



The *Mycoplasma hyorhinis* genome displays differential chromatin accessibility

Lewis Taylor^{a,1}, Steven Walsh^{a,1}, Anna Ashton^a, Norbert Varga^a, Sejal Kapoor^a, Charlotte George^b, Aarti Jagannath^{a,*}

^a Sleep and Circadian Neuroscience Institute (SCNi), Nuffield Department of Clinical Neurosciences, University of Oxford, New Biochemistry Building, South Parks Road, Oxford, OX1 3QU, UK

^b MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, Oxford, OX3 9DU, UK

ARTICLE INFO

Keywords:
Mycoplasma
Chromatin
Accessibility
Transcription

ABSTRACT

Whilst the regulation of chromatin accessibility and its effect on gene expression have been well studied in eukaryotic species, the role of chromatin dynamics and 3D organisation in genome reduced bacteria remains poorly understood [1,2]. In this study we profiled the accessibility of the *Mycoplasma hyorhinis* genome, these data were collected fortuitously as part of an experiment where ATAC-Seq was conducted on *mycoplasma*, contaminated mammalian cells. We found a differential and highly reproducible chromatin accessibility landscape, with regions of increased accessibility corresponding to genes important for the bacteria's life cycle and infectivity. Furthermore, accessibility in general correlated with transcriptionally active genes as profiled by RNA-Seq, but peaks of high accessibility were also seen in non-coding and intergenic regions, which could contribute to the topological organisation of the genome. However, changes in transcription induced by starvation or application of the RNA polymerase inhibitor rifampicin did not themselves change the accessibility profile, which confirms that the differential accessibility is inherently a property of the genome, and not a consequence of its function. These results together show that differential chromatin accessibility is a key feature of the regulation of gene expression in bacteria.

1. Introduction

In both prokaryotic and eukaryotic organisms, the organised compaction of the genome is necessary for its packaging into a small volume, whilst allowing functional accessibility for DNA-based processes. Three-dimensional organisation of the genome in eukaryotes as a result of nucleosomal patterning around histones is key to differential gene regulation. Increased accessibility is generally seen in nucleosome-free regions, such as seen around promoters and transcription start sites, whereas chromatin inaccessibility correlates with gene repression [3,4]. The equivalent 3D organisation that allows functional gene regulation in bacteria remains poorly understood [1, 2]. In general, the bacterial genome has a hierarchical structure, comprising of a single circular chromosome which is highly condensed to form the nucleoid, although multiple exceptions have been reported [5]. Bacterial genomes are relatively gene dense, lack intronic spaces and have short intergenic regions, which contrasts with eukaryotic genomes, in which up to 99% of the genome is

* Corresponding author.

E-mail address: aarti.jagannath@ndcn.ox.ac.uk (A. Jagannath).

¹ equal contribution.

<https://doi.org/10.1016/j.heliyon.2023.e17362>

Received 3 March 2023; Received in revised form 9 June 2023; Accepted 14 June 2023

Available online 15 June 2023

2405-8440/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

non-coding. Furthermore, in bacteria, such 3D organisation may be achieved through several factors, including DNA supercoiling, molecular crowding and nucleoid-associated proteins (NAPs), which act in poorly defined and fairly heterogenous ways to shape chromatin architecture [6]. The activities of NAPs are analogous to eukaryotic histone proteins, they bind to DNA to organise it into various conformations through DNA bridging, bending, looping or wrapping architecture [6,7]. A greater understanding of bacterial chromosome organisation has been achieved using various chromosome conformation capture approaches, which have shown clear topological and subcellular domains [2,8]. However, the net result of such organisation on bacterial chromatin accessibility and its bidirectional relationship with transcriptional patterns are not clearly understood.

Mycoplasma is most often encountered in the research setting as an unwanted contamination in mammalian culture, where it has been remarkably successful. Mycoplasma rarely results in overt changes in appearance or physiology, thus estimates suggest anywhere between 10 and 85% of cell lines commonly used in the laboratory carry a mycoplasma infection that is not evident until specifically assayed for. Indeed, our study was conducted under such conditions, and thus captures the biology of mycoplasma where it is thriving.

2. Experimental model and subject details

2.1. *In vitro* cell culture studies

Human U2OS cells and murine NIH3T3 cells were obtained from ATCC (ATCC® HTB-96™ and ATCC® CRL-1658™ respectively), maintained the lab for >5 years, and cultured in complete DMEM (DMEM supplemented with 10% FCS and 1% penicillin/streptomycin) at 37 °C, 5% CO₂ as described in Taylor et al. [9]. Cells were lifted by incubation with TrypLE express for 5 min, diluted with complete DMEM and counted by trypan blue exclusion. Cells were either passaged into new tissue culture flasks, or plated into multi well plates for *in vitro* experiments. *Mycoplasma hyorhinis* contamination was noticed upon ATAC-Seq on these cells.

2.2. ATAC-Seq library preparation

100,000 cells lifted with Tryple-E and were pelleted at 400 g for 5 min. Cells were lysed using a dounce homogenizer (Sigma, D8938) in 1 ml lysis buffer (20 mM Tricine-KOH, 25 mM MgCl₂, 0.25 M sucrose, 1 mM DTT, 0.15 mM spermine, 0.5 mM spermidine, 0.3% IGEPAL-630, EDTA-free protease inhibitor) for 10 strokes. The homogenate was filtered through a 40 µm strainer into a 50 ml falcon tube, transferred to a 1.5 ml eppendorf tube, underlaid with a cushion buffer (0.5 mM MgCl₂, 0.5 mM DTT, EDTA-free protease inhibitor, 0.88 M sucrose), and centrifuged at 3000 g for 20 min at 4 °C. The pellet was washed in 500 µl cold PBS and re-pelleted at 700 g for 10 min. The pellet was resuspended in transposition reaction mix (25 µL 2× TD Buffer (Illumina Cat #FC-121-1030); 2.5 µL Tn5 Transposase (Illumina Cat #FC-121-1030); 22.5 µL Nuclease Free H₂O) and incubated for 30 min at 37 °C for transposition to occur. DNA was then purified using the Qiagen MinElute Kit and eluted with 23 µL warm 10 mM Tris pH8. The fragments were amplified, and Illumina adapters added by preparing the following mix (10 µL Transposed DNA; 10 µL Nuclease Free H₂O; 2.5 µL 25 M Customized Nextera PCR Primer 1 Barcodes given in supplementary data); 2.5 µL 25 µM Customized Nextera PCR Primer 2 (Sigma Barcode Ad1_noMX); 25 µL Fast Sybr Green 2× PCR Master Mix (Qiagen Cat #4385612) and subjected to 15–17 cycles of steps 3–5 ((1) 72 °C, 5 min; (2) 98 °C, 30 s (3) 98 °C, 10 s; (4) 63 °C, 30 s; (5) 72 °C, 1 min; Libraries quantified using the KAPA library quantification kit (KK4824), pooled and diluted to 20pM, and sequenced on the Illumina NextSeq 500 system with a 70 cycle high output cartridge.

2.3. DNA extraction and qPCR

Genomic DNA (plus *Mycoplasma* genomic DNA) was extracted from cells using the Qiagen DNeasy Blood and Tissue Mini Kit following the manufacturer's instructions. DNA concentration and quality was determined using a NanoDrop ND-1000 spectrophotometer qPCR was then conducted using 100 ng of template DNA using the Quantifast SYBR Green PCR Kit and a StepOnePlus thermal cycler (Applied biosystems) with the following thermal profile: 95 °C for 5 min and then 40 cycles of 95 °C for 10s, 60 °C for 30s and 72 °C for 12 s. Quantification of transcript levels was conducted using the relative standard curve method, for comparing within a gene, or the $2^{-\Delta\Delta Ct}$ method for comparison across genes. Primer sequences used can be found in [Supplementary Table 1](#).

2.4. RNA sequencing library preparation

Following total RNA extraction using the Qiagen RNeasy mini kit, RNA sequencing libraries were prepared using the Illumina TruSeq Stranded Total RNA library prep gold kit following the manufacturer's instructions and 150 ng of starting material as described in Taylor et al. [9]. Briefly, 150 ng of total RNA was depleted of ribosomal and mitochondrial ribosomal RNA, cleaned up using RNA clean XP beads (Beckman Coulter, High Wycombe, United Kingdom) and then fragmented. Next, first and second stand cDNA synthesis was conducted, and the resultant cDNA purified using AMPure XP beads (Beckman Coulter) and adenylated at the 3' end. Illumina indexing adapters were then ligated to the cDNA, the fragments purified and then enriched by PCR. Following amplification, the libraries were purified and then their concentration determined using the KAPA Library Quantification Kit for Illumina Platforms following the manufacturer's instructions. The libraries were then diluted to 4 nM and pooled in equal volumes prior to sequencing. Paired end RNA sequencing was then conducted using the NextSeq 550 and a Nextseq 500/500 v2 75 cycle kit, with the library loaded at 1.8 pM.

2.5. ATAC-Seq and RNA sequencing data analysis

The raw reads were initially processed to remove adapter sequences and trim low quality ends. A reference genome sequence and annotation combining human (GRCh38) with *Mycoplasma hyorhinitis* (NC_014448 from the PATRIC Bacterial and Viral Bioinformatics Resource Center) was prepared. The ATAC-Seq reads were mapped using HISAT2 and quantified with featurecounts. The RNA-Seq reads were mapped using Salmon and gene counts generated using Kallisto, normalisation and differential expression were conducted on DeSeq2 (version 1.16). Coverage was determined using BAMCoverage (deepTools2.0) with which normalised wiggle plots were generated. Standard deepTools normalisation was applied here using the BPM (number of reads per bin/sum of all reads per bin in millions) option.

3. Results and discussion

3.1. The mycoplasma genome displays a differential chromatin accessibility landscape

We performed Assay for Transpose Accessible Chromatin Sequencing (ATAC-Seq - [10]) on mouse NIH3T3 fibroblast cells as part of a separate study and discovered upon data analysis that a large number of reads did not map to the human genome (~40%). A blast search showed matches to the *Mycoplasma hyorhinitis* HUB-1 genome instead, revealing a mycoplasma infection. This presented an opportunity to investigate the genome accessibility profile in this bacterial species whilst it was contained within its host cell line. The *Mycoplasma* genus of bacteria have a parasitic lifestyle and several species are pathogenic to their animal and human hosts. They are self-replicating organisms with the smallest reduced genome (0.8 Mbp for *Mycoplasma hyorhinitis* HUB-1), which is accompanied by a considerably reduced suite of transcription factors and NAPs [11,12].

Whilst we expected to see uniform accessibility across the genome to Tn5 as part of the ATAC-Seq approach, we were surprised to see regions of markedly increased accessibility (Fig. 1A). One region that showed two to three-fold increased reads over surrounding areas contained genes important to the viral life cycle (green region – Fig. 1A), and a selection of these including ribosomal subunits, DNA polymerase III and topoisomerases are illustrated (*rplJ*, *rplL*, *rpoB*, *rpoC* – Fig. 1B; *dnaX* – Fig. 1C; *gryA* – Fig. 1D; *parC*, *parE* – Fig. 1E respectively). Another region that showed eight to ten-fold increased reads over surrounding regions contained the surface antigens – *vpA*, *D*, *E*, *F* and *G* (Fig. 1F), which are responsible for antigenic diversity for host adaptation). Our results are in agreement with some previous studies. It has been reported that the *Mycoplasma* chromosome has a defined 3D structure as in larger bacteria [13], with Hi-C (a high-throughput genomic technique to capture chromatin conformation) and super-resolution microscopy demonstrating three-dimensional chromosome organisation into 15–30 kb domains, with genes within the same domain being coregulated.

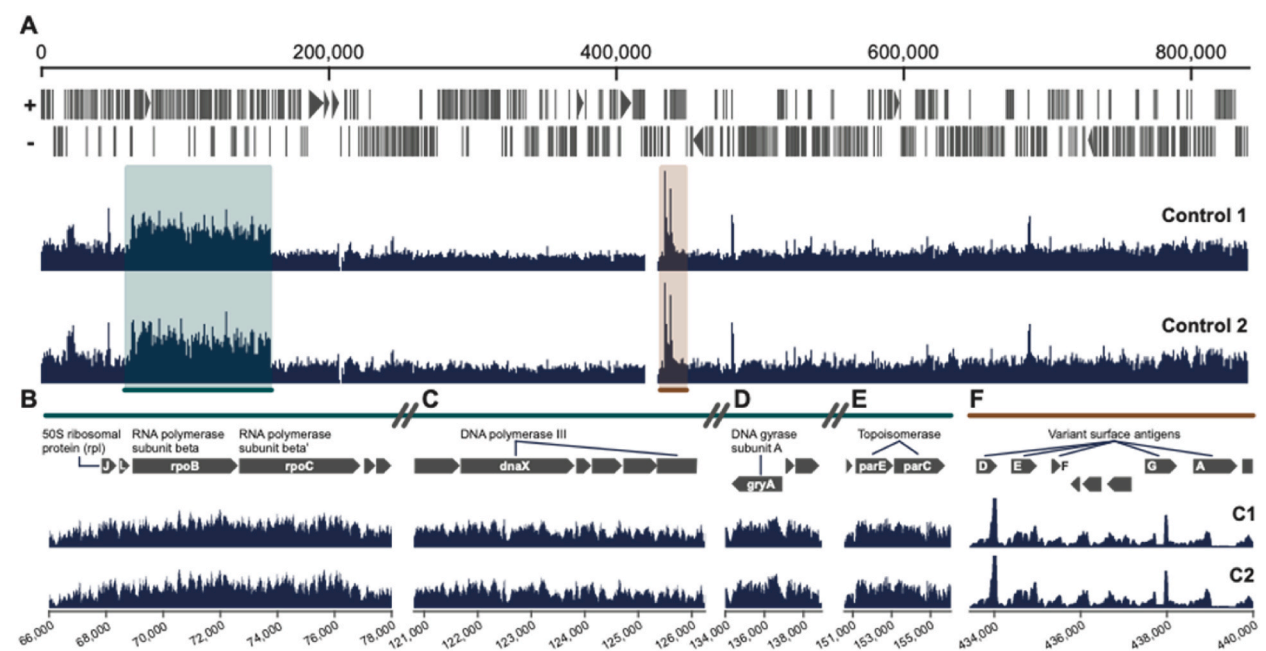


Fig. 1. The *Mycoplasma* genome differential DNA accessibility landscape. (A) ATAC-Seq data visualisation across the *Mycoplasma hyorhinitis* genome. Regions of interest containing genes essential for the bacterial life cycle highlighted in green and brown, including (B) *rplJ* and *L* – 50S ribosomal proteins, *rpoB* and *rpoC* – RNA polymerase subunits, (C) *dnaX* – DNA polymerase III, (D) *gryA* – DNA gyrase subunit A, (E) *parC* and *parE* – Topoisomerases, and (F) *vpA*, *D*, *E*, *F* and *G* – Variant surface antigens. Data are from two independent biological replicates (Control 1 and 2) and are displayed as normalised (using bedtools) ATAC counts. The Y-axis on all wiggle plots is group scaled number of reads. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Furthermore, the above study and others suggest local chromosome organisation is achieved through DNA supercoiling [14]. However, 5C analysis (similar to Hi-C) of the *Caulobacter crescentus* genome showed that repositioning elements of the chromosome such that the subcellular location of genes changed, did not alter gene expression [15]. Therefore the functional consequence of such 3D organisation remains unclear. Furthermore, *in silico* approaches suggest differential chromatin accessibility could indeed be a property of bacterial genomes that determines transcription factor occupancy [16], although other studies suggest transcription factor binding is determined by affinity for the sequence alone [17]. Our results confirm that accessibility across the *Mycoplasma* genome is indeed differential, but how this correlated with gene function and transcription were not clear.

3.2. The differential accessibility profile correlates with gene organisation

To assess whether the patterns of accessibility had any correlation with the functional organisation of the genome, we first quantified the ATAC-Seq reads at within genes versus in the intergenic regions (Fig. 2A) and found that overall, there was a close to fifteen-fold difference in the number of counts (Fig. 2B), indicating far greater accessibility within the coding regions. This is in contrast to the findings of Gomes et al., where intergenic regions were more accessible than protein-coding [16]. However, the non-coding regions of the *Mycoplasma* genome are fairly sparse, as is the case for most prokaryotes, therefore our results may be influenced by the relatively low sample size for intergenic regions. Furthermore, increased accessibility where there is protein-DNA interaction is a key feature of genome organisation. We predicted that transcription factor binding at promoters [3], transcriptional machinery at start sites and the stalling/slowing down of transcriptional machinery at the end of genes [18] would cause increased accessibility at the start and end of genes. Therefore we assessed the number of ATAC-Seq counts at the start and end of each gene, defined as 10% of the gene at each end, and the centre of the gene, which covers the remaining 80%. Counts were normalised for the length of each gene. This showed a more modest, but also statistically significant, enrichment of reads at the start/end of a gene versus its centre (Fig. 2C and D). Finally, we quantified the percentage of A or T bases in the regions with 50 or greater ATAC-Seq reads (Top),

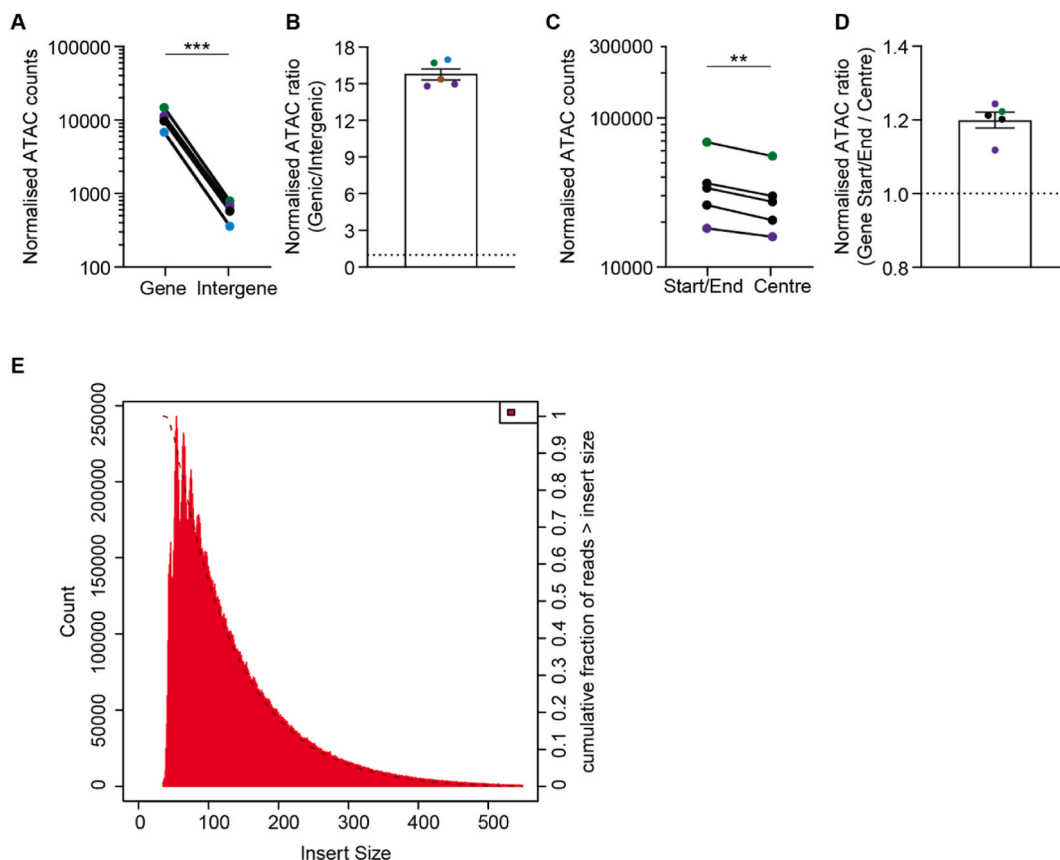
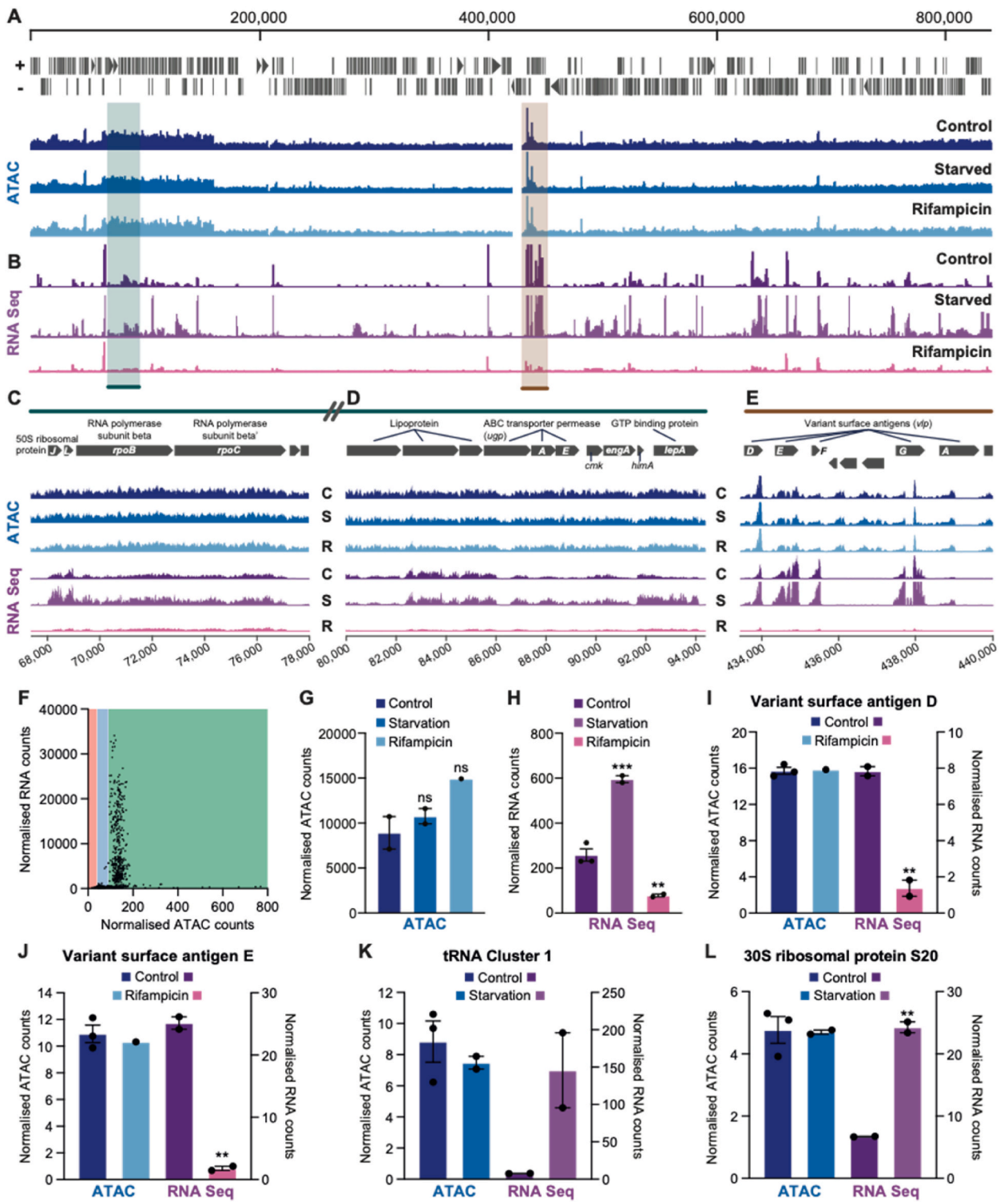


Fig. 2. The differential DNA accessibility of the *Mycoplasma* genome correlates with gene organisation. (A) Quantification of normalised ATAC-Seq counts at gene versus intergene regions across five samples and (B) the fold change of genic versus intergenic counts. (C) Quantification of normalised ATAC counts at the start or end vs the centre of a gene and (D) the fold change of gene start/end vs gene centre counts (A–D) data are mean only $n = 5$ independent samples. For (A and C) statistical analysis was conducted by a two-tailed ratio paired *t*-test. ns $P > 0.05$, ** $P < 0.01$ *** $P < 0.001$. (E) Fragment size distribution from one ATAC-Seq sample showing peaks distributed every 9–10 bases, with most fragments being 35–100 bp long. Error bars are standard error of mean.



(caption on next page)

Fig. 3. Differential DNA accessibility correlates with levels of transcription, but is not modified by transcription. *Mycoplasma* infected U2OS cells were either starved, treated with rifampicin (100 μ M, 4 h), or kept under normal conditions (Control) and ATAC and RNA sequencing conducted. Visualisation of (A) normalised ATAC counts and (B) RNA sequencing counts across the genome. ATAC and RNA sequencing counts at genes essential for the *mycoplasma* lifecycle (C and D – Green region) and infectivity (E – Brown region). Y-axis is group-scaled number of reads. (F) Correlation analysis between normalised RNA and ATAC counts showing a threshold amount of DNA accessibility was needed to allow for gene expression (green area – permissive, red and blue – not permissive). Quantification of the average ATAC signal (F) and RNA sequencing counts (G) across the whole genome and at (I) variant surface antigen D and (J) E, (K) tRNA cluster 1 and (L) 30S ribosomal protein S20. For (F–L) data are mean \pm SEM, $n = 1$ –3 independent samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

versus those with less (Rest). The Top regions were 78.78% AT (standard deviation 9.027), whilst the Rest were 71.99% AT (standard deviation 8.08) indicating the regions showing greater accessibility may have fractionally greater AT content, although this is not statistically significant. Finally, we measured read fragment lengths and found a large spread from 35 to 500 bases (Fig. 2E), with peaks occurring 10 bases apart. Histone like proteins, the most conserved and abundant of which is HU [1] have been described in many bacterial species including *Mycoplasma* [19], with the *Escherichia coli* HU protein appearing to bind DNA as dimer at 9 base-pair intervals apart regardless of sequence [20]. Therefore, it is tempting to speculate that this is similar to nucleosomal patterning seen in ATAC-Seq data from eukaryotes, where fragments occur in peaks approximately 200 bases apart. Together, these results indicate the differential accessibility pattern may correlate with the functional organisation of the *Mycoplasma* genome.

We then assessed a number of possibilities that the data observed were artefacts of the model or experimental technique. A key feature of the genes that code for surface antigens is a variable number of repeats contained within the sequence. As such, it is possible that the increased reads we found in these regions corresponded to multiple repeats in the reads mapping to a single repeat in the reference genome. This was not apparent in the mapped data, where we saw unique reads across the different repeats (Supplementary Fig. 1). We next addressed whether copy number variation (whether multiple copies of the same gene were found in the genome) could underlie the difference in ATAC signal across the *Mycoplasma* genome. We designed PCR primers to cover discrete genomic regions (Supplementary Fig. 2A) in three genes of interest – *ushA*, *rpoB* and *vlpD*. A quantification of ATAC-Seq reads within these regions is indicated in Supplementary Fig. 2B. The corresponding signal from PCR product normalised to host (human) housekeeping genes is shown in Supplementary Fig. 2C, and these were not found to correspond to one another. Furthermore, the distribution of ATAC-Seq counts over the genes shows a large spread for some genes, but not others (Supplementary Fig. 2D).

3.3. Differential DNA accessibility correlates with levels of transcription, but is not modified by transcription

We next sought to understand if the differential accessibility had any correlation with transcription. We performed RNA-Seq and ATAC-Seq on cells infected with *Mycoplasma* and mapped the *Mycoplasma* reads. We found overall that regions of higher accessibility did map to genes and also to regions of increased transcription (Fig. 3A and B). For example, the regions containing the RNA polymerase beta subunit genes (*rpoB* and *rpoC*) showed increased accessibility over surrounding regions, accompanied by increased transcription (Fig. 3C), as is also seen for a range of other genes including the variant surface antigens shown in Fig. 3D and E. However, there are clearly specific regions of where accessibility and levels of transcription do not correlate. This is illustrated (Supplementary Fig. 3) with the genes NADH oxidase and transposase IS861, where NADH oxidase appears more accessible, but it transcribed less. Furthermore, we observed high accessibility and transcription in a few regions where no genes were annotated (Supplementary Fig. 3). Interestingly, such findings have also been reported when profiling protein occupancy across the *E. coli* chromosome, where extensive occupancy was found at non-coding and transcriptionally silent areas, and these have been postulated to be organising centres that define topological domains in the bacterial genome [21].

Transcriptional activity shapes DNA organisation in many ways; active RNA synthesis induces nucleoid compaction and supercoiling, with negative supercoiling occurring upstream, and positive supercoiling occurring downstream of the transcription complex [22,23]. Additionally, the clustering of highly expressed genes into transcription foci defines chromosomal domain boundaries [24,25]. The relationship between DNA organisation and gene expression is actually bidirectional, as chromosomal structure dynamics can regulate transcription through NAP-mediated changes in DNA topology and supercoiling, leading to changes in RNA polymerase and transcription factor binding [26]. Furthermore, 3D genome structure in bacteria is dynamic and influenced by growth conditions and stress [27,28]. Cagliero et al. used high-resolution chromosome conformation capture-based methods to analyse spatial organization of the *Escherichia coli* nucleoid and found dynamic 3D organisation, with replication causing greater interactions in the origin domains, and amino acid starvation causing spatial clustering of downregulated genes [28]. In order to assess whether changes in transcriptional activity were driving changes in accessibility in the *Mycoplasma* genome, we performed various treatments that would alter levels of transcription. These treatments involved starving the host cells by incubating in PBS or halting transcription by incubation with 100 μ M rifampicin for 4 h each. Whilst these treatments clearly changed levels of transcription across the genome (Fig. 3B), no alterations in accessibility were seen (Fig. 3A).

An analysis of ATAC-Seq counts per base for variant surface antigen D/E (*vlpD/E*), the tRNA cluster and the 30S ribosomal protein S20 (*rpsT*) are illustrated in Fig. 3I/3J, 3K and 3L respectively, showing no change in accessibility in response to either starvation or rifampicin treatment, but a clear change in transcriptional responses, with starvation increasing transcription and rifampicin reducing it. Interestingly, some genes showed persistent expression even after rifampicin treatment (Fig. S3) and this may be due to longer degradation times for some transcripts. Several studies have tried to establish a possible role for transcription in chromosome organization using chemical inhibition of the initiation or elongation of RNA polymerase II (RNAPII) [29], and as detailed below. Some

drugs used to inhibit transcription also affect RNAPII levels at the promoter, whereas others do not [29]. Therefore, when interpreting results from this type of experiment, it is important to consider whether only transcription has been affected or whether the presence of proteins in the transcription complex — which may be responsible for mediating interactions contact domains — has been affected too. In the bacteria *Caulobacter crescentus* (also known as *Caulobacter vibroides*) and *Bacillus subtilis*, inhibition of transcription using rifampicin results in a loss of contact domains [8,30]. Interestingly, conflicting results regarding the effect of rifampicin on nucleoid structure - both nucleoid expansion and nucleoid compaction - have been reported [22,31], likely due to differences in experimental conditions, analyses, and/or strains used. In our study, we saw no changes in accessibility profiles after treatment with rifampicin, but a clear drop in transcription.

We then we measured the correlation between the ATAC-Seq counts and corresponding RNA-Seq counts at the gene level (Fig. 3F), summarised across the genome in Fig. 3G and H. We noticed that there was little to no transcription where there was low accessibility (red and blue regions), confirming that accessibility was indeed required for a gene to be transcribed. However, once a threshold level of accessibility was reached (green region), levels of transcription did not correlate with levels of accessibility. This is in agreement with what is known about regulation of gene expression in eukaryotic systems, where accessibility at the promoter indicates “priming” or readiness for transcription, not necessarily levels of transcription. Together, these results indicate that the accessibility of chromatin in *Mycoplasma hyorhina* is a requirement for gene expression, but is itself not modified by transcription.

A limitation of our study is that the conditions under which our experiments were performed were tailored to the culture of mammalian cells, and therefore may have been suboptimal for *Mycoplasma*. Indeed, the ATAC-Seq nuclei isolation and tagmentation protocol may not be optimal for capturing protein-DNA interactions in *Mycoplasma*. However, given the success of ATAC-Seq protocols in reliably capturing accessibility within mitochondrial DNA – indeed mitochondrial mapping reads are treated as an unavoidable experimental nuisance that have now been mined to provide useful accessibility data [32,33] - it is likely that the same applies to *Mycoplasma*. Furthermore, the extent of mycoplasma contamination across labs worldwide is a clear indicator that are conditions under which *Mycoplasma* thrive and are very successful, thus our observations are likely to underpin *Mycoplasma* gene regulation under more stringent conditions also.

In summary our data show that differential chromatin accessibility is a key feature of the regulation of gene expression in *Mycoplasma*. In general, it appears where there is no accessibility, there is no transcription, with genes that are key to the bacterial life cycle being the most highly accessible and transcriptionally active. Interestingly transcription itself does not alter accessibility, showing that it is not the dynamics of DNA coiling around RNA pol II that are responsible for the patterns seen. However, there are clear regions which are transcriptionally silent but highly accessible, suggesting this may be part of the mechanism by which complex topological organisation can be achieved [2,21].

Author contribution statement

Lewis Taylor, Aarti Jagannath: Conceived and designed the experiments; Performed the experiments; Analysed and interpreted the data; Wrote the paper.

Steven Walsh: Conceived and designed the experiments; Performed the experiments; Analysed and interpreted the data.

Anna Ashton: Analysed and interpreted the data; Wrote the paper.

Norbert Varga, Sejal Kapoor: Performed the experiments.

Charlotte George: Analysed and interpreted the data.

Funding statement

Dr Aarti Jagannath was supported by Biotechnology and Biological Sciences Research Council {BB/N01992X/1}.

Data availability statement

Data are available at <http://www.ncbi.nlm.nih.gov/bioproject/984476>.

Resource availability

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by Aarti Jagannath (aarti.jagannath@ndcn.ox.ac.uk)

Materials availability

This study did not generate new unique reagents.

Data and code availability

The datasets generated during this study will be deposited to the appropriate public repositories upon acceptance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the following sources of funding: BB/N01992X/1 David Phillips fellowship from the BBSRC to AJ, Elysium-Oxford Post-Doctoral Fellowship to LT. We would like to thank Prof. David Sherratt for his insightful comments and guidance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.heliyon.2023.e17362>.

References

- [1] B.A. Shen, R. Landick, Transcription of bacterial chromatin, *J. Mol. Biol.* 431 (2019) 4040–4066, <https://doi.org/10.1016/j.jmb.2019.05.041>.
- [2] M. Trussart, E. Yus, S. Martinez, D. Baù, Y.O. Tahara, T. Pengo, M. Widjaja, S. Kretschmer, J. Swoger, S. Djordjevic, et al., Defined chromosome structure in the genome-reduced bacterium *Mycoplasma pneumoniae*, *Nat. Commun.* 8 (2017), <https://doi.org/10.1038/ncomms14665>.
- [3] S. Venkatesh, J.L. Workman, Histone exchange, chromatin structure and the regulation of transcription, *Nat. Rev. Mol. Cell Biol.* 16 (2015), <https://doi.org/10.1038/nrm3941>.
- [4] K. Maeshima, S. Tamura, J.C. Hansen, Y. Itoh, Fluid-like chromatin: toward understanding the real chromatin organization present in the cell, *Curr. Opin. Cell Biol.* 64 (2020), <https://doi.org/10.1016/j.ceb.2020.02.016>.
- [5] G.C. diCenzo, T.M. Finan, The divided bacterial genome: structure, function, and evolution, *Microbiol. Mol. Biol. Rev.* 81 (2017), <https://doi.org/10.1128/mmr.00019-17>.
- [6] M.J. Rowley, V.G. Corces, Organizational principles of 3D genome architecture, *Nat. Rev. Genet.* 19 (2018), <https://doi.org/10.1038/s41576-018-0060-8>.
- [7] S.C. Dillon, C.J. Dorman, Bacterial nucleoid-associated proteins, nucleoid structure and gene expression, *Nat. Rev. Microbiol.* (2010), <https://doi.org/10.1038/nrmicro2261>.
- [8] T.B.K. Le, M. v Imakaev, L.A. Mirny, M.T. Laub, High-resolution mapping of the spatial organization of a bacterial chromosome, *Science* 342 (2013) 731–734, <https://doi.org/10.1126/science.1242059>.
- [9] L. Taylor, T. Palumaa, P. Reardon, S. Walsh, B. Johnson, S. Liberatori, S. Hasan, K. Clark, P. Cohen, S. Mohammed, et al., Light-regulated SIK1 Remodels the Synaptic Phosphoproteome to Induce Sleep, 2021. [BIORXIV/2021/462159](https://doi.org/10.1101/2021.04.21.462159).
- [10] J.D. Buenostro, P.G. Giresi, L.C. Zaba, H.Y. Chang, W.J. Greenleaf, Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position, *Nat. Methods* 10 (2013), <https://doi.org/10.1038/nmeth.2688>.
- [11] S. Torres-Puig, A. Broto, E. Querol, J. Piñol, O.Q. Pich, A novel sigma factor reveals a unique regulon controlling cell-specific recombination in *Mycoplasma genitalium*, *Nucleic Acids Res.* 43 (2015) 4923–4936, <https://doi.org/10.1093/nar/gkv422>.
- [12] R. Herrmann, B. Reiner, *Mycoplasma pneumoniae* and *Mycoplasma genitalium*: a comparison of two closely related bacterial species, *Curr. Opin. Microbiol.* 1 (1998) 572–579, [https://doi.org/10.1016/S1369-5274\(98\)80091-X](https://doi.org/10.1016/S1369-5274(98)80091-X).
- [13] M. Trussart, E. Yus, S. Martinez, D. Baù, Y.O. Tahara, T. Pengo, M. Widjaja, S. Kretschmer, J. Swoger, S. Djordjevic, et al., Defined chromosome structure in the genome-reduced bacterium *Mycoplasma pneumoniae*, *Nat. Commun.* 8 (2017), <https://doi.org/10.1038/ncomms14665>.
- [14] W. Zhang, J.B. Baseman, Transcriptional regulation of MG-149, an osmoinducible lipoprotein gene from *Mycoplasma genitalium*, *Mol. Microbiol.* 81 (2011) 327–339, <https://doi.org/10.1111/j.1365-2958.2011.07717.x>.
- [15] M.A. Umbarger, E. Toro, M.A. Wright, G.J. Porreca, D. Baù, S.H. Hong, M.J. Fero, L.J. Zhu, M.A. Marti-Renom, H.H. McAdams, et al., The three-dimensional architecture of a bacterial genome and its alteration by genetic perturbation, *Mol. Cell.* 44 (2011) 252–264, <https://doi.org/10.1016/j.molcel.2011.09.010>.
- [16] A.L.C. Gomes, H.H. Wang, T.M. Przytycka, The role of genome accessibility in transcription factor binding in bacteria, *PLoS Comput. Biol.* 12 (2016), 1004891, <https://doi.org/10.1371/journal.pcbi.1004891>.
- [17] J.T. Wade, N.B. Reppas, G.M. Church, K. Struhl, Genomic analysis of LexA binding reveals the permissive nature of the *Escherichia coli* genome and identifies unconventional target sites, *Genes Dev.* 19 (2005) 2619–2630, <https://doi.org/10.1101/gad.1355605>.
- [18] J.B. Rodríguez-Molina, S. West, L.A. Passmore, Knowing when to stop: transcription termination on protein-coding genes by eukaryotic RNAPII, *Mol. Cell.* 83 (2023), <https://doi.org/10.1016/j.molcel.2022.12.021>.
- [19] D. Kamashev, J. Oberto, M. Serebryakova, A. Gorbachev, Y. Zhukova, S. Levitskii, A.K. Mazur, V. Govorun, *Mycoplasma gallisepticum* produces a histone-like protein that recognizes base mismatches in DNA, *Biochemistry* 50 (2011) 8692–8702, <https://doi.org/10.1021/bi2009097>.
- [20] D. Kamashev, The histone-like protein HU binds specifically to DNA recombination and repair intermediates, *EMBO J.* 19 (2000) 6527–6535, <https://doi.org/10.1093/emboj/19.23.6527>.
- [21] T. Vora, A.K. Hottes, S. Tavazoie, Protein occupancy landscape of a bacterial genome, *Mol. Cell.* 35 (2009) 247–253, <https://doi.org/10.1016/j.molcel.2009.06.035>.
- [22] J.E. Cabrera, C. Cagliero, S. Quan, C.L. Squires, J.J. Ding, Active transcription of rRNA operons condenses the nucleoid in *Escherichia coli*: examining the effect of transcription on nucleoid structure in the absence of transcription, *J. Bacteriol.* 191 (2009) 4180–4185, <https://doi.org/10.1128/JB.01707-08>.
- [23] S. Chong, C. Chen, H. Ge, X.S. Xie, Mechanism of transcriptional bursting in bacteria, *Cell* 158 (2014) 314–326, <https://doi.org/10.1016/j.cell.2014.05.038>.
- [24] T.B.K. Le, M.V. Imakaev, L.A. Mirny, M.T. Laub, High-resolution mapping of the spatial organization of a bacterial chromosome, 2013. PMID: 24158908.
- [25] I. Junier, O. Martin, F. Képès, Spatial and topological organization of DNA chains induced by gene co-localization, *PLoS Comput. Biol.* 6 (2010), <https://doi.org/10.1371/journal.pcbi.1000678>.
- [26] S.C. Verma, Z. Qian, S.L. Adhya, Architecture of the *Escherichia coli* Nucleoid, 2019, <https://doi.org/10.1371/journal.pgen.1008456>.
- [27] J.E. Cabrera, C. Cagliero, S. Quan, C.L. Squires, J.J. Ding, Active transcription of rRNA operons condenses the nucleoid in *Escherichia coli*: examining the effect of transcription on nucleoid structure in the absence of transcription, *J. Bacteriol.* 191 (2009) 4180–4185, <https://doi.org/10.1128/JB.01707-08>.
- [28] C. Cagliero, R.S. Grand, M.B. Jones, D.J. Jin, J.M. O'sullivan, Genome conformation capture reveals that the *Escherichia coli* chromosome is organized by replication and transcription, 2013. PMID: 23632166.
- [29] H. Mosaei, N. Zenkin, Inhibition of RNA polymerase by rifampicin and rifamycin-like molecules, *EcoSal Plus* 9 (2020), <https://doi.org/10.1128/ecosalplus.esp-0017-2019>.
- [30] X. Wang, H.B. Brandão, T.B.K. Le, M.T. Laub, D.Z. Rudner, *Bacillus subtilis* SMC complexes juxtapose chromosome arms as they travel from origin to terminus, *Science* (1979) 355 (2017) 524–527, <https://doi.org/10.1126/science.aai8982>.

- [31] E.W. Harrington, N.J. Trun, Unfolding of the bacterial nucleoid both in vivo and in vitro as a result of exposure to camphor, *J. Bacteriol.* 179 (1997) 2435–2439, <https://doi.org/10.1128/jb.179.7.2435-2439.1997>.
- [32] C.A. Lareau, V. Liu, C. Muus, S.D. Praktijnjo, L. Nitsch, P. Kautz, K. Sandor, Y. Yin, J.C. Gutierrez, K. Pelka, et al., Mitochondrial single-cell ATAC-seq for high-throughput multi-omic detection of mitochondrial genotypes and chromatin accessibility, *Nat. Protoc.* (2023), <https://doi.org/10.1038/s41596-022-00795-3>.
- [33] L.S. Ludwig, C.A. Lareau, J.C. Ulirsch, E. Christian, C. Muus, L.H. Li, K. Pelka, W. Ge, Y. Oren, A. Brack, et al., Lineage tracing in humans enabled by mitochondrial mutations and single-cell genomics, *Cell* 176 (2019), <https://doi.org/10.1016/j.cell.2019.01.022>.