
Methanosarcinales
Methanomicrobiales
Methanocellales
Archaeoglobales
Methanomassiliicoccales
Marine group II
Methanopyrales
Methanobacteriales
Thermococcales
Methanococcales
Euryarchaeota
Heimdallarchaeota
Odinarchaeota
Thorarchaeota
Lokiarchaeota
Asgard
Korarchaeota
Thaumarchaeota
Bathyarchaeota
Verstraetearchaeota
Thermoproteales
Desulfurococcales
Fervidicoccales
Sulfolobales
TACK
Absent
Present
Distant homologs
Distant homologs in some species
B
p: 0.932
p: 0.144
p: 0.361
Generation time (min)
WT
Δsmc
Δrad50
Δsph4
C
WT
Δsmc
Δrad50
Δsph4
DIC
Hoechst 33342
Merge
D
ori-pHV4
oriC3
oriC2
rRNA
rRNA
oriC1
pHV3
pHV1
0
0.02
bin: 5 kb
Δsmc
WT
E
Chromosome position (Mb)
Δsmc
+2
log2
-2
WT
F
a b c d e f g h i
log2fc (Δsmc / WT)
Chromosome position (Mb)
G
ori-pHV4
oriC3
Chromosome position (kb)
WT
Δsmc
0.01
0
H
Chromosome position (kb)
Δsmc
+2
log2
-2
WT
bin: 5 kb
I
log2fc (Δsmc/WT)
Chromosome position (kb)
J
WT
Δ smc
Δ rad50
Δ sph4
Chromosome position (kb)
bin: 5 kb
0
0.01
Borders
Loops
K
Δsmc / WT
Δrad50 / WT
Δsph4 / WT
Chromosome position (kb)
WT
Δsmc
-2 log2 +2
WT
Δrad50
-2 log2 +2
WT
Δsph4
-2 log2 +2
L
n = 26
μ = 0.039
n = 12
μ = 0.018
n = 25
μ = 0.041
n = 26
μ = 0.032
Border score
WT
Δsmc
Δrad50
Δsph4
M
n = 64
μ = 0.331
n = 20
μ = 0.227
n = 48
μ = 0.308
n =53
μ = 0.314
Loop score
WT
Δsmc
Δrad50
Δsph4

Euryarchaeota



Optimized Hi-C
1 kb resolution



No A/B compartments

Self-interacting domains (CIDs)

DNA loops

Transcription and SMC mediated structuring