

Chromatin reorganization, signaling responses and non-coding RNA regulators in homeostasis and senescence

Dissertation

for the award of the degree
Doctor rerum naturalium (Dr. rer. nat)
of the Georg-August-Universität Göttingen

within the doctoral program
International Max Planck Research School for Genome Science
of the Georg-August University School of Science (GAUSS)

submitted by

Nataša Josipović

Göttingen, June 2021

Thesis Advisory Committee

Prof. Argyris Papantonis (Supervisor)

Institute of Pathology, University Medical Center Göttingen, 37075 Göttingen, Germany

Prof. Johannes Söding

Quantitative and Computational Biology, Max-Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

Prof. Markus T Bohnsack

Department of Molecular Biology, University Medical Center Göttingen & Göttingen Centre for Molecular Biosciences, Georg-August-University, 37077 Göttingen, Germany

Members of Examination Board

1st referee: **Prof Argyris Papantonis** (Supervisor)

Institute of Pathology, University Medical Center Göttingen, 37075 Göttingen, Germany

2nd referee: **Prof. Johannes Söding**

Quantitative and Computational Biology, Max-Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

Further members of Examination Board

Prof. Patrick Cramer

Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

Dr. Marieke Oudelaar

Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany

Dr. Ufuk Günesdogan

Department of Developmental Biology, Göttingen Center for Molecular Biosciences, University of Göttingen, Göttingen, Germany; Department of Molecular Developmental Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Date of oral examination: September 15th, 2021

Acknowledgements

Reflecting on this long journey made me realize that no words can do justice to the overwhelming happiness and appreciation I feel for being a part of such an amazing group. Our paths might have crossed because of the same journey we chose, but the friendships we developed I will cherish wherever the future takes us.

Boss, I chose to address you like this throughout the years, because you have brought to this word the meaning, I believe it should carry. From our first interview, your curiosity and excitement about the science made me wish to be a part of it. Through all the good, bad, and random times, you found a moment to offer a pep talk, kind word, advice, tissue or even a drink. You respected and nurtured our individualities and had a way to bring out the best from each of us. I am lucky to have had your full attention and guidance when we were smaller group, and knowing that Papantonis lab will soon become a whole department, I hope your new kids will get to be fortunate to have a full Boss experience.

Since this journey started in Cologne, I would like to thank you Miloš, first for getting a PhD in Papantonis lab himself, because then I learned about it and decided to try and skew the serbian-to-greek ratio in the group. Second, because you mentored me and helped develop my first bioinformatic skills. Having you as my family in a new city and new country have eased the transition in so many ways. Theodore, I remember our days in Cologne with a great laugh, and although our relationship was challenging at times, it would not be the same without you there.

Kostas, we were on this journey together since the beginning of mine, and over the years you became my best friend. With you always there to cheer me up, listen and discuss, I never felt alone. We pushed each other to be better, and despite me sometimes being tough on you, you were always trying to be considerate of my feelings and patient with your words. I thank you for that and many, many, many great nights out in Göttingen. Of course, these nights would be not as fun without our Spiro. Spirić, you know you have special place in my heart, and that I consider you a lifelong friend. Nadine, you started with us in rebellion against injustice and stayed a warrior every single day after. But your sensitive side and care for others is what I appreciate the most and I am happy to have met you. In the spirit of strong girls of our lab, Adi, you were the most surprising friendship I have made. Overwhelming at times, but mostly adorable and a person with a big heart, whose determination I am in awe of. The rest of you guys, forgive me for not pointing each one, but know that I cherish I met all of you and I thank you for all the memories, the family spirit in the lab and care for each other. I could not think of a better bunch of people.

Finally, I would like to dedicate few words to my family. Mama, od učenja prvih slova, preko pisanja sastava, maturiranja, polaganja ispita i diplomiranja, ovo je bio tvoj trud koliko i moj. Konačno smo došle i do kraja našeg školovanja, a početka novih poglavlja. Dudane, hvala što si uvek tu da se nasmejemo, kukamo i čavljamo. Ovih par godina su nas možda razdvojile fizički, ali smo, osećam, bliži nego ikad. Mama, Dušane, tata, ovu disertaciju posvećujem vama, jer bez vas ne bi bilo ni nje.

Table of Contents

Acknowledgements.....	3
Summary	6
Introduction	7
1. Interplay between chromatin structure and transcription.....	8
1.1 Chromatin structure conveys the function	8
1.2 Mechanisms of chromatin interactions	14
1.3 Role of transcription in shaping the genome.....	17
2. Senescence as a model for studying genome organization principles	22
2.1 HMG family of architectural proteins	27
2.2 About HMGB1	29
3. Regulatory non-coding RNAs	34
3.1 circRNAs are versatile class of lncRNAs	36
3.2 circRNA biogenesis.....	37
3.3 circRNAs in homeostasis and disease	41
Chapters.....	44
Chapter 1 – RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis.....	44
Abstract.....	45
Introduction	46
Results.....	47
Discussion.....	59
Materials and methods.....	60
Appendix – Supplementary figures 1-6.....	69
Chapter 2 - HMGB1 coordinates SASP-related chromatin folding and RNA homeostasis on the path to senescence.....	79
Abstract.....	81
Introduction	82
Results.....	83
Discussion.....	98
Materials and methods.....	100
Appendix – Supplementary figures 1-6.....	109

Chapter 3 - circRNAs modulate cell proliferation and the immediate-early response by reorganizing gene expression and mRNA processing 119

- Abstract..... 121
- Introduction 122
- Results..... 122
- Discussion..... 134
- Materials and methods..... 135
- Appendix – Supplementary Figures 1-4 142
- Discussion..... 147
- Outlook 154
- Literature 156
- Curriculum Vitae* 189

Summary

In mammalian cells, a relatively limited fraction of genetic information is utilized in a variety of ways to give rise to distinct cellular phenotypes. Thus, the genotype essentially acts as input for multilayered regulatory networks, which coordinate cellular activities and ultimately shape cell identity and function. Early studies, focusing on individual components of such regulatory networks, have unraveled core principles of cellular functions. Therein, individual, or small groups of genes are regulated by a defined set of locally acting factors. The idea of such individual regulatory units persists, despite new experimental and computational evidence pointing to extensively intertwined “pan-genomic” regulatory networks. In such emerging models, intricate cellular traits originate from numerous interactions, and the mutual dependencies between distinct regulatory units are further governed by such factors as 3D chromatin architecture, non-coding RNA, and even regulatory crosstalk between cells. Therefore, understanding the mechanisms of accurate spatiotemporal regulation of gene expression requires not only elucidating the roles of individual factors at local nodes, but also their contribution to a collective pan-genomic network.

In light of such a view, I combined genomics and functional studies in primary human cells to study the regulatory modes of three diverse molecular factors in the context of cellular homeostasis. In **Chapter 1**, I examine the contribution of active RNA polymerase II (RNA Pol II) in the intricate process of setting up chromatin organization following exit from mitosis. In **Chapter 2**, I examine the role of HMGB1 as an architectural and RNA-binding factor mediating key events in cells exiting homeostatic cell cycle progression to enter replicative senescence. In **Chapter 3**, I explore the regulatory potential of two circRNAs, circCAMSAP1 and circRAB3IP, in navigating the regulation of cell cycle progression and inflammation in human cells. Taken together, these three layers of genomic regulation – via RNA Pol II transcription, a dual-function transcription factor, and a pair of circular non-coding RNAs – exemplify the idea of multi-layered regulation of pan-genomic networks inside human cells.

Introduction

Establishment and maintenance of cellular homeostasis is founded in the cells' ability to register and adapt to various environmental and intrinsic cues. Throughout the processes of division, differentiation, signaling, aging, as well as in disease, cells undergo morphological and functional changes which ultimately represent the read-out of a specific gene expression program. Therefore, it is of great importance to understand how transcription is regulated to ensure the correct spatiotemporal execution of expression programs.

Such coordination of gene expression upon activation of signaling cascades is achieved through multi-faceted regulatory mechanisms. Local gene regulation operating on linear chromatin fiber, involves cooperative action of *cis*-acting regulatory elements (1D) and is further embedded in larger, three-dimensional (3D) networks, encompassing various *trans*-acting factors. *Cis*-acting gene regulatory elements, such as gene promoters, enhancers, transcription-factors, and histone modifications, finetune gene regulation through the establishment of local permissive or repressive transcriptional environments, often acting upon small groups of genes. Multiple gene groups pertaining a joined regulatory program are further organized into larger transcriptionally active or inactive environments in nucleus and placed under shared regulatory mechanisms. Regulatory elements often reside at great distances in nucleus, implying the necessity of dynamic chromatin movement in nuclear space. Traditionally viewed as a somewhat static 1D transmitter of genetic information, chromatin is now referred to as highly dynamic and multi-dimensional platform for the orchestration of essential cellular processes. Multifaceted regulation and compartmentalization of transcriptional programs, all the while confined within a micron-sized nucleus, implies an intimate and dynamic relationship between chromatin function and its structure.

Therefore, a further understanding of complex and intertwined gene regulatory networks, which emerge from coordinated activity of tens of thousands of molecules and regulatory sequences, which are maintained and modified in strictly regulated temporal manner, calls for careful consideration of all its contributing factors.

1. Interplay between chromatin structure and transcription

1.1 Chromatin structure conveys the function

Electron microscopy studies coupled with biochemical approaches, more than half a century ago, revealed a periodical association of DNA with histone proteins in a structure we know today as nucleosome (Oudet *et al*, 1975). In human, about 30 million nucleosome octamers, in most cases composed of two sets of H2A, H2B, H3 and H4 histone proteins (Luger *et al*, 1997), are spaced by nucleosome-free linker DNA and stabilized by the H1 histone at the nucleosome entry/exit. The role of histone proteins extends beyond structural confinement of ~2-meter-long DNA into micron scale nuclei and into the functional chromatin regulation. N-terminal tail of histone proteins can be post-translationally modified, often by acetylation and methylation, which in turn affects their positioning and affinity for DNA binding. Compacted DNA, bound by histones, is largely inaccessible for interactions with other protein complexes (Oudet *et al*, 1975). Conversely, nucleosome depleted genomic regions are available for interaction with replication- and transcription-related molecular machineries. Due to their role in governing local chromatin accessibility, nucleosome arrays are highly dynamic and subject to coordinated regulation. Through interaction with various ATP-dependent chromatin remodelers and histone chaperons, chromatin and histone proteins are substrates for remodeling and repositioning along the chromatin fiber (Lai & Pugh, 2017; Lawrence *et al*, 2016).

At the largest structural scale, interphase chromatin is non-randomly distributed in nuclei; In fact, interphase chromosomes occupy distinct nuclear territories (Cremer & Cremer, 2001) and are generally organized into two large compartments named euchromatin and heterochromatin (Passarge, 1979; van Steensel & Furlong, 2019; Hildebrand & Dekker, 2020). This finding is supported by principal component analysis (PCA) of population and single-cell Hi-C data (Lieberman-Aiden *et al*, 2009; Stevens *et al*, 2017). Such analysis revealed two types of Mb-sized compartments named “A” and “B”, which largely overlap and share the properties of euchromatin and heterochromatin, respectively. Euchromatin is generally less condensed, rich in GC content and histone modifications indicative of active transcription, such as H3K4me3, H3K27ac, H4K8ac and H4K16ac (Lawrence *et al*, 2016). Genomic regions embedded in euchromatin tend to be positioned centrally in the nucleus, apart from nuclear pore regions (Fišerová *et al*,

2017) located at the nuclear membrane. In contrast, heterochromatin is more compacted, thereby less accessible, largely transcriptionally inactive and decorated with H3K9me3 (facultative heterochromatin) and H3K27me3 (constitutive heterochromatin) histone marks (Penagos-Puig & Furlan-Magaril, 2020). Heterochromatin is positioned around and tethered to the nuclear envelope (lamina-associated domains - LADs) and nucleoli (nucleolus-associated domains - NADs). Another prominent feature of chromatin compartmentalization is a strong preference for association between compartments of same type and essentially no contacts between different compartments. Notable single-gene loci exemptions involve activity of TFs (Yin Yang 1 or YY1; Harr *et al*, 2015) and chromatin modifiers (CCCTC-binding factor or CTCF; Yusufzai *et al*, 2004) that results in compartment switching. Association of compartments with a particular transcriptional status implies a functional relationship between the two. Growing body of evidence indicates that transcription is a contributing but not deterministic factor in compartmentalization, and in turn, compartments define large permissive/repressive transcriptional environments that are further regulated by the means of local transcription regulation (van Steensel & Belmont, 2017; Therizols *et al*, 2014).

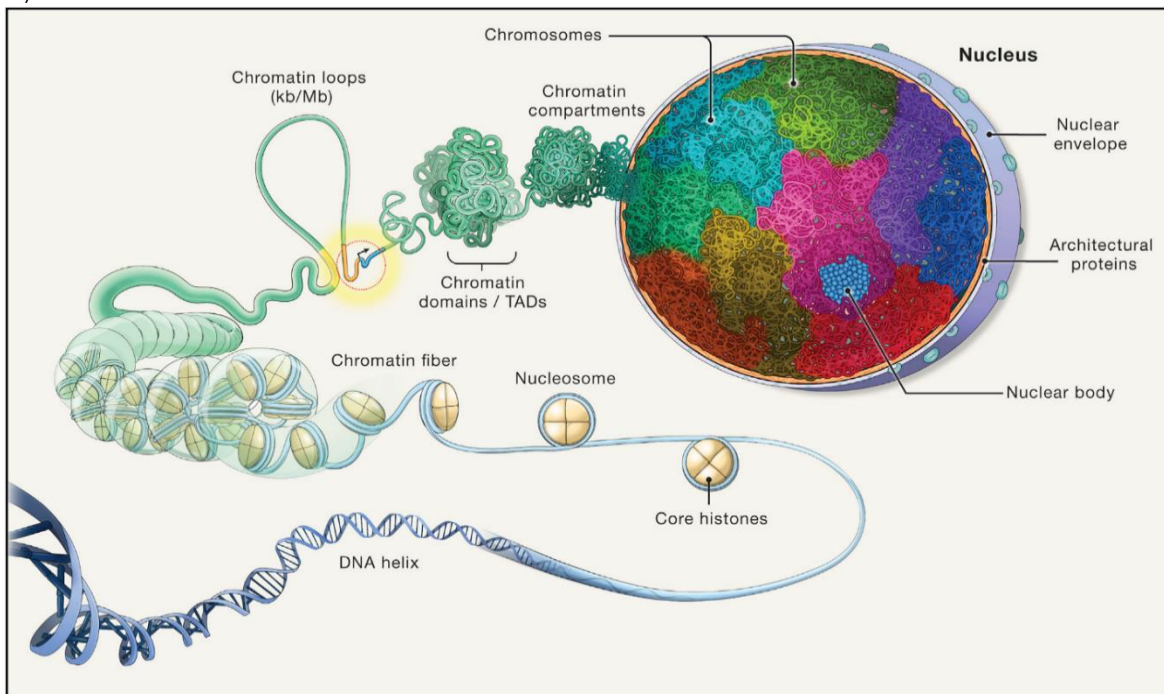


Figure 1. Multiple layers of chromatin folding in eukaryotic nucleus.

DNA is wrapped around core histone octamer to form nucleosomes, which further organize into chromatin fiber. Interactions of chromatin fiber in 3-dimensional space are represented by chromatin loops, which give rise to larger chromatin domains and compartments and eventually chromosomes. Reprinted from (Misteli T., 2020)

Chromatin compaction beyond the linear DNA fiber and below the chromosomal compartments is currently an area of very active research. Folding of chromatin in 3D space has to be non-random due to the topological constraints of nucleus, while also being flexible enough to accommodate very dynamic essential cellular processes, such as transcription or replication. Formation of higher order chromatin structures can be, at least in part, explained by polymer-like properties of chromatin fiber. Genomic regions found apart in linear chromatin randomly collide in the nuclei, where the probability of two regions colliding, or looping, drops as their distance on linear DNA increases (**Fig. 1**; Bouwman & de Laat, 2015). Such genomic interactions in cells are additionally modulated by plethora of regulatory and/or architectural proteins, enabling contacts at larger distances, restricting contacts, or maintaining existing ones. This principle was exploited for development of proximity-ligation based approaches that shed light on chromatin folding and revealed previously undetected layers of chromatin hierarchy. Ligation-based techniques, collectively named chromosome conformation capture-based methods (Übelmesser & Papantonis, 2019), proved particularly useful in probing targeted or genome-wide contacts between genomic regions (Dekker *et al*, 2002; Simonis *et al*, 2006; Dostie *et al*, 2006; Lieberman-Aiden *et al*, 2009; Hsieh *et al*, 2015; method variations reviewed in de Wit & de Laat, 2012; Goel & Hansen, 2021). Earlier mentioned Hi-C is a genome-wide derivative of the chromatin conformation capture based methods, relying on proximity-ligation of distal genomic regions whose interactions are mediated by proteins. High throughput sequencing of such ligated chimeric fragments results in a genome-wide interaction frequency map (**Figure 2**). Square or triangle contact maps reflect contact frequency of interacting genomic regions in a checkboard manner, where the color intensity of contacts is proportional to their frequency. Since 2009, Hi-C has gained a momentum as a technique of choice to study principles of chromatin folding, supplemented with other sequencing-based techniques, such as chromatin immunoprecipitation (ChIP-seq), Assay for Transposase-Accessible Chromatin (ATAC-seq) and their derivatives, as well as super-resolution microscopy and gene editing. ChIP-seq is widely applicable method for investigation of DNA-protein interactions, based on crosslinking and immunoprecipitation of protein of interest and associated chromatin, which is then subjected to high-throughput sequencing (Park, 2009). ATAC-seq is a popular method for investigating accessible chromatin by exploiting the general availability of such regions to the activity of bacterial Tn5-transposase (Buenrostro *et al*, 2015). With the increasing research efforts over the last two decades our understanding of complex interplay between chromatin structure and its function is continuously updated. This has inevitably led to some discrepancy in interpretation and naming conventions for higher order chromatin structures however, they all converge in the idea of genome

structured in hierarchical, nested layers of loops, domains, and chromatin compartments (Schoenfelder & Fraser, 2019).

Chromatin loops represent both structural and functional feature of chromatin architecture, can range from few kilo-bases (kb) to mega-bases (Mb) in size and are universal feature of eukaryotic genome (Misteli, 2020). At the smallest scale, up to 200 kb, chromatin loops typically facilitate interaction of one or few genes with their regulatory elements and are known as promoter-enhancer loops (**Fig. 2**; Tolhuis *et al*, 2002; Sanyal *et al*, 2012). Enhancer regions serve as a platform for recruitment of sequence and context specific transcription factors (TFs), chromatin modifiers and RNA Polymerase II (RNA Pol II). They promote gene transcription through local chromatin remodeling and recruitment of other transcription cofactors (Vermunt *et al*, 2019). Given that enhancers vastly outnumber the genes and are often located at a great distance from target promoters (Schoenfelder & Fraser, 2019), they might loop one to another forming larger clusters which in turn regulate more genes, in a context-specific manner. Hi-C studies revealed that looping is a widespread phenomenon (Rao *et al*, 2014), wherein some loops show invariability across developmental stages and differentiation and others are highly dynamic and susceptible to modulation (Vermunt *et al*, 2019). Stable pre-looping of enhancer-promoter contacts in anticipation of signaling is functionally significant for immediate/early signaling responses (Jin *et al*, 2013; Kolovos *et al*, 2016) and suggests that ongoing transcription is not a requirement for initiation of enhancer-promoter contacts. Which universal factors do, in fact, contribute to loop formation, is not fully understood. Few proteins are observed enriched at the loop anchors – Mediator complex, YY1, CTCF and cohesin complex, although to what extent they contribute to gene expression via looping remains heavily debated (Rao *et al*, 2014; Beagan *et al*, 2017; Weintraub *et al*, 2017; El Khattabi *et al*, 2019).

Enhancer-promoter loops are often embedded in longer Mbp-sized domain (**Fig. 2.**; Dekker *et al*, 2013; Nora *et al*, 2012; Symmons *et al*, 2014; Hsieh *et al*, 2015) named topologically associated domains (TADs). TADs were initially detected through 40kb resolution Hi-C maps, and are defined by the high interaction propensity among the genomic regions they encompass compared to a lack of interactions with regions in other domains (Dixon *et al*, 2012; Nora *et al*, 2012; Sexton *et al*, 2012). While promoter-enhancer loops seem to be dynamic, cell type-specific and subject to context-dependent remodeling, TAD boundaries are thought to be stable across cellular divisions, cell types, even species (Dixon *et al*, 2016). Given that genome-wide proximity ligation approaches are probabilistic in nature and reflect population average, existence and stability of TADs had recently been brought to question. Single-cell studies coupled with oligo-paint microscopy have revealed that TADs are *bona fide* structures, but the strength of their

boundaries vary in a cell population and only strong interactions seem to be conserved (Nagano *et al*, 2013; Bintu *et al*, 2018). TAD distribution, in general, coincides with chromatin signatures, gene expression, replication timing, lamina and chromocenters (Dixon *et al*, 2012; Nora *et al*, 2012; Sexton *et al*, 2012; Wijchers *et al*, 2015). Functionally, TADs are thought to confine promoter-enhancer interactions of genes so as to ensure coordinated expression of neighboring groups of genes, while simultaneously restricting contacts with other genomic regions (Symmons *et al*, 2014). Notably, TAD boundaries are enriched in active genes and genes belonging to the same TAD tend to be coregulated (Nora *et al*, 2012; Symmons *et al*, 2014). Moreover, increasing distance of enhancers from target gene does not affect their interaction as long as they are in the same TAD, while TAD boundary disruption leads to increase in inter-TAD contacts (Symmons *et al*, 2014). A signature feature of a substantial number of TAD boundaries is their demarcation by CTCF (Rao *et al*, 2014; Dixon *et al*, 2015). Notably, not all the CTCF binding sites in the genome are found at TAD boundaries and ubiquitous CTCF binding sites mediate boundary formation in some tissues while not in others (Vermunt *et al*, 2019).

Often coinciding with TAD boundaries are anchors of loop domains (**Fig. 2**; Rao *et al*, 2014), which represent CTCF-anchored loops, 185 kb in length on average, with significant function in regulation of gene expression through control of encompassed enhancer-promoter interactions. Alteration of CTCF binding sites or their orientation can affect insulation strength between two TADs and allow for aberrant inter-TAD interactions (de Wit *et al*, 2015; Hanssen *et al*, 2017). It is well accepted that CTCF-CTCF domains are formed through cooperative activity of CTCF and some loop extruding factor (LEF), most likely cohesin or condensin complexes (Fudenberg *et al*, 2017; Davidson *et al*, 2019; Nora *et al*, 2020). In the “loop extrusion” model, a ring-shaped LEF, which contain structural maintenance of chromosomes (SMC) proteins, bind in non-specific manner and slide along the chromatin fiber in ATP dependent fashion. Loaded LEF then extrudes the chromatin fiber, forming the loop, until it reaches two convergent CTCF proteins which hinder the further movement. Such model offers a possible explanation of how long-range chromatin interactions occur beyond the simple diffusion and how CTCF/LEF can guide enhancer-promoter interactions, thereby modulating transcription. TAD-like domains are detected across evolution, however, not all TAD-like domains rely on CTCF proteins. In *D. melanogaster* (fruit fly; Gambetta & Furlong, 2018), *S. cerevisiae* (yeast; Hsieh *et al*, 2015), *Caulobacter crescentus* (bacteria; Le *et al*, 2013) and *C. elegans* (nematode; Crane *et al*, 2015) such domains are either not enriched for CTCF at their boundaries, or CTCF play a non-essential role in their establishment. A common feature of such domains is strong enrichment for highly expressed genes, suggesting that active transcription could be acting as a driving force for domain establishment, and delineating additional category of domains, known as compartmental domains (Rowley *et al*, 2017; Beagan

& Phillips-Cremins, 2020). Compartmental domains are defined through Hi-C analysis, albeit with high 5-10 kb resolution using a different detection algorithm. The structure of compartmental domain is defined by the transcriptional status and identified in Hi-C maps through PCA, where boundaries between adjacent compartmental domains reflect the switch in transcriptional activity. Contrary to loop domains, compartmental domains lack the corner dot signal in Hi-C and are not CTCF dependent, as seen in CTCF and cohesin acute depletion experiments (Nora *et al*, 2017; Rao *et al*, 2017). Therefore, compartmental domains are not the same as TADs, but TADs can be either CTCF loops, compartmental domains or combination of the two (Rowley & Corces, 2018).

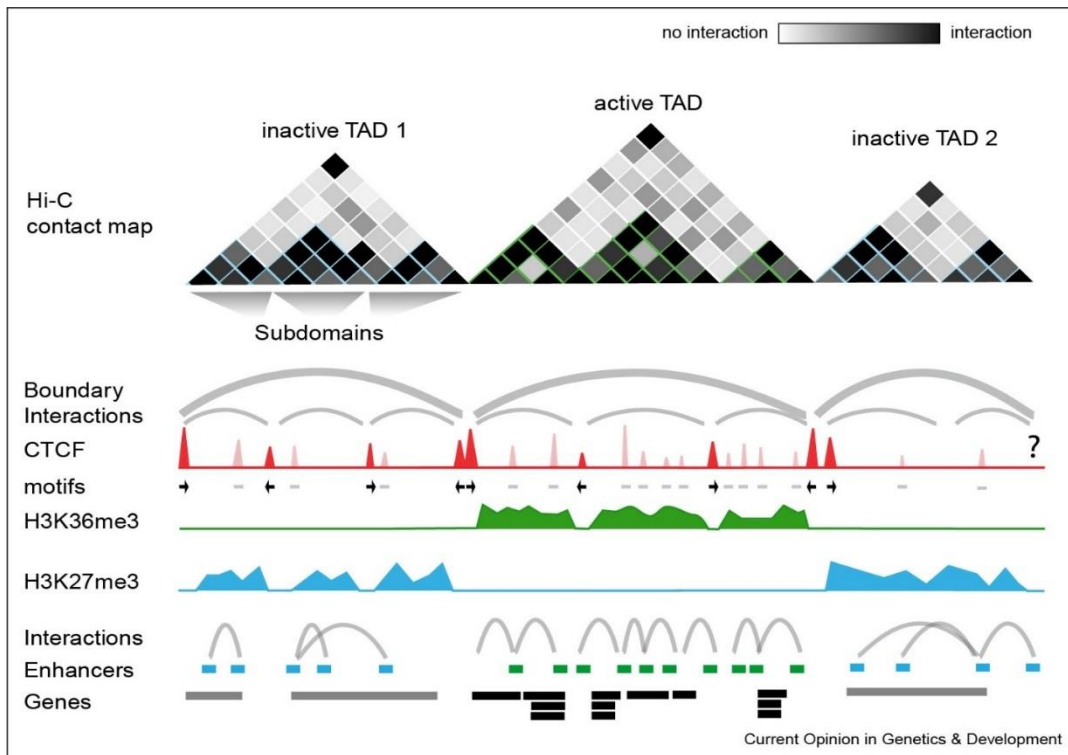


Figure 2 Hi-C derived contact map outlines intermediate chromatin structure

Schematic representation of Hi-C interaction map (*top*) where contact domains are represented as triangles emerging from Hi-C derived interaction frequencies. Prominent interactions occur predominantly within TADs while TAD boundaries are demarcated by absence of interactions and observed as triangle edges. TAD boundaries depend on the binding of CTCF (shown as a schematic ChIP-seq track in *red*) in the convergent orientation (*black arrows* indicate motif orientation), while CTCF sites not contributing to TAD boundaries formation are shown in pale red and grey motifs. As not all TAD boundaries are bound by CTCF it is likely that additional factors may be involved in their function (indicated by question mark). Schematic ChIP-seq track for of active (H3K36me3; *green*) and a repressive (H3K27me3; *blue*) histone modifications indicate their distribution reflects gene activity in TADs. Active TADs are gene-rich (black bars for active genes) in contrast to gene-poor repressed domains (*grey bars*). Enhancer-promoter interactions are smaller in size than CTCF-loops and nested in TADs. Adapted from (Ali *et al*, 2016).

1.2 Mechanisms of chromatin interactions

Recent developments of chromatin biology field led to the description of chromatin as a structure emerging from a series of nested interactions. Although such highly ordered structure implies rigidity of chromatin contacts, it is becoming increasingly appreciated that it is in fact a very dynamic property of the cell, where fluctuations of local and genome-wide chromatin interactions are controlled through a plethora of factors. What sort of forces drive the organization of chromatin folding, and which factors are modulating its dynamics is still an open question for the chromatin field.

Studies of polymer physics and computational modeling have revealed that the polymeric nature of chromatin fiber could explain some of its 3D folding features (Parmar *et al*, 2019; Misteli, 2020). Chromatin essentially acts as a semi-flexible and self-avoiding polymer, which limits possible diffusion distances and self-interaction. Such behavior can be easily modeled in a simple scenario where transient chromatin contacts form through random collisions in nuclei. Indeed, a movement of particular chromatin locus by diffusion, as observed by live-cell imaging, allows it to explore its immediate surrounding (Misteli, 2020). The stability of random genomic interactions can further be increased by diverse molecular factors, as shown through modeling of enhancer-promoter loop formation (Bianco *et al*, 2018), TADs and compartments (Nuebler *et al*, 2018; Parmar *et al*, 2019). A fundamental mechanism of protein and other macromolecule mobility in the nucleus is thermally induced random movement or Brownian motion, whereby single molecules roam through the nuclear environment. This movement is restricted to confined spaces by some macromolecular complexes in the nucleus, which reduce the effective exploration space, and facilitated by others (Halford & Marko, 2004).

Notably, not all features of macromolecular movement can be described by diffusion suggesting that other, active interactions play a role. This is amply evident in the case of TADs, where convergent CTCF sites at domain boundaries are favored over other orientations. Our understanding of emergence of TAD boundaries and recapitulation of original Hi-C maps therefore needs to be supplemented by other models, such as loop-extrusion through CTCF/LEF (Banigan *et al*, 2020; Starkov *et al*, 2021). Indeed, simulations of loop extrusion by CTCF/LEF complex are remarkably successful in recapitulating the Hi-C contact frequency maps (Fudenberg *et al*, 2017, 2016). Reeling of the chromatin through SMC containing LEF, where each motor unit operates in opposite direction to give rise to a chromatin loop was recently observed in *in vitro* experiments in *S. cerevisiae* (Ganji *et al*, 2018), *H. sapiens* (Davidson *et al*, 2019) and *X. laevis* eggs (Golfier

et al, 2020). This is supported by depletion of CTCF (Nora *et al*, 2017) and cohesin complex (Rao *et al*, 2017; Schwarzer *et al*, 2017; Gassler *et al*, 2017), which lead to disruption of TAD boundaries and in case of cohesin, all CTCF-loops. Formation of larger-scale compartments, on the other hand, is thought to occur through self-association of chromatin loci with similar transcriptional and chromatin states (Sexton *et al*, 2012; Rowley *et al*, 2017; Szabo *et al*, 2018). In this case, homotypic interactions of histone modifications and TFs/cofactors are driving the compartmentalization process. The idea of two competing and often opposing processes of loop extrusion and compartmentalization as drivers of intermediate level of chromatin folding has been tested using cohesin depletion experiments. It was shown that cohesin removal does not impede compartmentalization; in fact, its depletion strengthens existing compartments, suggesting that the two organizational processes are disparate (Haarhuis *et al*, 2017; Rao *et al*, 2017; Schwarzer *et al*, 2017). Exemplified in the case of TAD and compartment domain formation, various molecular players, such as architectural proteins, histone modifiers, components of transcriptional machinery, etc., could play a role in determining the patterns of genomic interactions. It should be noted however, that protein diffusion in nuclear space is rapid and their residence time on chromatin is typically measured in minutes, or even seconds. The apparent stochasticity of chromatin movement and high turnover of chromatin binding molecules imposes a question of how an apparent steady-state of chromatin structure emerges from such dynamic components?

A possible explanation lies in non-covalent, transient, and multivalent interactions of protein complexes in formation of membraneless microenvironments through the process of “phase separation” (Hnisz *et al*, 2017; reviewed in Boeynaems *et al*, 2018). Since both nucleic acids and proteins display features of liquid behavior - they are dynamic, spherical in shape and can interact among themselves, liquid-liquid phase separation had recently become a point of interest in chromatin folding studies (Banani *et al*, 2017; Palikyras & Papantonis, 2019). Liquid-liquid phase separation mediated by proteins relies on the presence of low-complexity intrinsically disordered regions (IDR) of such proteins. IDRs in the proteins can be nucleating agents that drive the formation of liquid droplets via weak multivalent interactions. A typical property of such droplets is a high local concentration of encompassed molecules and their insulation from the surrounding environment. Well known examples of nuclear bodies that are thought to form via phase separation are the nucleolus (Brangwynne *et al*, 2011; Frottin *et al*, 2019), Cajal bodies (Handwerger *et al*, 2005), stress granules (Wheeler *et al*, 2016) and nuclear paraspeckles (Yamazaki *et al*, 2018).

The potential significance of phase separation extends beyond nuclear bodies and into functional genome folding. Constitutive heterochromatin is enriched for H3K9me2/3 and is deposited and maintained

through coordinated activity of heterochromatin protein 1 α (HP1 α) and SUV39H1, which can read such histone modifications, and the HP1 α -scaffolding protein TRIM28. It was recently demonstrated that in human (Larson *et al*, 2017) and *D. melanogaster* (Strom *et al*, 2017), when phosphorylated, HP1 α forms phase-separated droplets *in vitro* that are permissive for repressing factors and induce constitutive heterochromatin formation, but whether this holds true in physiological conditions remains to be seen (Erdel *et al*, 2020). Additionally, multivalent H3K9me3 interactions with chromodomains of “reader” proteins are shown to phase separate, revealing a possible mechanistic explanation for contribution of repressive histone marks in heterochromatinization (Wang *et al*, 2019). Polycomb group repressive complex 1 (PRC1) is another heterochromatinization driver that is known to utilize phase separation for its function. Canonical PRC1 relies on the Chromobox 2 (CBX2) subunit for recruitment onto H3K27me3 sites marked by PRC2, in order to induce chromatin compaction. Recently it was reported that PRC1 readily phase-separates in Polycomb bodies, owing to disordered protein regions of CBX2 (Plys *et al*, 2019; Tatomosian *et al*, 2019). Similarly, another PRC1 subunit named Polyhomeotic (Ph) also phase-separates, but Ph-mediated condensates facilitate a different PRC1 role, that of ligating ubiquitin onto histone H2A (Seif *et al*, 2020).

Formation of heterochromatin is characterized by compaction, therefore phase separation is an intuitive mechanism driving this process. This is not the case for euchromatin, which is often associated with more open, transcriptionally active, and dynamic structure. However, the idea of the transcriptional machinery being compartmentalized at the distinct positions in nuclei, wherein high local concentration of molecules coordinates transcription of coregulated genes, is certainly not new (Iborra *et al*, 1996; Ghamari *et al*, 2013; Ghavi-Helm *et al*, 2014). Often referred to as “transcriptional factories” or hubs (Mitchell & Fraser, 2008; Sutherland & Bickmore, 2009; Papantonis & Cook, 2013; Tsai *et al*, 2019), these multi-protein supra-molecular nuclear bodies are enriched for components of the transcriptional machinery, especially RNA Pol II and cofactors (Ghamari *et al*, 2013). Recently it was demonstrated that the long C-terminal domain (CTD) of RNA Pol II is, in fact, an IDR capable of inducing liquid-liquid phase separation. The efficiency of droplet formation was shown to be dependent on the phosphorylation status of RNA Pol II, supporting the role of active transcription as driving force of chromatin folding (Hnisz *et al*, 2017; Boehning *et al*, 2018). Additionally, phosphorylation status of RNA Pol II CTD may govern the switch between RNA Pol II association with mediator condensates associated with transcription initiation (hypophosphorylated CTD) and condensates enriched for splicing factors (hyperphosphorylated CTD; Guo *et al*, 2019). This suggests that association of RNA Pol II with a particular condensate is regulated by conventional CTD modifications with known regulatory role. Besides RNA Pol II itself, various TFs contain low complexity domains and very

dynamically phase separate in a TF type specific and concentration dependent manner (Chong *et al*, 2018; Boija *et al*, 2018; Lu *et al*, 2020; Cai *et al*, 2019). This was shown for superenhancer regions (Hnisz *et al*, 2017), which are associated with cell-specific gene expression and characterized by highly dense binding of master TFs such as Mediator complex and BRD4; both TFs contain IDR and are able to compartmentalize the transcription apparatus at superenhancers (Sabari *et al*, 2018). The lifetime of TF-formed condensates is extremely short, centered around 30 sec (Chong *et al*, 2018) and dependent on transcriptional output, which in turn positively correlates with condensate lifetime (Cho *et al*, 2018). It is not only that active transcription makes condensates persistent, but also the converse is true, as demonstrated through synthetic TF-driven condensation, where transcriptional output shows almost exponential growth with the increased condensate induction time (Schneider *et al*, 2021).

1.3 Role of transcription in shaping the genome

In the light of mechanistic insights into localized transcription and given that the structural composition of the genome is reflective of its transcriptional activity, it is intuitive to consider transcription as a contributor to chromatin folding. Additionally, individual cell types display distinct patterns of genome organization due to activity of disparate gene expression programs. In line with this, undifferentiated cell types, such as embryonic stem cells are characterized by relatively homogeneous chromatin organization, wherein large groups of genes are expressed at low levels. Upon differentiation, cells adopt distinct identity through specification of gene expression patterns accompanied by the extensive chromatin remodeling (Bonev *et al*, 2017). Active transcription seems to be a reliable predictor of higher-order chromatin structures, such as chromatin compartments (Rowley *et al*, 2017). In addition to this, transcription strongly correlates with local chromatin folding, as observed through chromatin remodeling coinciding with transcriptional bursts and chromatin compaction in the absence of transcription (van Steensel & Furlong, 2019). Thus, examining the contribution of transcription and all its molecular components to chromatin folding could help our understanding of their mutual relationship in different cellular contexts.

During the cell cycle, chromatin undergoes major structural changes alternating between the highly condensed mitotic and transcriptionally active interphase structures. The mitotic cell is characterized by chromosome condensation, loss of transcriptional activity and long-range chromatin contacts (Vagnarelli, 2013; Naumova *et al*, 2013). While mitotic cells are devoid of typical composite chromatin

structure, throughout the rest of the cell cycle genome compartmentalization and TAD structures remain largely stable (Nagano *et al*, 2017). Post-mitotic compartmentalization is thought to be driven by carryover histone modifications, “bookmarking” TFs (Teves *et al*, 2016) and homotypic chromatin interactions (Hsiung *et al*, 2015). Some of the active histone marks, such as H3K4me3 and H3K2me1 remain on the mitotic chromatin, while H3K27ac is mostly lost. As cells are exiting the mitosis, H3K27ac is deposited at the sites of cell specific genes and transcription is later re-established (Kang *et al*, 2020). When it comes to TAD remodeling in mitosis, reports are controversial. While earlier studies report complete loss of domains in prometaphase (Naumova *et al*, 2013; Gibcus *et al*, 2018), recent evidence shows that TADs are significantly lost, but still above the detection limits, and enriched for active histone marks (Zhang *et al*, 2019). Zhang and colleagues reported that domains are gradually established from smaller to larger TADs through deposition of CTCF and later cohesin, followed by increase in domain size and insulation strength. CTCF binding is at a complete loss in pro/metaphase of mitosis (Oomen *et al*, 2019), but gets replenished by telo/anaphase (Zhang *et al*, 2019), with cohesin complex loading following later. The temporal separation between CTCF and cohesin loading reveals that late-established loops are larger than early, and their formation correlates with cohesin deposition on chromatin (Zhang *et al*, 2019). However, most early chromatin loops are not enriched in CTCF, appear sooner than loops with converging CTCF, and their stability is hindered by formation of CTCF loops. Formation of early loops coincides with the early transcriptional burst and RNA Pol II occupancy, later followed by reduction in transcription intensity and their disruption (Zhang *et al*, 2019). Exit from mitosis corresponds to transcriptional spike at the genes involved in cell growth, but how is the transcriptional activity restarted and how it contributes to re-emerging genome structure is not clear. It is possible that transcription inhibition in condensed mitotic chromatin is established at the elongation stage of transcription and not necessarily initiation (Vaňková Hausnerová & Lanctôt, 2017). In this case, RNA Pol II accumulates on genes and is released into productive transcription as soon as chromatin decondenses, thus ensuring the correct temporal transcription of cell growth-related genes.

Global gene repression also occurs in larger time scales than cell cycling, e.g., during differentiation and development. Differentiation by human or mouse ES cells is characterized by extensive A/B compartment switching, wherein overall tendency towards a higher level of compaction is observed and repositioned genes are largely lineage-restricted (Dixon *et al*, 2015; Bonev *et al*, 2017). Compartment level changes are followed by remodeling of intra-TAD interactions, wherein TADs with increased interaction frequency correlate with upregulated genes and vice versa. Genome-wide gene silencing is largely achieved through deposition of repressive histone marks by Polycomb-group proteins, whose

function is executed through Polycomb response elements in *Drosophila*, but interestingly, no such elements are found in mammals. Instead, in mammalian cells it was demonstrated that long non-coding RNA *Xist* can bind to and instruct the Polycomb machinery in the process of X-chromosome inactivation or even invoke repositioning of X chromosome to nuclear lamina (Chen *et al*, 2016), connecting the active transcription with gene repression (Engreitz *et al*, 2013). Similar example is in mouse T-cell progenitors, where activation of *Bcl11b* is driven through lncRNA ThymoD that drives repositioning of the entire locus to the A-compartment (Isoda *et al*, 2017). Additionally, forced tethering of TFs to LacO promoter and formation of transcriptional preinitiation complexes demonstrated that transcription could cause nuclear repositioning of the genes (van Steensel & Furlong, 2019). The loss of transcriptional activity of pluripotency-related genes, such as Oct4, Nanog or Klf4 drives their repositioning to nuclear periphery in differentiation (Peric-Hupkes *et al*, 2010). Gene silencing seems to directly contribute not only to the loss of silenced enhancer-promoter contacts (Bonev *et al*, 2017; Chopra *et al*, 2012) but also to the gain of new contacts (Jing *et al*, 2008). Newly acquired cell-specific transcriptional programs are followed by establishment of enhancer-promoter contacts mediated by cell-specific TFs in neural and hematopoietic differentiation (Bonev *et al*, 2017; Chen *et al*, 2016).

Genome-wide chromatin remodeling is observed in acute signaling events such as T-lymphocyte activation. Indeed, the most dynamic parameter of genome structure are enhancer-promoter contacts, which display ~80% of cell-type specificity (Stadhouders *et al*, 2019), therefore, more often than not contribution of transcriptional activity on the genome folding is detected at this level. Activation of CD4+ and CD8+ lymphocytes is accompanied by increased expression of lineage specific genes and TFs and extensive chromatin remodeling (Bediaga *et al*, 2021). T-cell activation invokes partitioning of TADs into finer shorter chromatin domains that are decorated with increased intra-domain contacts, especially in regions encompassing marker genes of active T-cells. On the other hand, TAD boundaries associated with low chromatin accessibility and high nucleosome occupancy become strengthened. Interestingly, compartmental, TAD and enhancer-promoter loops are not significantly remodeled upon cytokine stimulation of endothelial cells, however inter-chromosomal chromatin contacts seem to change, showing a prominent increase upon stimulation (Calandrelli *et al*, 2020). Here, driver genes of inflammation (*SERPINE1*, *THBS1* and *VWF*) are transcribed first and accumulate at the inflammation relevant super-enhancers and mediate inter-chromosomal contacts. Attenuation of such RNA-chromatin interactions leads to abolishment of inflammatory signaling. Human neutrophil genomes display a dramatically segmented compartment organization, where key neutrophil response-specific genes reside in moderately euchromatin regions, well separated from highly active euchromatic regions. Upon microbe infection, such

moderately euchromatin regions are promptly repositioned to highly active compartments to execute acute response-related transcriptional program (Denholtz *et al*, 2020). This long-range movement is associated with loss of insulation at euchromatic subdomain boundaries and formation of new promoter enhancer loops at the immune-response genes. Interestingly, observed changes are driven by the cooperative activity of H3K27ac deposition and cohesin complex repositioning.

Studies of interplay between transcription and chromatin structure in perturbing environments outline the mutual dependency of the two processes upon appropriate stimuli. However, to what extent active transcription is necessary for establishing or maintaining steady-state chromatin is a question often addressed in RNA Pol II perturbation studies. Genome-wide chemical block of active transcription, with several RNA Pol II inhibitors (α -amanitin, triptolide or flavopiridol), seems to impair compartmentalization in bacteria (Le *et al*, 2013), fruit fly (Li *et al*, 2015b) and mice (Ke *et al*, 2017). These studies imply the necessity of active transcription for establishing compartments, although decoupling transcription *per se* from a generally permissive genomic surrounding, governed by histone modifications and TF binding, is a difficult task. Additionally, genome-wide depletion of RNA through RNase treatment leads to weakening of B-compartmentalization, although RNA itself can act as an effector of chromatin interactions and therefore, indirectly, RNA Pol II activity participates also in these events (Barutcu *et al*, 2019). Studies of embryonic development offer further insight, as genes are generally inactive up to a certain developmental stage. In fruit fly (Hug *et al*, 2017) and mice (Du *et al*, 2017c; Ke *et al*, 2017) appearance of transcription coincides with the formation of chromatin compartments and chemical RNA Pol II block in embryo leads to decrease in compartment formation. In contrast, co-occurrence of transcription and compartmentalization in development is not observed in zebrafish (*Danio rerio*), where compartments can be detected before establishment of a transcriptional program (Kaaij *et al*, 2018). However, it should be emphasized that the choice of chemical block and its dosage have variable effects on the amount of RNA Pol II remaining bound to chromatin, thus leading to nuanced differences in the observed effects on compartmentalization.

Interestingly, neither chemical inhibition of RNA Pol II activity nor inducible depletion of RNA Pol II subunit, seem to greatly affect mammalian TAD arrangement, albeit they weaken the insulation strength (Barutcu *et al*, 2019; Hsieh *et al*, 2020; Jiang *et al*, 2020) and lead to general increase in chromatin dynamics (Nagashima *et al*, 2019). Additionally, single-gene studies demonstrate that CRISPR-dCas9-mediated transcriptional activation is not sufficient for TAD formation (Bonev *et al*, 2017). TADs are, however, maintained through CTCF/cohesin in mammalian cells and cohesin complexes are enriched at the sites of active transcription. Here, contribution of transcription to maintenance of TADs could be indirect, as it is

shown that blocking the active transcription or altering transcriptional program, leads to cohesin repositioning (Busslinger *et al*, 2017; Olan *et al*, 2020). Movement of cohesin can further disrupt existing CTCF loops and induce remodeling of chromatin interactions (Heinz *et al*, 2018). Notably, X-chromosome inactivation through *Xist*, which leads to global repression of transcription, also results in cohesin eviction and TAD remodeling (Minajigi *et al*, 2015). Unlike TADs, compartmental domains are not formed through CTCF/cohesin and are enriched at active promoters, similar to domains observed in *Drosophila* genome (Rowley *et al*, 2017). Active genes in compartmental domains are distributed across the domain rather than at the boundaries. Here, triptolide treatment reduces the frequency of interactions inside the domain and the effect correlates with RNA Pol II occupancy (Rowley *et al*, 2017; Hsieh *et al*, 2020). In addition to RNA Pol II occupancy, active genes in compartmental domains display enrichment of other transcriptional facilitators such as mediator complex and YY1, which was shown to be reliable predictor of gene loops (Kagey *et al*, 2010). YY1 occupies promoters and enhancers, it is known to homodimerize and editing of YY1 binding sites or acute depletion of YY1 protein leads to decrease in enhancer-promoter contacts (Weintraub *et al*, 2017). This finding suggests that cooperative effect of RNA Pol II with variable TFs contributes to the chromatin folding. Functional studies of single gene loci suggest that chromatin modifiers and transcription factors influence chromatin topology irrespective of active transcription. In one such example, transcription factor tethering was shown to be sufficient to drive compartment switching, whereby nuclear repositioning depends on the choice of TF (Wijchers *et al*, 2016). In fact, during cellular differentiation, many enhancer-promoter contacts are created *de novo*, but others are pre-existing and subject to chromatin remodeling (Schoenfelder & Fraser, 2019). Similar is demonstrated in early cellular response to TNF- α signaling, where enhancers are found pre-looped to their cognate promoters in anticipation of signaling (Jin *et al*, 2013; Kolovos *et al*, 2016). Such contacts can induce repositioning of gene locus into a transcriptionally permissive A-compartment, as demonstrated for Oct4 during cellular reprogramming (Stadhouders *et al*, 2018). The pre-looping phenomenon seems to hold true across different cell lines and signaling factors, suggesting that other modulators of transcription and their interactions can drive chromatin folding even before the productive transcription.

2. Senescence as a model for studying genome organization principles

Cell cycling is a continuous process which occurs a finite number of times under physiological conditions (Hayflick & Moorhead, 1961). Repetitive cycling of the cells through processes of replication and division negatively affects the authenticity of the replicated genome, as genome instability and DNA damage are propagated and accumulated over time. To control the potential damage and prevent the possible shift into malignancy, number of cell cycles hinges on telomere shortening with each cycle that finally results in permanent cell cycle arrest. This phenomenon of replicative exhaustion is termed replicative senescence, and is a physiological event that leads to senescent cell accumulation in normal tissues with age (Calcinotto *et al*, 2019). Senescence is a good example of “pan-genomic” regulatory network driven process (Brackley *et al*, 2020), as it is decorated by intricate traits emerging from numerous molecular interactions, and orchestrated by various factors such as 3D chromatin architecture, non-coding RNA, and even regulatory crosstalk between cells.

Throughout their life, cells are exposed to various types of physiological stress, such as injury, metabolic shock, hypoxia, radiation, or genotoxic stress, all of which can induce premature senescence as means to cope with the local insult and prevent further damage to surrounding cells. In addition to being a tool of stress management, senescence plays a significant role in non-stress induced physiological events, such as development, tissue remodeling and repair (Davaapil *et al*, 2017). Finally, senescence manifests itself through permanent cell cycle arrest and can be induced in malignant tissues to prevent aberrant proliferation, making senescence also relevant for development of anti-cancer therapies. Dysregulation of the senescence program, on the other hand, can hinder the regenerative capacity of the tissues and create a pro-inflammatory milieu supportive of age-related diseases and cancer (Calcinotto *et al*, 2019). This duality of the senescence program implies its dependence on balancing mechanisms that can shift the activity of senescent cells either positively or negatively for tissue homeostasis. It seems that persistence of the senescent phenotype in a given microenvironment is a control point; clearance of senescent cells by activity of immune cells is a necessary step for the resolution of ensuing cascades (Ito *et al*, 2017).

Senescent phenotypes vary greatly between cell types and depend on the initiating event, which might explain how senescence is rather dynamic and not a definitive cellular state. The common feature of different senescent programs is exit from the cell cycle, but proliferation arrest is merely an entry into senescent process, followed by a cascade of other cellular events (**Fig. 3**). Initiation of senescence, marked

by proliferation arrest, invokes extensive chromatin remodeling, morphology changes, modulation of metabolism and secretion of compendium of signaling molecules, as it progresses into early senescence (Herranz & Gil, 2018). Late senescence comes about in cases where the senescent state is retained for extended periods of time without provoking cell death.

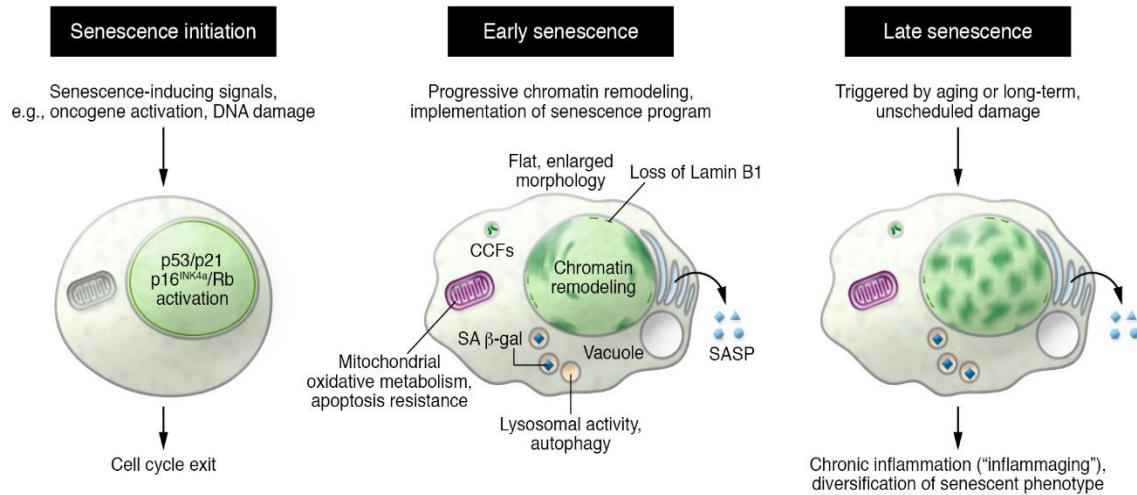


Figure 3. Cellular progression into senescence

Outline of gradual establishment of senescent phenotype. Senescence is initiated by gene expression coordinated entry into cell cycle arrest. Early senescence involves large scale changes in chromatin, mitochondrial and lysosomal activity, and release of pro-inflammatory signals through SASP. Late senescence is a long-term senescent state that leads to chronic inflammation. Reprint from (Herranz & Gil, 2018)

Cell cycle exit upon senescence initiation can be triggered by intrinsic (oxidative stress, telomere shortening or hyperproliferation) and extrinsic cues (ultraviolet, γ -irradiation, chemotherapeutic drugs). These stressors invoke persistent DNA damage response (DDR) that eventually leads to activation of p53/p21CIP1 axis (Groesser *et al*, 2011). Activity of p53/p21 occurs transiently in physiological non-senescent processes of DNA repair, where upon resolving the damage, cells progress into normal cell cycling. In the case of senescence, however, initiation steps rely on p53/p21, but maintenance of cell cycle arrest is achieved through activity of INK4/ARF, encoding three proteins – p14ARF (in human, p19ARF in mouse) and p16INK4a encoded by the *CDKN2A* gene, and p15INK4b encoded by *CDKN2B* (Kim & Sharpless, 2006). ARF protein promotes p53 activity in response to stress, through inactivation of the MDM2 protein, a negative regulator of p53. Notably, other modes of action were ascribed to ARF protein in cell cycle arrest, though interaction with different proteins such as E2F-1, MDMX, HIF-1 α , topoisomerase I, MYC, and nucleophosmin (Kim & Sharpless, 2006). INK4 proteins inhibit the activity of cyclin D-dependent kinases,

CDK4 and CDK6, and consequently CDK4/6-mediated phosphorylation of retinoblastoma (Rb) family members. Hypophosphorylated Rb binds to transcription factor E2F, which in turn leads to downregulation of E2F target genes and cell-cycle arrest. Depletion of p16 or Rb activity or their mutation in cell is enough to bypass senescence and conversely, overexpression of either leads to initiation of senescence program (Jarrard *et al*, 1999; Sage *et al*, 2003; McConnell *et al*, 1998). The expression profile of p16 coincides with age and is often used as senescence biomarker; it is barely detectable in young proliferative tissues and progressively upregulated in aging (Burd *et al*, 2013).

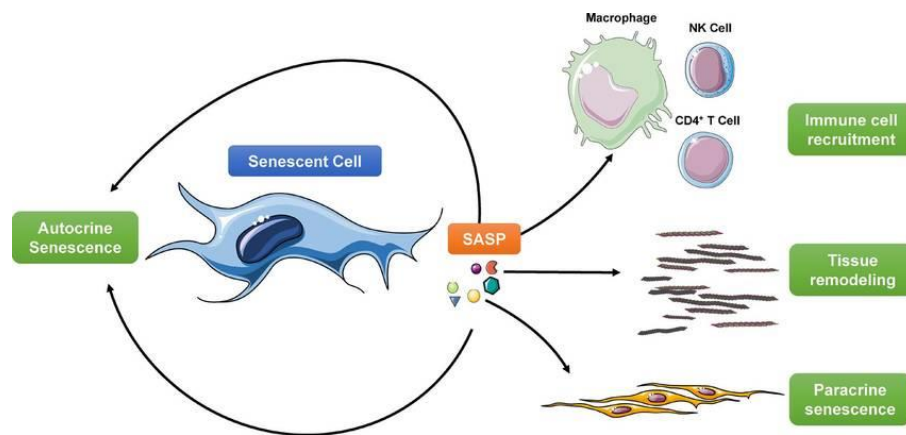


Figure 4. Effects of SASP signaling in senescence

Pro-inflammatory molecules released as a part of senescent associated secretory pathway mediate significant portion of senescent phenotype and act in paracrine and autocrine manner. Reprint from (Hugh and Gil, 2018)

Early senescence stage ascends beyond the cell-autonomous mechanisms through production and secretion of soluble and insoluble mixture of factors in a phenomenon known under the term senescence associated secretory phenotype (SASP, **Fig. 4**; Kuilman & Peeper, 2009; Coppé *et al*, 2010). SASP manifests variably in different cell types and depending on senescence-inducing stimulus, but usually comprises of cytokines, chemokines, growth factors, extracellular matrix proteases and other. The consequences of SASP are twofold: activation of senescence in malignant tissue can spread the anti-proliferative effect on neighboring malignant or non-malignant cells and participate in antitumoral response (Xue *et al*, 2007). On the other hand, SASP components are mostly pro-inflammatory, thereby they have potential to drive tumor progression. Chronic inflammation is typical feature of aging tissues, which could be explained through accumulation of senescent cells and elevation of SASP signaling in aging. Convergenly, elimination of SASP positive senescence cells leads to decrease in cytokines of aged mice (Baker *et al*, 2016). The positive effect of SASP signaling is manifested in processes such as wound healing (Demaria *et al*, 2014), wherein activity

of extracellular matrix remodelers and pro-inflammatory cytokines accelerate the tissue recovery, or promotion of tissue fibrosis (Krizhanovsky *et al*, 2008). Other metabolic changes in senescence are reflected in glycolytic and mitochondrial activity increase, which is thought to support energy requirements for SASP production (Dörr *et al*, 2013), as well as autophagy increase to provide supply of amino acids (Narita *et al*, 2011). The exact contribution of autophagy and mitophagy in senescence maintenance is controversial. In some cases, inhibition of autophagy facilitates senescence (Kang *et al*, 2015) and mitophagy is decreased (García-Prat *et al*, 2016). Mitophagy loss, however, leads to reactive oxygen species (ROS) production due to the increase in dysfunctional mitochondria, which in turn promotes senescence.

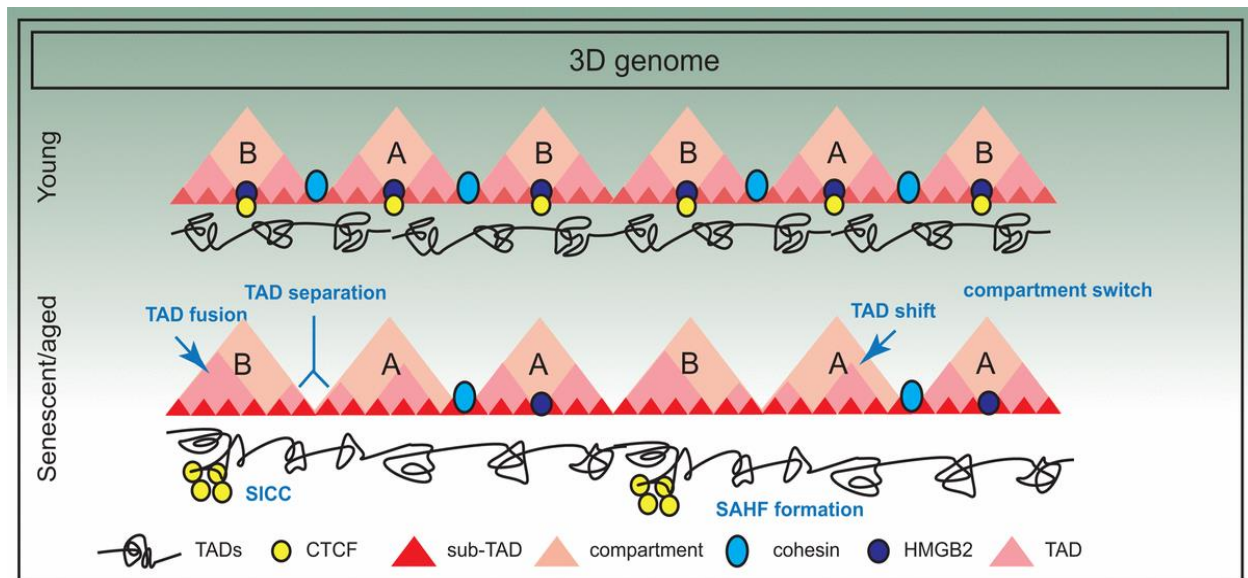


Figure 5. Changes in chromatin structure during senescence

Senescence progression is associated with extensive compartment switching and TAD remodeling due to significant loss of CTCF and other architectural proteins. Reprinted from (Yand and Sen, 2018)

Changes in cellular metabolism and morphology are accompanied by remarkable reorganization of chromatin in senescence (**Fig. 5**). Alterations in chromatin methylation, accessibility, histone marks, as well as noncoding RNAs vary between the senescence models, indicative of the diversity of senescent phenotypes (Pal & Tyler, 2016; Parry & Narita, 2016; Buschbeck & Hake, 2017; Nacarelli *et al*, 2017). DNA methylation at CpG islands affects gene silencing under physiological conditions and plays a role in cancer. DNA methylation in replicative senescence is decreased genome-wide, but focally increased at a subset of CpG islands, not unlike cancer models (Cruickshanks *et al*, 2013), offering a possible explanation for dual role of senescence as tumor suppressor and a cancer risk factor. The pattern of DNA methylation in senescence shows high reproducibility in contrast to cancer and suggests the process is programmed for

senescence, although it still results in genome instability (Xie *et al*, 2018). Complementary to methylation, global accessibility in senescence is increased with some focal declines; this tendency was also observed in oncogene-induced senescence, but differs between oncogene induced senescent types (OIS; De Cecco *et al*, 2013; Parry *et al*, 2018).

Perhaps the most striking feature of senescent chromatin reorganization is the emergence of senescence-associated-heterochromatic foci (SAHF), most prominent in Ras oncogene-induced senescence (Narita *et al*, 2003). SAHF form via reorganization of heterochromatic regions, accumulation of high mobility group A proteins and loss of linker histone H1 (Chandra *et al*, 2012; Salama *et al*, 2014). SAHF are enriched in former LAD-belonging regions which detach from the nuclear envelope upon senescence induction due to the loss of Lamin B1 (LMNB1; Sadaie *et al*, 2013). Unlike typical constitutive heterochromatin, SAHF do not include centromeres or telomeres (Narita *et al*, 2003). This suggests that they do not form through genome-wide expansion of heterochromatin, but rather through spatial rearrangement of existing constitutive (H3K9me3) and facultative (H3K27me3) heterochromatic histone marks (Chandra *et al*, 2012). Redistribution of repressive histone marks suggests the functional role of SAHF in silencing the cell-cycle relevant genes. Indeed, formation of SAHF relies on activation of p16/Rb pathway and subsequent silencing of E2F family regulated genes (Narita, 2007). In contrast to SAHF, senescence-associated distension of satellites (SADS) represents unfolding of the constitutive heterochromatin on satellite regions (Swanson *et al*, 2013). These two opposite processes are thought to be jointly mediated by the loss of LMNB1 and the collapse of the nuclear envelope. Depletion of LMNB1 in senescence is achieved through silencing of Rb-dependent genes and autophagic degradation, which leads to compromised nuclear integrity and leakage of chromatin into cytoplasm.

How is then higher-order chromatin folding altered in senescence and what are the functional consequences of those changes? Oncogene induced senescence seems to be characterized by significant alterations of loop domain structures, wherein cohesin is depleted from CTCF-anchored loops and enriched at newly formed, transcription-dependent enhancer-promoter contacts (Olan *et al*, 2020). Here, cohesin is deposited at the 3' ends of a subset of OIS-relevant genes, such as *IL1B*, promoting formation of new loops and gene upregulation. This suggests a possible positive feedback loop whereby newly formed contacts promote transcription, which in turn allocates more cohesin to active genes and formation of new contacts, contributing to phenotype specific gene expression profile. Whether this principle holds true in other senescence types, such as replicative senescence, remains to be seen. On a larger chromatin scale, TADs are marked by a loss of local insulation and intra-TAD interactions at the LMNB1 dependent loci which are

enriched in H3K9me3, in concordance with SAHF formation (Chandra *et al*, 2012). In replicative senescence, specifically in a model of “deep” senescence, changes of chromatin landscape were associated with general chromatin compaction, conservation of TADs, but prominent B-to-A switching of compartments that encompass senescence-relevant genes (Criscione *et al*, 2016).

2.1 HMG family of architectural proteins

High mobility group (HMG) proteins are most abundant non-histone proteins discovered in calf thymus chromatin and named according to their rapid mobility in gel electrophoresis (Goodwin *et al*, 1973). This large group of proteins is characterized by relatively small size, high content of charged amino acids and chromatin binding capacity, which facilitates their function as architectural proteins (Stros *et al*, 2009). Based on the structural and biochemical properties, HMGs are divided into three distinct families (Fig. 6): the AT-hook (HMGA), nucleosome-binding (HMGN), and HMG-box family (HMGB).

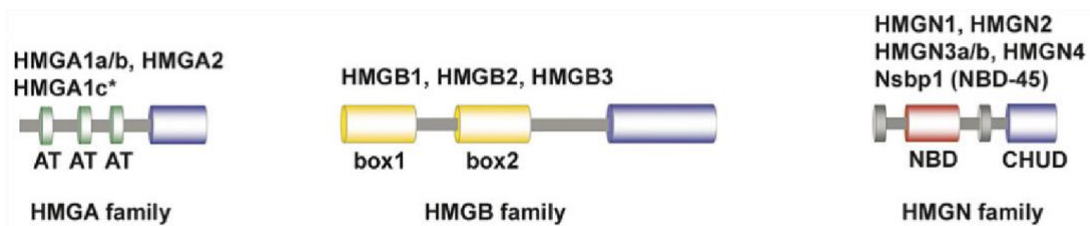


Figure 6. HMG protein family

HMG groups of proteins differ by the domain composition. HMGA family is characterized through presence of “AT hook” domains, HMGB family contains one or more HMG-box domains and HMGN family carries a nucleosome binding domain. Adapted from (Hock *et al*, 2007)

The HMGA family contains 4 proteins coded by the *HMGA1* and *HMGA2* genes, with typical “AT-hook” domains through which they bind AT stretches in DNA in a sequence-independent but non-random manner to invoke conformational changes in DNA (Maher & Nathans, 1996). In addition to DNA-binding, though their acidic C-tails HMGA proteins interact with other proteins, such as TFs and are known to be a part of enhanceosomes (Reeves & Beckerbauer, 2001) and mediators of enhancer-promoter loops (Reeve, 2003). Abundance of HMGA proteins in undifferentiated and proliferating cells in early embryo, and their suppression in adult tissues, suggest they have a prominent role in development (Hock *et al*, 2007). Indeed,

HMGA1 deficient mice suffer cardiac hypertrophy, malignancies and type 2 diabetes (Fedele *et al*, 2006), while HMGA2 deficient mice display “pigmy” phenotype and due to dysregulation of spermatogenesis, the males are sterile (Zhou *et al*, 1995; Chieffi *et al*, 2002). Overexpression of HMGA proteins is typical feature of several types of malignant and benign tumors (Sgarra *et al*, 2004). HMGA1 seems to be important in embryonic epithelial-mesenchymal transition, cell movement and metastasis, while HMGA2 is mediator of TGF β signaling in the processes of differentiation, tumor invasiveness and metastasis (Reeves, 2001; Thuault *et al*, 2006). HMGA proteins often carry out their roles through protein partners, such as NF κ B, p53 and Rb, which might explain why HMGA proteins are often overexpressed in cancer (Hock *et al*, 2007). They also inhibit nucleotide excision repair process in cells, which renders DNA unstable and prone to accumulation of DNA damage. High expression of HMGA proteins is observed in senescence, where they contribute to formation of SAHF and display antiproliferative activity (Narita *et al*, 2006).

HMGN family proteins are characterized by a nucleosome-binding domain, a nuclear localization signal and regulatory domain in their C-terminus and are only found in vertebrates (Hock *et al*, 2007). Their major function is binding and regulating nucleosomes, and *HMGN* expression levels correlate with differentiation (Körner *et al*, 2003): two of the five family members, HMGN1 and HMGN2, are found in all embryonic mouse tissues and are progressively depleted throughout differentiation (Furusawa *et al*, 2006). Interestingly, they both remain active in renewing cell types, such as the basal layer of epithelia. Ectopic *HMGN1* overexpression prevents differentiation during erythropoiesis, osteogenesis and limb bud development (Hock *et al*, 2007). Overexpression of HMGNs is also observed in mouse model of Down syndrome, suggesting they have a role in postnatal development (Potier *et al*, 2006). On the other hand, mice deficient in HMGN1 display high sensitivity to stress (UV and irradiation), decrease in fertility, high incidence of tumors and defective corneal epithelium maturation (Birger *et al*, 2003). These observations could be explained through HMGNs role in unwinding the chromatin and providing instructive role for DNA repair machinery recruitment (Birger *et al*, 2003). HMGN1 can act as a part of damage-associated molecular pattern with pro-inflammatory effect, when released from the cells, albeit through an unknown mechanism (Yang *et al*, 2012).

HMGBs are the most abundant among HMG proteins; they are highly conserved and comprise of four members (HMGB1-4; Goodwin *et al*, 1973). The first three members share a high degree of amino-acid sequence similarity, but their disparate phenotypes in knockout mice indicates that *in vivo* roles are distinct. Recently discovered HMGB4, on the other hand, lacks the typical acidic tail (Stros, 2010). Each HMGB contains two DNA binding-box domains – A and B, and an acidic C-terminal tail (except for HMGB4;

Thomas & Travers, 2001). Notably, variable numbers of HMG-box domains are detected in other types of proteins, which carry out diverse functional roles (Kang *et al*, 2014). The binding of HMGB proteins to DNA is not sequence specific and is carried out through the integration of box domains into the minor groove of B-type DNA (Ueda & Yoshida, 2010). The binding of HMGBs to DNA is highly dynamic, which is thought to play a role in transcriptional control. HMGBs could facilitate TF binding at their respective loci by unwinding or bending DNA before TFs are loaded. Alternatively, HMGBs can promote cooperative TFs binding after initiating TFs are locked in, by unwinding or bending DNA, or they may directly interact with TFs (Stros, 2010). HMGBs participate in cellular processes such as transcription, replication, DNA repair, inflammation, and tumor invasiveness (Lotze & Tracey, 2005). HMGB expression profiles throughout development are variable; HMGB1 is ubiquitously expressed throughout the development and remains in adult mouse tissues, HMGB2 is highly expressed in embryonic tissues and lost in adult mice, except for lymphoid organs and testes, while HMGB3 mainly in hematopoietic cells (Lotze & Tracey, 2005; Ronfani *et al*, 2001; Nemeth *et al*, 2003). HMGB4 is present in some stages of embryonic development and remains expressed in adult testes, like HMGB2, where it has a role as a transcriptional repressor (Catena *et al*, 2009). Despite sequence conservation and indications of functional redundancy, loss of HMGB1 cannot be compensated by HMGB2 activity (Ronfani *et al*, 2001). HMGB1-knockout mice display hypoglycemia and die soon after birth, while HMGB2 deficient mice are viable but suffer germ cell differentiation malfunction which renders them infertile (Ronfani *et al*, 2001). HMGB3 deficient mice display non-lethal erythropoiesis dysregulation resulting in overproduction of erythrocytes and lack of myeloid and B-cells (Nemeth *et al*, 2003).

2.2 About HMGB1

HMGB1 is a highly conserved mammalian protein with homologues detected across the evolution, among the others in yeast, fruit fly, nematode, fish, plant and bacteria (Bustin, 2001; Giavara *et al*, 2005; Wu *et al*, 2003). The homology of 215 amino acid long HMGB1 in human and rodents is almost absolute, with the difference of two amino acids in the acidic C-tail. Like other HMGB proteins, HMGB1 contains two DNA binding domains, A-box and B-box, which carry nuclear emigration signal (**Fig. 7**). Additionally, HMGB1 contains two nuclear localization signals, enabling the HMGB1 proteins to shuttle between nucleus and cytoplasm (Bonaldi *et al*, 2003). Through specific residues in their primary sequence, HMGB1s interact with various proteins, such as RAGE, TLR4 and p53 transactivating domains (Huttunen *et al*, 2002; Li *et al*, 2003),

mediating cell migration, inflammation, and transcription, respectively. Here B-box domain carries out pro-inflammatory activity, while A-box acts as an antagonist together with C-terminal tail (Gong *et al*, 2010). C-terminal domain modulates HMGB1 binding to DNA by interacting with DNA-binding domains and histones H1 and H3 (Stros, 1998; Cato *et al*, 2008; Kawase *et al*, 2008). The two DNA-binding domains fold into “L” shaped structure, with slight differences in composition of amino acids, where concave side of the domains is thought to be responsible for transiently binding the minor groove of DNA molecule, in “hit and run” fashion (Gerlitz *et al*, 2009). HMGB1 bends the DNA (Pil *et al*, 1993) inducing a change in DNA conformation through looping, unwinding, or compacting of the DNA molecule (Stros *et al*, 1994; Yoshida *et al*, 1984; Javaherian *et al*, 1978). DNA bending is thought to be carried out through intercalation of bulky amino acid residues of HMGB1 DNA-binding boxes into successive base pairs in the minor groove of DNA, bending it toward major groove (Stros, 2010). The bending activity and affinity of HMGB1 towards DNA molecule is tightly regulated through posttranslational modifications of HMGB1, flanking sequences of HMG boxes (Stros, 1998; Ugrinova *et al*, 2001) and acidic tail (Yoshida *et al*, 1984).

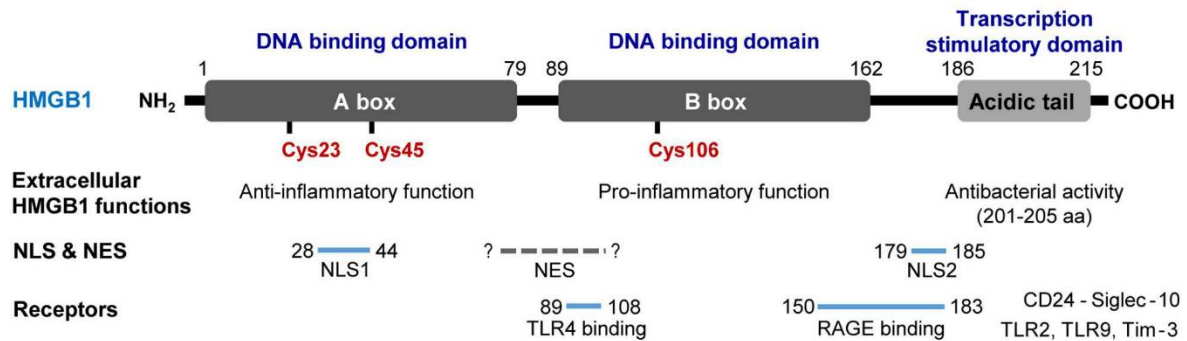


Figure 7. Detailed structure of HMGB1

Two DNA binding domains, A and B box are followed by the acidic tail at the C-terminus. A and B box domains govern the DNA binding of HMGB1, and acidic tail modulates their interaction. B box is associated with pro-inflammatory actions of HMGB1 while A box is involved in termination of inflammatory response. HMGB1 carry two nuclear localization signals (NLS), one nuclear export signal (NES) and domains for binding to inflammatory receptors. Adapted from (Kwak *et al*, 2020)

Expression of HMGB1 is ubiquitous and most prominent in spleen and thymus, with high levels in most human tumors and low during aging (Müller *et al*, 2004; Prasad & Thakur, 1990). HMGB1s are ubiquitously expressed in adult tissues; lymphoid tissues and testes contain high levels of HMGB1 whereas low expression was found in brain and liver. High levels of HMGB1 are generally detected in less differentiated cells, a finding which may explain a high expression of HMGB1 in neoplastic cells which often undergo dedifferentiation. Elevated HMGB1 expression is found in proliferating tissues, dividing cells and

cells treated with estrogen. Levels of HMGB1 decrease with age in rat thymus, aging of mouse brain, however their expression is elevated in astrocytes. In *X. laevis* and *D. melanogaster* development, HMGB1s are substitute proteins for H1 histone that is absent in embryogenesis in these models. In mouse this was not the case, as HMGB1 levels changed from stage to stage.

Nuclear HMGB1 is associated with DNA replication, repair, recombination, transcription, and genomic stability, and extracellular HMGB1 plays a role in inflammation, immunity, cell growth, proliferation, and cell death (**Fig. 8**). In nuclei, HMGB1 can bind to linker DNA between the nucleosomes, having an opposing effect to that of H1 histones (Cato *et al*, 2008). While H1 stabilizes the nucleosome position by binding to DNA entry/exit site of the nucleosome, HMGB1 binding induces the increase in chromatin accessibility by distorting the nucleosome. It directly interacts with H1 histone through its acidic tail and this interaction can be modulated by changes in redox state, pH and local ionic concentration (Kohlstaedt & Cole, 1994). Not only does HMGB1 influence nucleosome stability, but also it can recruit chromatin modifiers and induce nucleosome repositioning (Bonaldi *et al*, 2002), affect nucleosome biogenesis through chaperone activity (Celona *et al*, 2011) and prevent nuclear catastrophe and nucleosome release while bound to chromatin during inflammation (Kang *et al*, 2014). The role of HMGB1 in DNA replication in mammals (Bonne-Andrea *et al*, 1986) and viruses (Cotmore & Tattersall, 1998) is manifested through its unwinding activity, which is modulated through the posttranslational modifications; phosphorylated HMGB1 decreases the rate of replication (Bonne *et al*, 1979) while acetylated HMGB1 promotes it (Alexandrova & Beltchev, 1988). Interestingly, such activity of HMGB1 in DNA replication was demonstrated *in vitro* in isolates from regenerating rat liver but not the normal liver (Duguet *et al*, 1977), leaving the functional significance of HMGB1 in replication elusive. Loss of nuclear HMGB1 is associated with decrease in DNA damage repair efficiency; HMGB1 mediates mismatch repair (Genschel & Modrich, 2009), base excision repair (Prasad *et al*, 2007), nucleotide excision repair (Lange & Vasquez, 2009) and double strand break repair (Yumoto *et al*, 2011). Other roles attributed to nuclear HMGB1 include promotion of telomerase activity through upregulation of telomerase RNA in mouse (Giavara *et al*, 2005) and stimulation of gene transfer when it operates as cofactor of sleeping beauty transposase (Zayed *et al*, 2003). Role of HMGB1 as transcription factor is versatile (Singh & Dixon, 1990). Besides influencing nucleosome dynamics and thereby modulating general gene expression, HMGB1 directly binds TATA binding protein (TBP; Lee *et al*, 2014) and stimulates TBP binding to TATA regulatory elements. Lymphokine genes, regulated by conserved lymphokine elements-0 (CLE0) which respond to CLE0 binding protein, are also activated by HMGB1 binding due to the high homology between HMGB1 and CLE0 binding protein (Marrugo *et al*, 1996). HMGB1 can promote or repress gene expression through interference of sequence-

specific TF interactions with their target genes, as has been demonstrated for several TFs (Kang *et al*, 2015), including p53. The Both HMG-box domains and C-tail have binding capacity for p53, however A-box of HMGB1 displays the highest affinity for p53 binding; DNA bending by HMGB1 and recruitment of p53 therefore positively regulates p53 target genes and implicates HMGB1 in various stress-related responses (Stros *et al*, 2002; Rowell *et al*, 2012).

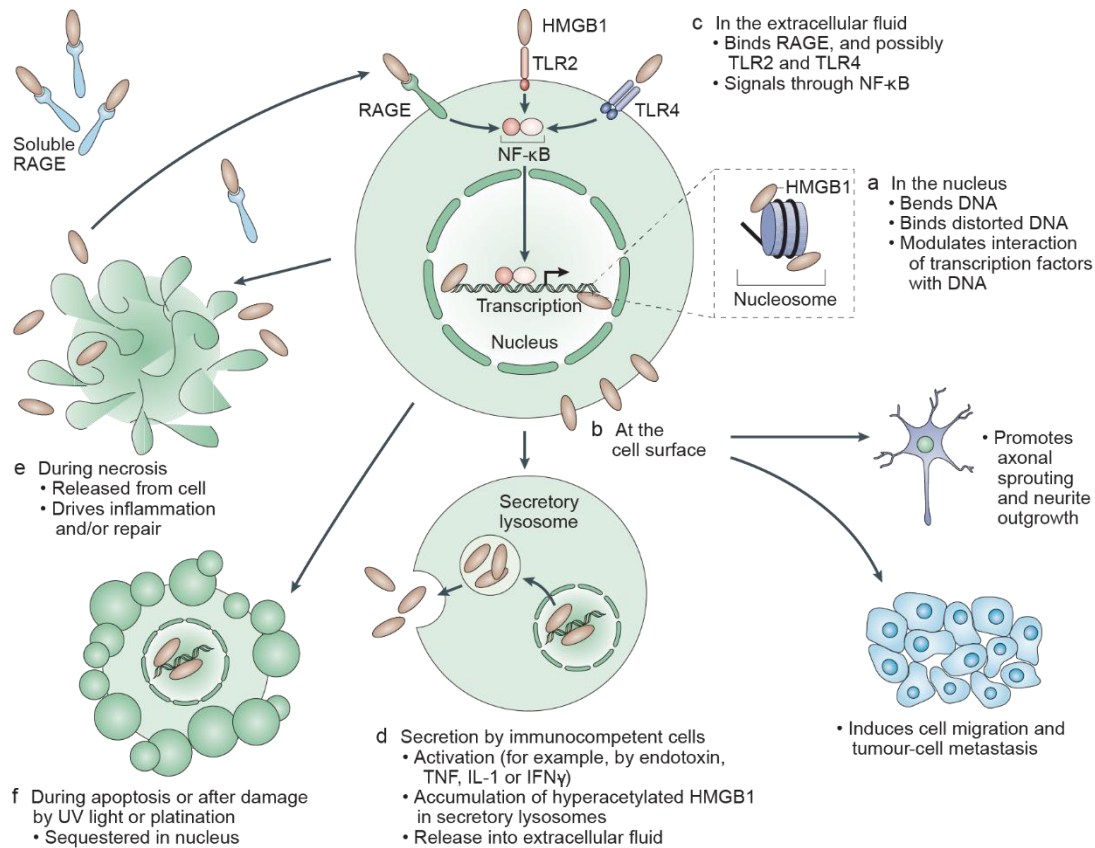


Figure 8. Overview of HMGB1 roles

(A) Nuclear HMGB1 binds to chromatin and induces conformational changes that reflect on the gene expression. (B) Present at the cellular membrane, HMGB1 is implicated in facilitating neuronal system development and cell migration. (C) Extracellular HMGB1 interacts with RAGE and TLR receptors to induce pro-inflammatory signaling. (D) Secreted through the vesicles, HMGB1 acts as pro-inflammatory signaling molecule. (E) Necrotic cells release HMGB1 in extracellular fluid where HMGB1 act as a paracrine signaling factor. (F) In stress HMGB1 can be sequestered in the nucleus to prevent its release. Reprinted from (Lotze and Tracey, 2005)

HMGB1 is actively secreted by immune cells or passively diffused out of injured or apoptotic cells (Fig. 8). When released into medium, extracellular HMGB1 was shown to promote murine erythroleukemia cell differentiation, malignant erythroblasts whose differentiation into red blood cells is impaired (Melloni

et al, 1995; Sparatore *et al*, 1996). The pro-inflammatory role of extracellular HMGB1 was established in 1999 and initiated a series of experimental work correlating the activity of HMGB1 with pro-inflammatory effects in tens of different cell types, through stimulation, production, and release of all major cytokines, chemokines, growth factors, adhesion and other pro-inflammatory molecules (reviewed in Kang *et al*, 2015). The role of HMGB1 in immunity is complemented with its activity as a chemokine factor to recruit innate immune and stem cell to the sites of insult and promote tissue repair and regeneration (Kang *et al*, 2015; Degryse *et al*, 2001; Palumbo & Bianchi, 2004). HMGB1 mediated cell-migration is carried out through binding to immunoglobulin family RAGE receptors (Bauer *et al*, 2013), but in some cases, such as pulmonary artery endothelial cells, HMGB1 can inhibit cell migration through TLR4. HMGB1 is implicated in promoting cellular proliferation in few cell types, while in cancer cells extracellular HMGB1 in reduced state promotes proliferation and oxidized induces apoptosis (Sundberg *et al*, 2009). This suggests the effect of HMGB1 on proliferation is modulated by redox state of the protein. Apoptosis can be induced through accumulation of extracellular HMGB1, where cytotoxicity is achieved through accumulation of HMGB1 in the mitochondria and depletion of mitochondrial DNA (Gdynia *et al*, 2010). The peculiar distribution of HMGB1 in testes can potentially be explained on the basis of its antibacterial activity via the TLR4 receptor to induce neutrophil extracellular traps (Tadie *et al*, 2012; Zetterström *et al*, 2006).

During senescence, irrespective of its type (replicative, p16 overexpression-induced, X irradiation or oncogene-induced), HMGB1 is evicted from nuclei and released in extracellular space (Davalos *et al*, 2013). The loss of HMGB1 in senescence can be considered a hallmark of senescence and not a cell-type specific event, as it was observed in several human fibroblasts, mammary and prostate epithelial cell types. Extracellular HMGB1 acts as a part of SASP signaling, however release of HMGB1 in senescence is an early event, before SASP occurrence and the released HMGB1 form is mostly soluble (Davalos *et al*, 2013). In concordance with this observation, downregulation of HMGB1 can induce senescence in mouse and human fibroblasts and human endothelial cells, corroborating the HMGB1 eviction as an inducing event in senescence program. Interestingly, overexpression of HMGB1 does not induce senescence entry bypassing, but it actually invokes senescence. This finding could be explained through notion that senescence in its essence is a stress-response program and previously observed cytotoxic effects of non-physiological accumulation of HMGB1 (Gdynia *et al*, 2010). The release of HMGB1 is p53 dependent and released HMGB1 in oxidized cytokine-inducing form binds to TLR4 and induces IL-6 production in target cells which in turn promotes inflammation. Any residual HMGB1 in nuclei of senescent cells seems to be important for inducing the expression of IL6, and subsequently the SASP, given that near absolute siRNA-mediated depletion of HMGB1 is enough to induce senescence but not the SASP. As a cancer and senescence-

associated biomarker, HMGB1 is also explored in the context of anticancer therapies. Here, it was shown to act via the STING/STAT/p21 axis which involves suppression of TRIM30a, a negative STING regulator (Lee *et al*, 2021).

Previous studies suggested that HMGB1 plays a functional role in induction and propagation of senescence. HMGB1 nuclear eviction is an early event initiating senescence program, suggesting that upstream acting mechanisms, such as conformational chromatin changes might influence relocalization of HMGB1 (Davalos *et al*, 2013). On the other hand, the wealth of data suggests significant role of HMGB1 as an architectural protein that regulates various aspects of cellular homeostasis including the transcriptional program. Assessing the positioning of HMGB1 on chromatin which could offer more insight into these two aspects of HMGB1 functions, had been proved to be challenging. The dynamic mobility and low retention time of HMGB1 on chromatin deters classical crosslinking approaches from capturing HMG-box factors bound to DNA (Teves *et al*, 2016). Therefore, alternative approaches are needed to obtain more insight into the contribution of HMGB1 in senescence.

3. Regulatory non-coding RNAs

Traditional postulate of molecular biology, where genetic information flow is unidirectional and starts with DNA encoding for RNA, which is eventually being translated into protein, has evolved in the recent years. The advancement of sequencing technologies revealed that albeit 2 percent of the genome encodes for proteins, more than 90 percent of the human genome is transcribed, suggesting that transcriptome may act beyond its protein coding role (Wilhelm *et al*, 2008). The identification of non-coding class of RNAs effectively divided the DNA transcription into coding and non-coding, where the non-coding transcriptome seems to play a significant role in gene regulation (**Fig. 9**). Realization of non-coding RNA abundance, variability, and presence across all the domains of life led to the concept of 'pervasive transcription'. Pervasive transcripts display features distinct from coding RNAs: they rarely have assigned function, often overlap or are a part of coding genes and they can be sense or antisense derived (Wade and Grainger, 2014). The lack of functional information and experimental verification has led to controversy when it comes to the pervasive transcription, and while the existence of such transcription is acknowledged, the question of its significance remains open.

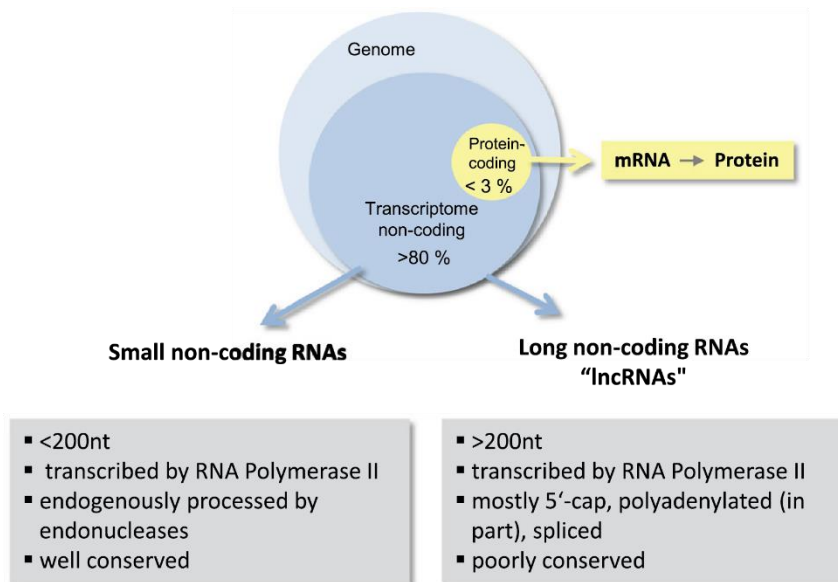


Figure 9. Non-coding RNAs represent the largest fraction of eukaryotic transcriptome

Non-coding transcriptome makes up for more than 80% of transcriptional output. Non-coding RNAs can be roughly divided into small (<200 nt) and long (>200 nt) RNAs. Reprinted from (Uchida and Dimmeler, 2015).

Nonetheless, several categories of non-coding RNAs are attributed to the non-coding transcriptome: microRNA, piwi-interacting RNA, small-interfering RNA and long non-coding RNA. Among these, long non-coding RNAs (lncRNA) are functionally and structurally most diverse class of non-coding RNAs (Kung *et al*, 2013). Broadly defined as relatively stable, longer than 200 nucleotide RNAs which are often poly-adenylated and spliced, but generally not translated (Iyer *et al*, 2015). lncRNAs can be transcribed from existing protein coding genes, promoters, enhancers or intergenic sequences (Marques & Ponting, 2009). lncRNAs are poorly conserved across evolution, where ~30% seems to be primate specific (Hezroni *et al*, 2015; Derrien *et al*, 2012) and they are expressed in cell and context specific manner (Clark & Mattick, 2011). The reports on the number of different lncRNA varies from ~16 000 to more than 100 000, depending on the annotation and experimental approach, however only few hundred lncRNAs are experimentally validated (Uszczynska-Ratajczak *et al*, 2018). Although the expression levels of lncRNA compared to protein coding genes is lower, the dynamic range of expression is higher for lncRNA and a greater proportion of lncRNAs is retained in the nucleus, especially in chromatin fraction (Derrien *et al*, 2012). As lncRNAs are transcribed by RNA Pol II, but with different RNA Pol II phosphorylation dynamic compared to protein coding genes, they are inefficiently spliced, delivering lncRNA transcripts that transiently accumulate at the sites of transcription but are rapidly degraded by nuclear exosome complex (Schlackow *et al*, 2017). However, there might be other means of lncRNA accumulation in the nucleus;

lncRNAs can contain RNA binding sites and associate with U1 snRNP (Yin *et al*, 2020), binding sites for recruitment of other nuclear factors (Azam *et al*, 2019) or alternative polyadenylation sites which regulate their subcellular localization (Xiang *et al*, 2014). Large fraction of lncRNAs is exported to the cytoplasm, where they are often bound by ribosomes (Carlevaro-Fita *et al*, 2016), which is another type of interaction that could lead to lncRNA degradation.

Preference for nuclear location, relative stability of lncRNA and their regulated degradation and export suggest a functional role for these non-coding RNAs. Indeed, a growing body of evidence implicates lncRNAs in variety of cellular processes ranging from transcriptional and posttranscriptional gene regulation, across scaffolding, immunity, cancer and others (Statello *et al*, 2021). Notable example includes one of the most studied lncRNA, Xist, involved in X-inactivation through direct interaction with X chromosome, whereby Xist coats the chromosome surface and induces transcriptional silencing through recruitment of polycomb repressive complex 2 (Clemson *et al*, 1996; Zhao *et al*, 2008). lncRNAs can also regulate gene expression by competing with protein-coding genes for transcriptional machinery, resulting in transcriptional silencing (Latos *et al*, 2012). An interesting example of lncRNA function in gene activation is p53-induced lncRNA named LED. Binding of LED to p53-enhancers induces local histone acetylation and production of enhancer RNAs, which in turn bring enhancer regions in proximity of their respective promoters to induce transcription (Léveillé *et al*, 2015). Regulatory actions of lncRNA are important in induction of key immunity genes prior to the inflammatory stimulus, as exemplified by gene-priming lncRNA UMLILO in monocytes, which facilitates deposition of active chromatin mark H3K4me3 at the promoters of several chemokine genes in a key event in immune gene priming in trained immunity (Fanucchi *et al*, 2019).

3.1 circRNAs are versatile class of lncRNAs

Tremendous development of non-coding transcriptome research in the past two decades had for the most part settled the question of non-coding RNA authenticity and, in many cases, their functionality. Regardless of the type and despite their variability, lncRNAs are detected as linear transcripts that can be relatively easily discovered in RNA-seq data. Relying on linearity as an inherent property of lncRNA resulted in overlooking another lncRNA type, circular RNAs (circRNAs). It has been long known that RNAs can assume a circular configuration, which is a typical genomic arrangement of RNA viruses. However, the first observation of circRNAs in eukaryotes, through electron microscopy, dates to 1979 (Hsu & Coca-Prados,

1979). In the following years, a handful of circRNAs were reported, such as that from the mouse *Sex-determining Region Y (Sry)* gene (Koopman *et al*, 1990). *Sry* produces a functional protein important for mouse testes development, but in adult mice the circRNA is the dominant product of *Sry* transcription and acts as a miRNA sponge (Hansen *et al*, 2013). Genome-wide detection of circRNAs developed with the advances and affordability of RNA sequencing, leading to an explosion in the numbers of detected circRNAs and the subsequent reports on their functionality. It soon became clear that circRNAs are ubiquitous constituents of human (Salzman *et al*, 2012; Jeck *et al*, 2013; Guo *et al*, 2014), mouse (Memczak *et al*, 2013; Rybak-Wolf *et al*, 2015), fruit fly (Salzman *et al*, 2013; Westholm *et al*, 2014), nematode (Memczak *et al*, 2013; Ivanov *et al*, 2015), yeast and plant cells (Wang *et al*, 2014). Moreover, circRNA expression patterns seem to be cell type- (Salzman *et al*, 2013), tissue- (Maass *et al*, 2017; Xia *et al*, 2017) and context-specific (Farooqi *et al*, 2021; Knupp & Miura, 2018; Venø *et al*, 2015). Unlike other lncRNAs, circRNAs seem to be highly conserved across evolution (Jeck & Sharpless, 2014; Wu *et al*, 2020). Taken together, compelling evidence suggests that circRNA expression is highly regulated, advocating for their functionality.

3.2 circRNA biogenesis

The circular conformation of this lncRNA type allows few peculiarities as regards their structure and biogenesis. Unlike linear RNAs decorated with 5' caps and 3' poly-A tails, circRNAs are covalently closed through association of their 5' and 3' ends, rendering them difficult to detect in conventional mRNA-seq approaches. In poly-A depleted RNA-seq datasets however, circRNA presence was observed as unexpected signal accumulation in various exonic or intronic regions, named excised exons/introns (Yang *et al*, 2011). Although excised exons/introns were thought to be a byproduct of splicing that gets rapidly degraded after transcription, it is now understood that these transcripts form through back-splicing and are stabilized through circularization (Salzman *et al*, 2012). Canonical mRNA splicing is a 5' to 3' directed process that together with other co-transcriptional events leads to transformation of pre-mRNA into mature mRNA transcript. Conversely, back-splicing follows the opposite direction, where downstream 5' donor splice site is joined to an upstream 3' acceptor site which results in covalently closed circular transcript (**Fig. 10**). An additional product of this event is the quickly degraded and alternatively spliced linear RNA, not carrying a circRNA sequence, fitting to the observed signal distribution of RNA-seq data (Jeck & Sharpless, 2014). Back-splicing is carried out through conventional splicing machinery, as demonstrated by the diminishing of back-

splicing events upon spliceosome inhibition with isoginkgetin leading to global depletion of circRNAs (Starke *et al*, 2015). Mutating canonical back-splicing sites also interferes with circRNA production, either through activation of cryptic splice sites or direct inhibition their biogenesis (Ashwal-Fluss *et al*, 2014). Whether canonical splicing happens first, resulting in linear transcript that is subject to back-splicing, or back-splicing drives formation of circRNA and linear transcript that is later canonically spliced, it is not clear, but both of the scenarios seem plausible *in vivo* (Jeck & Sharpless, 2014; Lasda & Parker, 2014).

Nevertheless, circRNA biogenesis is connected to canonical splicing machinery, although how the choice between canonical and back-splicing is regulated, remains elusive. CircRNAs are often less abundant than their linear counterparts, suggesting that the back-splicing might be less favorable compared to the canonical splicing. Indeed, expression of most circRNAs is lower than that of their linear counterparts (Jeck *et al*, 2013). However, in some cases circRNA expression is preferred over linear counterpart (Salzman *et al*, 2012). Various *cis*-elements and *trans*-factors could play a role in regulating the balance between two splicing mechanisms, by bringing the donor and acceptor site in close proximity (Zhang *et al*, 2014; Jeck *et al*, 2013; Ivanov *et al*, 2015; Ashwal-Fluss *et al*, 2014). Until now, three hypothetical models are proposed: intron-pairing driven, RNA binding protein (RBP)-mediated and lariat-driven circularization. In intron-pairing driven circularization model, flanking sequences of back spliced transcripts, which are often complementary to each other, lead to RNA pairing that enhances the production of circRNAs (Zhang *et al*, 2014); same principle is observed in canonically spliced linear transcript variants, which means that competition between canonical and back splicing remains in play, even when splicing is driven by complementary base pairing. Observations of trans-factor mediated circularization are sparse, but a notable example of such activity is the muscleblind (MBNL1) splicing factor, which is known to regulate adult splicing patterns. MBNL1 autoregulates its expression through binding to conserved RNA motifs, which are present in its own mRNA among the other transcripts. This binding facilitates the back splicing event rate of the MBNL1 transcript, offering an example of circRNA function in negative feedback loop of MBNL1 expression (Ashwal-Fluss *et al*, 2014). Similar protein-mediated promotion of circularization was shown for NF90/NF110 (Li *et al*, 2017) and quaking (Conn *et al*, 2015), while adenosine deaminase (Ivanov *et al*, 2015) and DEAH-box helicase 9 (Aktas *et al*, 2017) are recognized as negative modulators of back splicing through interference in joining donor and acceptor sites. One of the steps in canonical splicing is the formation of lariats, intermediate circular structures that connect acceptor and donor splicing sites. In some cases, lariats flanked by specific motifs can skip debranching and stabilize as intronic circRNAs (Yang *et al*, 2011; Zhang *et al*, 2013).

CircRNAs are a substrate for splicing themselves, in a process referred to as alternative circularization. In repeat element-rich genes, competition for RNA pairing which would drive splicing events can deliver multiple circRNAs from a single gene (Zhang *et al*, 2014). Alternatively, multiple circRNA variants can be produced through internal splicing, where exons and introns are included or excluded (Salzman *et al*, 2013). Enrichment of alternative circRNA isoforms is another cell type-specific feature indicating regulated circRNA production. CircRNAs can contain one or more exons and many multiexonic circRNAs are composed exclusively out of exonic sequences (Zhang *et al*, 2014). In some cases, introns are retained resulting in exon-intron circRNA (elciRNA; Li *et al*, 2015b), while some circRNAs are derived solely out of intron sequences (ciRNA; Zhang *et al*, 2013). elciRNAs and ciRNAs are often retained in nuclei whereas exonic circRNAs are dominantly cytoplasmic (Li *et al*, 2015d; Jeck *et al*, 2013; Memczak *et al*, 2013).

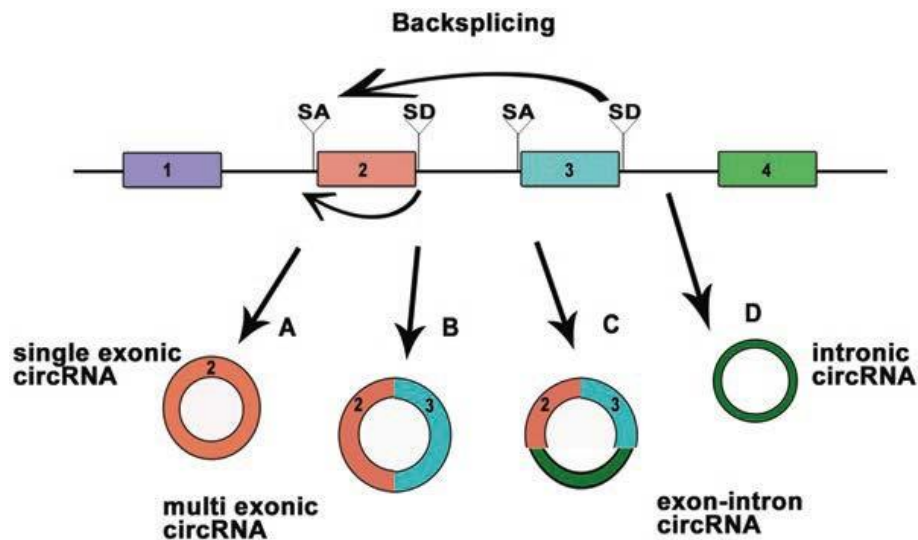


Figure 10. Overview of circRNA biogenesis.

circRNA is produced through back-splicing events, where downstream donor splicing site is connected to upstream acceptor site (top). Back-splicing gives rise to circRNAs of variable composition depending on the spliced region. circRNAs can contain one or more exons (A and B), exons and introns (C) or be completely intronic (D). Reprint from (Bose and Ain, 2018)

Given that most of circRNAs localize to the cytoplasm, it is reasonable to consider whether they are also translated. Early studies suggested that lack of 5'-cap in the circRNA and often incomplete open reading frame in their sequence make circRNA translation a highly unlikely event (Jeck *et al*, 2013; Guo *et al*, 2014). Canonical initiation of translation requires 5'-cap and poly-A tail, but alternative translation initiation mechanism exist, such as from internal ribosome entry sites (IRES). Insertion of IRES in circRNAs was demonstrated to be sufficient for translation of circRNA in vitro (Chen & Sarnow, 1995) and IRES-like

sequences are often present in circRNA. Very recent research efforts have demonstrated that IRES can be efficiently utilized for circRNA translation and few circRNA derived peptides were identified (Yang *et al*, 2018; Zhang *et al*, 2018; Legnini *et al*, 2017; Zheng *et al*, 2019). One such example is circFBXW7, shown to encode for 21-kDa protein which inhibits cell proliferation and whose knock-down promotes malignant phenotype (Yang *et al*, 2018). Interestingly, circ β -catenin is highly expressed circRNA in liver cancer that encodes for a variant of β -catenin protein. Translation of the protein is initiated at canonical start codon that is used in linear β -catenin RNA; however circularization of circ β -catenin creates a new stop codon at the back-splicing junction, which results in translation of a shorter peptide. This novel β -catenin isoform can bind the full-length β -catenin and prevent its phosphorylation and degradation, thereby facilitating the activation of the Wnt pathway (Liang *et al*, 2019). Another initiation alternative is N6-methyladenosines (m6A) present in the 5' untranslated region (UTR), where only one modified nucleotide seems to be enough to recruit translational machinery. To this date, no m6A initiated circRNA translation with identified peptide has been reported, but a recent study suggests that a large fraction of circRNAs carry this modification and associate with polysomes, therefore could be subjected to m6A initiated translation (Yang *et al*, 2017). On the other hand, m6A modification is implicated in circRNA degradation (Park *et al*, 2019), due to it being recognized by YTHDF2 protein which recruits endoribonuclease, as well as circRNA biogenesis (Zhou *et al*, 2019a), through promotion of back splicing instead of canonical one. Therefore, it is still unclear to what extent m6A enrichment in circRNAs plays a role in translation of these RNAs.

Turnover of linear RNAs in the cell is largely controlled by exonucleases, however covalently closed circular structures renders circRNAs resistant to the exonuclease cleavage. It is possible that circRNAs serve as a substrate for endonucleolytic activity, linearizing transcripts so that they are now further processed by conventional RNA degradation machinery (Guo *et al*, 2020; Hansen *et al*, 2011; Park *et al*, 2019; Liu *et al*, 2019). Another clearance mechanism for circRNAs is their export from the cells through exosomes, as it was recently discovered that circRNAs are selectively exported and significantly concentrated in exosomes derived from human liver cancer cells (Li *et al*, 2015c) and KRAS mutant colorectal cancer cell lines (Dou *et al*, 2016). How is this export regulated remains unclear; however, miRNA and RNA-binding proteins are suggested to play a role in selecting the circRNA species for the transport. Exosome export of circRNAs also raises a possibility of circRNAs acting as paracrine signaling molecules, but this remains to be further explored.

3.3 circRNAs in homeostasis and disease

Biological functions have until now only been investigated in a small fraction of identified circRNAs. Nonetheless, since circRNA expression seems to be strictly regulated in both a spatial and a temporal manner, and given the capacity of circRNAs to interact with other RNAs or proteins, the putative regulatory roles of circRNAs are many (**Fig. 11**). As circRNAs exhibit a potential to act through RNA or protein associations, they have found an important place in “pan-genomic” regulatory networks (Brackley *et al*, 2020), reinforcing the idea of multivalent small-scale interactions governing the large-scale cellular processes.

One of the first described molecular functions for circRNAs is miRNA sponging. miRNAs exert their role on a number of mRNAs by inducing their cleavage and degradation. Specific miRNA binding sites are also found in circRNAs, some of which are resistant to miRNA induced degradation, therefore circRNAs can sequester the miRNAs from available RNA pool and indirectly regulate mRNA stability. For example, ciRS-7 (also named CDR1as) is a circRNA abundantly expressed in neuronal tissues and likely involved in neuronal differentiation. It contains more than 70 conserved binding sites for miR-7 miRNA, acting as a decoy for miR-7 mediated RNA degradation (Hansen *et al*, 2013). ciRS-7 itself is resistant to degradation initiated by miR-7, and the two regulatory RNAs tend to be co-expressed. Perturbance in regulatory ciRS-7/miR-7 axis is observed in Alzheimer disease, leading to defects in ubiquitin-proteasome function and subsequent amyloid accumulation, outlining the contribution of circRNA regulation in disease (Zhao *et al*, 2016). Sponging function was demonstrated for Sry circRNA and miR-138 (Hansen *et al*, 2013), cir-ITCH and miR-7, miR-17 and miR-214 (Li *et al*, 2015a), and 25 other circRNAs (reviewed in Panda, 2018), while disruption of the sponging function is implicated in series of cancer types and other pathological conditions. CircRNA functions as a miRNA sponge is not typical of circRNAs; in order to significantly alter the availability of miRNAs, circRNAs need to carry multiple conserved miRNA binding sites in their sequence, which, more often than not, is not the case.

However, circRNAs can also act as decoys for proteins and regulate cellular processes at transcriptional and posttranscriptional level. Previously mentioned principle of MBNL1 autoregulation through circMbl binding (Ashwal-Fluss *et al*, 2014) was also observed in case of circPABPN1, where translation of PABPN1 is prevented due to sequestering of Hu-antigen R, a positive modulator of PABPN1 translation (Abdelmohsen *et al*, 2017). In another study, self-renewal capability of mouse hematopoietic stem cells was shown to rely on circRNA-protein interactions. Hematopoietic stem cell “stemness” is

preserved throughout the life of an organism in a dormant state until they encounter a pro-inflammatory signal and enter the cell cycle. Cytosolic DNA sensor that regulates production of such inflammatory signals is cyclic GMP-AMP synthase (cGAS) which is kept sequestered and inactive in dormant stem cells through binding of cia-cGAS circRNA (Xia *et al*, 2018). Therefore, downregulation of cia-cGAS leads to activation of cGAS, activation and decrease of hematopoietic stem cell pool.

Binding of multiple proteins to circRNAs exerts their function as molecular scaffolds. In human cardiomyocytes, Circ-Amotl1 facilitates PDK1-dependent phosphorylation of AKT1 in this manner (Zeng *et al*, 2017). Phosphorylated AKT is then translocated to nucleus as cell survival promoting factor. Another prominent example correlates circFoxo3 function with cellular senescence regulation. G1 to S cellular phase transition relies on phosphorylation of Rb by cyclin E/cyclin-dependent kinase 2 (CDK2) complex. CDK2 is in turn negatively regulated by p21 and this interaction is facilitated by circFoxo3 through formation of ternary complex that inhibits cyclin complex activity, thereby inducing the cell cycle arrest (Du *et al*, 2016). Interestingly, circFoxo3 is also upregulated in starvation and ROS induced stress conditions in mouse cardiomyocytes, promoting the cellular senescence through alternative action. Here, circFoxo3 was shown to bind the several cell cycle promoting proteins, sequestering them in cytoplasm, preventing their

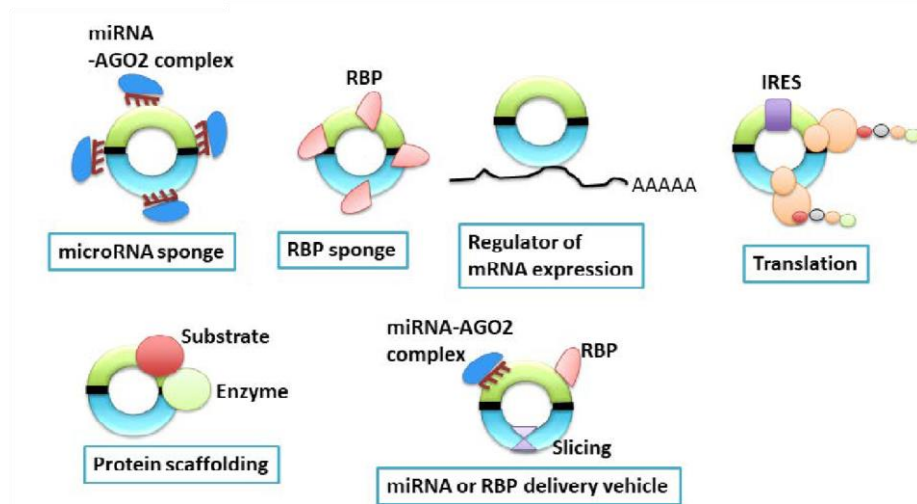


Figure 11. Overview of circRNA functions.

circRNAs can bind other RNA species or RNA binding proteins and act as molecular sponges and scaffolds. Through modulation of splicing events circRNAs contribute to regulation of mRNA expression. circRNAs can be translated into peptides and display indirect roles in cellular homeostasis. circRNAs can translocate through cellular environment and be exported, thereby functioning as delivery vehicle for associated proteins. Adapted from (Momen-Heravi and Bala, 2018)

translocation to nucleus and thus interfering with cellular stress response and inducing senescence (Du *et al*, 2017b).

CircRNAs are well established regulators of transcription, both at initiation and elongation stage, through interactions with transcriptional and splicing machinery. Notably, regulatory roles of circRNA in transcription have been demonstrated until now mostly for expression of circRNA parental genes and sporadically for others. For example, ci-ankrd52 was shown to accumulate at the ankrd52 transcription site where it interacts with RNA Pol II complex to stimulate its gene expression (Zhang *et al*, 2013). Recently, two elciRNAs were found to localize in nuclei and mediate RNA Pol II interaction with U1 small ribonucleoprotein to promote transcription of their parental genes (Li *et al*, 2015d). Another exonic circRNA, FECR1 originating from *FLI1*, was implicated in promotion of tumor metastasis (Chen *et al*, 2018b). FLI1 is ETS family transcription factor whose upregulation is associated with poor prognosis in cancer patients. FECR1 promotes tumor growth at the gene expression level, by upregulating the expression of FLI1 through recruitment of TET1 demethylase and subsequent demethylation of FLI1 region. Interestingly, FECR1 could also bind in trans at DNA methyltransferase DNMT1 promoter region to repress its transcription, but mechanism of this interaction is not known.

Chapters

Chapter 1 – RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis

My contribution is reflected in the following experimental/bioinformatic analysis as outlined below:

- Factory-seq experiments for DLD1 mAID-RPB1, with N. Übelmesser, and HCT116 TO2B^{-/-}-TOP2A-mAID cells
- ATAC-seq experiments for DLD1 mAID-RPB1 cells
- Computational analysis of all the Factory-seq, ATAC-seq, ChIP-seq and Cut&Tag-seq experiments

The following figure panels were prepared by Dr. A. Papantonis or myself with the data input from experiments prepared and/or analyzed by me:

- Figure 1.1B. Chip-seq and ATAC-seq profiles for experiments performed by N. Übelmesser and me
- Figure 1.3D. Cut&Tag signal coverage for CTCF, SMC1A and Rad21 at CTCF peaks
- Figure 1.3E. Top panel showing SMC1A and Rad21 at CTCF-bound, transcriptionally active/inactive and intergenic regions
- Figure 1.3F-G. Signal coverage for ATAC-seq at CTCF and transcription start sites
- Figure 1.S1D. PCA plot (left) and log₂(fold-change) distribution derived from differential gene expression analysis for DLD1 mAID-RPB1 Factory-seq data
- Figure 1.S1E. Exemplary coverage plots for DLD1 mAID-RPB1 Factory-seq data
- Figure 1.S1F. GSEA analysis for differentially expressed genes in DLD1 mAID-RPB1 Factory-seq data
- Figure 1.S4C. Same as 1.S1E and 1.S1F but for HCT116 TO2B^{-/-}-TOP2A-mAID Factory-seq

RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis

Shu Zhang^{1,#}, Nadine Übelmesser^{1,#}, Natasa Josipovic^{1,#}, Giada Forte², Johan A. Slotman³, Michael Chiang², Henrike Gothe⁴, Eduardo Gade Gusmao¹, Christian Becker⁵, Janine Altmüller⁵, Adriaan B. Houtsmuller³, Vassilis Roukos⁴, Kerstin S. Wendt⁶, Davide Marenduzzo², Argyris Papantonis^{1,7,*}

¹ Institute of Pathology, University Medical Center Göttingen, 37075 Göttingen, Germany.

² School of Physics and Astronomy, University of Edinburgh, EH9 3FD Edinburgh, United Kingdom.

³ Optical Imaging Centre, Erasmus Medical Center, 3015 GD Rotterdam, Netherlands.

⁴ Institute of Molecular Biology, 55128 Mainz, Germany.

⁵ Cologne Center for Genomics, University of Cologne, 50931 Cologne, Germany.

⁶ Department of Cell Biology, Erasmus Medical Center, 3015 GD Rotterdam, The Netherlands.

⁷ Center for Molecular Medicine Cologne, University of Cologne, 50931 Cologne, Germany.

These authors contributed equally to this work.

*Corresponding author: A.P

Abstract

Mammalian chromosomes are three-dimensional entities shaped by converging and opposing forces. Mitotic cell division induces drastic chromosome condensation, but following reentry into the G1 phase of the cell cycle, chromosomes reestablish their interphase organization. Here, we test the role of RNA Pol II in this transition using a cell line that allows its auxin-mediated degradation. *In situ* Hi-C showed that RNA Pol II is required for both compartment and loop establishment following mitosis. RNAPs often counteract loop extrusion and, in their absence, longer and more prominent loops arise. Evidence from chromatin binding, super-resolution imaging, and *in silico* modeling allude to these effects being a result of RNA Pol II-mediated cohesin loading upon G1 reentry. Our findings reconcile the role of RNA Pol II in gene expression with that in chromatin architecture.

Introduction

The evolution and expansion of chromosome conformation capture (3C) technologies (Tjalsma & Laat, 2020) has profoundly renewed our understanding of the spatial organization of eukaryotic chromosomes and of how it underlies their function and maintenance (Marchal *et al*, 2019; Szabo *et al*, 2019). It is now well accepted that chromosomes are dynamic entities (Hansen *et al*, 2018), and that their dynamics result from converging and opposing forces acting on chromatin (Rada-Iglesias *et al*, 2018). These forces include tethering to nuclear landmarks like lamina or the nucleolus (Canat *et al*, 2020), the interplay between transcription factor-bound *cis*-elements (Kim & Shendure, 2019), and the dynamic extrusion of loops via cohesin complexes (Rada-Iglesias *et al*, 2018; Nuebler *et al*, 2018).

Cohesin-extruded loops, almost invariably anchored at convergent CTCF-bound sites, are found along mammalian chromosomes representing a prominent feature of 3D genome organization. The combination of high-resolution Hi-C with acute and reversible degradation (Yesbolatova *et al*, 2019) of chromatin-organizing factors has shed light on loop emergence. CTCF degradation causes loss of insulation at thousands of topologically-associated domain (TAD) boundaries (Nora *et al*, 2019). Cohesin depletion leads to the elimination of essentially all CTCF-anchored loops (Rao *et al*, 2017; Schwarzer *et al*, 2017). Depletion of the cohesin-release factor WAPL promotes loop enlargement and aberrant looping by also engaging non-convergent CTCF-bound anchors (Haarhuis *et al*, 2017). These observations, together with the recently documented ability of cohesin to extrude loops *in vitro* (Davidson *et al*, 2019; Kim *et al*, 2019) and the finding that CTCF-STAG interactions protect cohesin from chromatin release (Li *et al*, 2020; Wutz *et al*, 2020), have crystalized a model for how architectural loops form and dissolve.

In addition to cohesin, another molecular motor known for its ability to translocate DNA is the RNA polymerase (Papantonis & Cook, 2011). However, its contribution to chromatin folding is still debated. Different lines of evidence point to a connection between RNA Pol II binding to chromatin and the differential formation of spatial interactions. To cite some recent examples, allele-specific Hi-C showed that the mouse inactive X chromosome lacks active/inactive compartments and TADs, which however form around “escapee” genes and in the active allele (Giorgetti *et al*, 2016); the transcriptional state of variably-sized domains across eukaryotes, from *C. elegans* and *D. melanogaster* to *A. thaliana* and mammals, is a robust predictor of interactions mapped via Hi-C and explains chromatin partitioning to a great extent (Ulianov *et al*, 2016; Rowley *et al*, 2017); and TAD emergence coincides with transcriptional activation in zygotes (Hug *et al*, 2017). Pharmacological abrogation of transcription compacts chromatin (Naughton *et*

al, 2013), weakens, but does not alleviate, TAD boundaries (Hug *et al*, 2017; Barutcu *et al*, 2019), and treating native (Brant *et al*, 2016) or fixed nuclei (Barutcu *et al*, 2019) with RNase does not affect TADs, but eliminates specific contacts. Single-nucleosome imaging upon acute RNA Pol II depletion showed that polymerases act to constrain and direct chromatin movement in 3D space (Nagashima *et al*, 2019), compatible with the idea of transcription-based chromatin organization.

In contrast, RNA Pol II and Mediator-complex components were found to be dispensable for bringing *cis*-elements into spatial proximity (El Khattabi *et al*, 2019) and inhibition of transcription in parallel with RAD21 reintroduction in cohesin-depleted cells did not affect loop reestablishment (Rao *et al*, 2017). Nevertheless, CTCF or cohesin depletion from mammalian cells had rather limited impact on gene expression (Nora *et al*, 2019; Rao *et al*, 2017), and upon cohesin elimination, a comparable number of loops formed on the basis of chromatin identity (Rao *et al*, 2017) or did not dissolve at all (Thiecke *et al*, 2020). Most recently, Micro-C, a sub-kbp Hi-C variant, unveiled thousands of fine-scale loops connecting transcriptionally-active loci in mouse and human cells, often without association to CTCF/cohesin (Hsieh *et al*, 2020). Thus, the direct effects of active RNAPs on chromatin folding remain unclear.

On top of its potentially direct effects, RNAPs and the act of transcription may remodel genome folding via interplay with cohesin-CTCF complexes. For example, transcription can relocate cohesin by many kilobases (Busslinger *et al*, 2017). Such transcription-mediated displacement can even disrupt prominent CTCF loops and rewire spatial interactions (Heinz *et al*, 2018; Olan *et al*, 2020). In addition, RNAPs are essential for domain formation and often counteracted by condensin complexes (Rowley *et al*, 2019; Brandão *et al*, 2019). This and other data highlight the need to dissect and reconcile the contribution of RNA Pol II to chromatin organization. To this end, and as pharmacological inhibition of RNAPs is inefficient, we exploited a human cell line that allows rapid and reversible RNA Pol II depletion (Nagashima *et al*, 2019). We combined *in situ* Hi-C and super-resolution imaging of RNA Pol II with *in silico* models to disentangle the role of RNA Pol II in gene expression from that in genome architecture to reconcile the aforementioned observations.

Results

Acute RNA Pol II depletion impacts loop-level interphase chromatin folding

RNA Pol II is essential for cell viability, so its depletion may only be transient. Thus, we exploited a human DLD-1 colorectal cancer line, in which the largest RNA Pol II subunit, RPB1, is N-terminally tagged with a

mini-AID domain. This allows for its acute and reversible degradation upon addition of auxin (and of doxycycline to activate the plant ubiquitin ligase TIR1 recognizing this mAID domain; see refs Yesbolatova *et al*, 2019; Brant *et al*, 2016). In our hands, 2 h of dox/auxin treatment reduce RNA Pol II protein levels by >60%, while 14 h of treatment result in >80% degradation as assessed by western blotting – without affecting RNA Pol I or RNA Pol III levels (**Figs 1.1A** and **Appendix Fig. 1.S1A**). Degradation is less impactful on chromatin-embedded RNAPs (e.g., those marked by phosphorylated Ser5 residues in their C-terminal domain). Washing out auxin in the presence of its competitive inhibitor, auxinole, largely restores RNA Pol II-Ser5 levels (**Appendix Fig. 1.S1A**), suggesting that soluble (non-phosphorylated) RNA Pol II is more susceptible to degradation and that any residual polymerases will be chromatin bound (as was the case for CTCF-mAID; see ref. Nora *et al*, 2019). Thus, quantitative RNA Pol II removal can be achieved via this system in order to assess its contribution to genome folding.

To further characterize this line, we performed ChIP-seq using an antibody targeting the mClover tag in RPB1. Compared to public RNA Pol II ChIP-seq data from DLD-1 (GEO: GSM2769059), mClover-tagged polymerases occupied the same positions and could be depleted from chromatin genome-wide upon auxin treatment (**Fig. 1.1B**). Polymerase degradation was accompanied by a strong decrease in chromatin accessibility at TSSs (**Fig. 1.1B**), similar to that recently seen using mESCs (Jiang *et al*, 2020). We also queried the H3K27ac (marking active chromatin) and H3K27me3 histone modifications (marking facultative heterochromatin). Upon RNA Pol II depletion, significant H3K27ac reduction was observed concomitant with increased H3K27me3 levels (**Appendix Fig. 1.S1B,C**). Last, we monitored changes in nascent RNA levels using “factory” RNA-seq (Caudron-Herger *et al*, 2015). Control and auxin-treated samples separated well in PCA plots, with ~1500 genes changing their transcription levels significantly (**Appendix Fig. 1.S1D,E**). Of these, >90% were downregulated and mainly involved in chromatin assembly and gene expression regulation (**Appendix Fig. 1.S1F**).

We next asked whether the spatial organization of interphase chromatin is also altered upon RNA Pol II depletion. We applied *in situ* Hi-C to G1-sorted DLD1-RPB1-mAID cells treated or not with auxin for 14 h or to cells in which auxin was complemented by triptolide treatment (an inhibitor abrogating transcriptional initiation to further enhance RNA Pol II degradation; see ref. (Wang *et al*, 2011). Using G1 cells removes heterogeneity arising from S-/G2-phase cells to generate Hi-C maps of greater detail (Sati *et al*, 2020). Following data analysis, we saw only marginal differences in A-/B-compartments (**Fig. 1.1C,D**). TADs also showed only mild disruptions (**Fig. 1.1E,F**), with <20% of the 4,110 identified in control Hi-C data changing in cells lacking RNA Pol II. 227 new TADs could be detected in auxin-/triptolide-treated cells and displayed

reinforced insulation at their boundaries (Fig. 1.1F). Average contact profiles in/around TADs, revealed stronger definition of their borders at the expense of intra-TAD interactions (Fig. 1.1G).

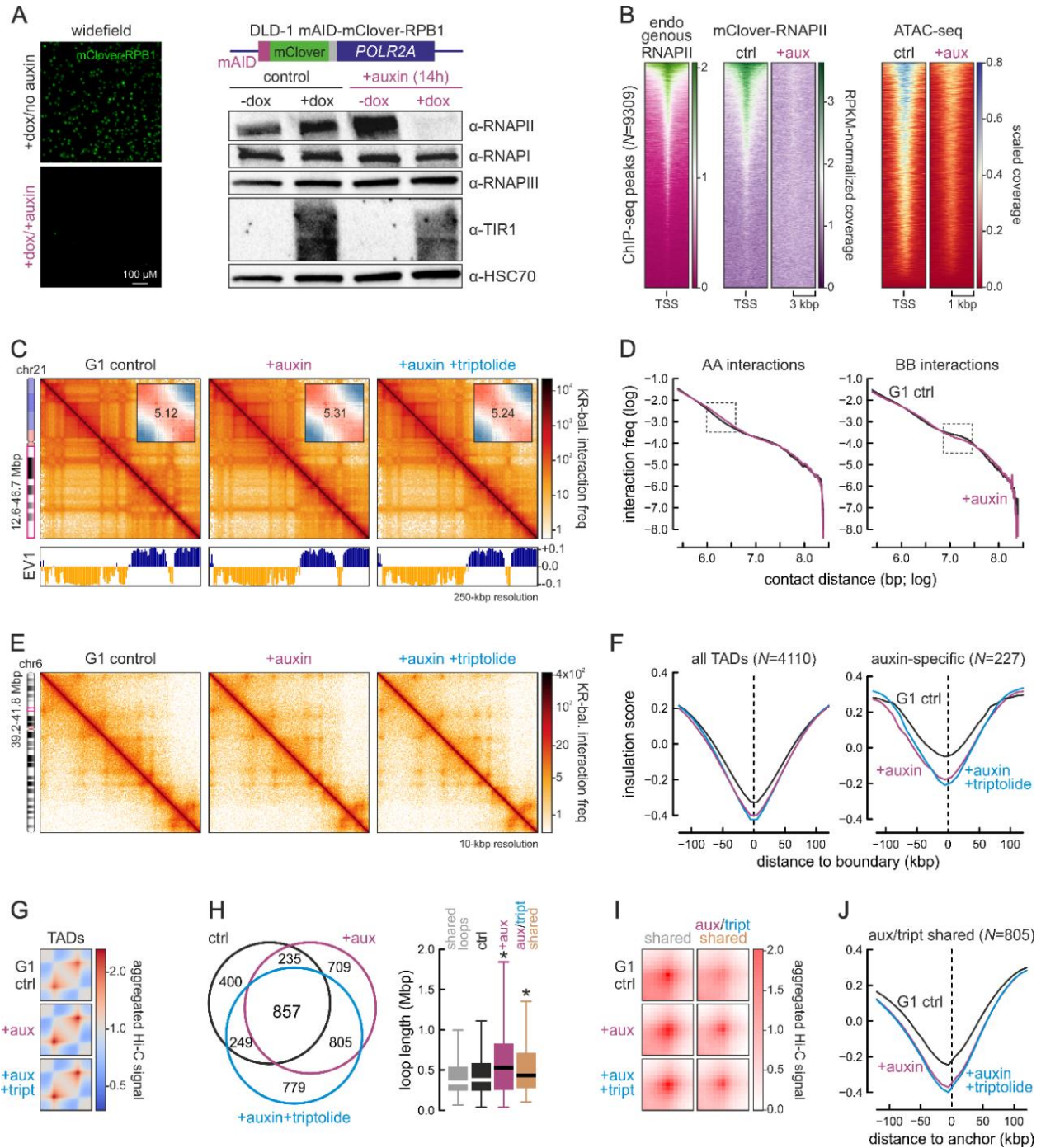


Figure 1.1. Effects of acute RNA Pol II degradation on interphase chromatin folding.

(A) *Left*: DLD1-mAID-RPB1 cells lose mClover signal fused to the large RNA Pol II subunit (schematic) upon doxycycline/auxin treatment for 14 h. *Right*: Selective TIR1-mediated RNA Pol II degradation confirmed by Western blotting; HSC70 levels provide a loading control. (B) Heatmaps showing mClover-RNA Pol II ChIP-seq signal (*middle*) overlapping positions bound by RNA Pol II in parental DLD-1 (*left*) and its loss upon auxin treatment concomitant with decreased chromatin accessibility (*right*). (C) Exemplary Hi-C maps of chr21 from G1-sorted control (*left*), 14-h auxin-treated (*middle*), and auxin plus triptolide-treated cells (*right*) at 250-kbp resolution aligned to first eigenvector values (EV1; below). Insets: saddle plots showing A/B-compartment insulation. (D) Decay plots showing Hi-C interaction frequency amongst A- (*left*) or B-compartments (*right*) as a function of genomic distance in control (*black line*) or auxin-treated cells (purple line). Dashed rectangles indicate distances at which the two lines deviate most. (E) Exemplary Hi-C maps of chr6 from control (*left*), auxin-treated (*middle*), and auxin+triptolide-treated cells (*right*) at 10-kbp resolution. (F) Line plots showing mean insulation from control (*black line*), auxin-treated (*purple line*), and auxin+triptolide-treated cells (*blue line*) in the 240 kbp around all (*left*) or degron-specific TAD boundaries (*right*). The number of TADs queried (N) is indicated. (G) Heatmaps showing aggregated TAD-level interactions in control (*top*), auxin-treated (*middle*), and auxin-/triptolide-treated cells (*bottom*). (H) *Left*: Venn diagram showing shared and unique loops to control (*black*), auxin-treated (*purple*), and auxin+triptolide-treated Hi-C (*blue line*). *Right*: Loop lengths displayed as boxplots (*right*). *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (I) Aggregate plots showing mean Hi-C signal for the loop groups in panel I. (J) As in panel F, but for loop anchors shared by auxin-treated and auxin+triptolide-treated cells.

Our data so far agreed well with recent observations in mESCs (Jiang *et al*, 2020). However, we also detected >1,500 loops exclusively in auxin-treated cells. These were significantly larger than those in control cells or than those shared between conditions (Fig. 1.1H,I). Of these, 805 loops shared by auxin- and auxin-/triptolide-treated cells were also larger and displayed increased insulation at their anchors (Fig. 1.1H-J). To understand the emergence of these stronger loops, we calculated the cumulative nascent RNA expression levels within the loop domains they form. Compared to all or shared loops from control cells, these 805 were significantly more associated with top-quantile loop domains (Appendix Fig. 1.S1G,H), suggesting that highly active RNA Pol II can counteract loop extrusion and its removal leads to stronger loop extrusion. However, although more “stripes” were detected in RNA Pol II-depleted Hi-C (indicative of loop extrusion; see ref. Vian *et al*, 2018), these were on average shorter (Appendix Fig. 1.S1I). In addition, almost 10% of these 805 loops also associated with zero-expression domains, but were still enhanced upon RNA Pol II depletion (Appendix Fig. 1.S1G,H). We attribute this to Polycomb-mediated interactions that become accentuated upon RNA Pol II depletion. This is in agreement with the elevated H3K27me3 levels (Appendix Fig. 1.S1C), the enrichment of H3K27me3 ChIP-seq signal at loop anchors (Appendix Fig. 1.S1J,K), as well as with recent literature (Rhodes *et al*, 2020).

Finally, note that analysis of Hi-C data from mAID-RPB1 cells treated with auxin for 2 h (where 60-70% RPB1 is degraded; Appendix Fig. 1.S1A), did not reveal changes at any level of spatial genome organization

(Appendix Fig. 1.S2A-J). Together, our data reveal subtle yet discernible effects at the levels of TAD and loop organization occurring upon acute RNA Pol II depletion in interphase.

Reestablishment of spatial chromatin organization after mitosis requires RNA Pol II

Since RNA Pol II depletion did not dramatically affect interphase chromatin organization in asynchronized cell populations, we hypothesized that it may be implicated in reestablishing chromatin folding upon exit from mitosis. This was based on two observations. First, on the detailed description of chromatin refolding dynamics in the mitosis-to-G1 transition, where contacts among *cis*-elements form early and rapidly, often not related to CTCF/cohesin (Abramo *et al*, 2019; Zhang *et al*, 2019). Second, on the fact that, early in this transition, >50% of all active enhancers and genes exhibit a strong spike in transcription (Hsiung *et al*, 2016).

To study chromatin refolding following mitotic exit, we synchronized mAID-RPB1 cells at the G2/M checkpoint using the CDK1 inhibitor RO3306 (blocking ~90% cells in G2; Fig. 1.2A,B), before releasing them via mitosis into G1. 6 h after washing out the inhibitor, >70% cells reentered G1 and were collected by FACS (Fig. 1.2A,B). RNA Pol II degradation, initiated by adding auxin to cells arrested in G2, was maintained throughout mitosis and G1 reentry without compromising progression past early G1 (as also exemplified by cell cycle markers; Appendix Fig. 1.S3A). RNA Pol II degradation was confirmed by fractionation and whole-cell blots (Fig. 1.2C and Appendix Fig. 1.S3B), immunodetection of Ser5-phosphorylated RNAPs, and EUTP-labeling of nascent RNA (Appendix Fig. 1.S3C,E). Like in asynchronous cells, the decrease in H3K27ac levels was accompanied by increased H3K27me3 levels in G1-reentry cells depleted of RNA Pol II (Appendix Fig. 1.S3C,D). At the same time, the levels of abundant SWI/SNF chromatin remodeler subunits in chromatin were altered, but CTCF incorporation was largely unchanged (Fig. 1.2C).

We next performed Hi-C on G1-reentry cells treated or not with auxin. We obtained >740 million and >1 billion Hi-C contacts from control and auxin-treated samples, respectively. Our first observation was that RNA Pol II-depleted cells showed increased inter-chromosomal contacts at the expense of intra-chromosomal ones (Appendix Fig. 1.S3F,G) also confirmed by high throughput 3D-DNA FISH (Appendix Fig. 1.S3H). At the same time, compartment boundaries were markedly blurred (Fig. 1.2D and Appendix Fig. 1.S3I), with interactions at distances of >1 Mbp between A- and B-compartment segments becoming stronger (Fig. 1.2E).

Loss of interactions at the TAD scale (<1 Mbp) in auxin-treated cells (Appendix Fig. 1.S3J) led us to analyze 10-kbp resolution Hi-C maps. There, we observed strong and widespread erosion of domain structure, local insulation and loop formation (Fig. 1.2F,G and Appendix Fig. 1.S3K). Insulation was weakened across all ~4,500 TADs identified in control cells, and even more in the ~10% of TAD boundaries

that were not reestablished in the absence of RNA Pol II (**Fig. 1.2H**). RNAP-depleted reentry cells also had fewer and larger TADs than control cells (**Fig. 1.2I**), indicative of boundary collapse and TAD merging. These effects are in line with RNA Pol II and active histone mark enrichment at TAD boundaries (**Fig. 1.2J**).

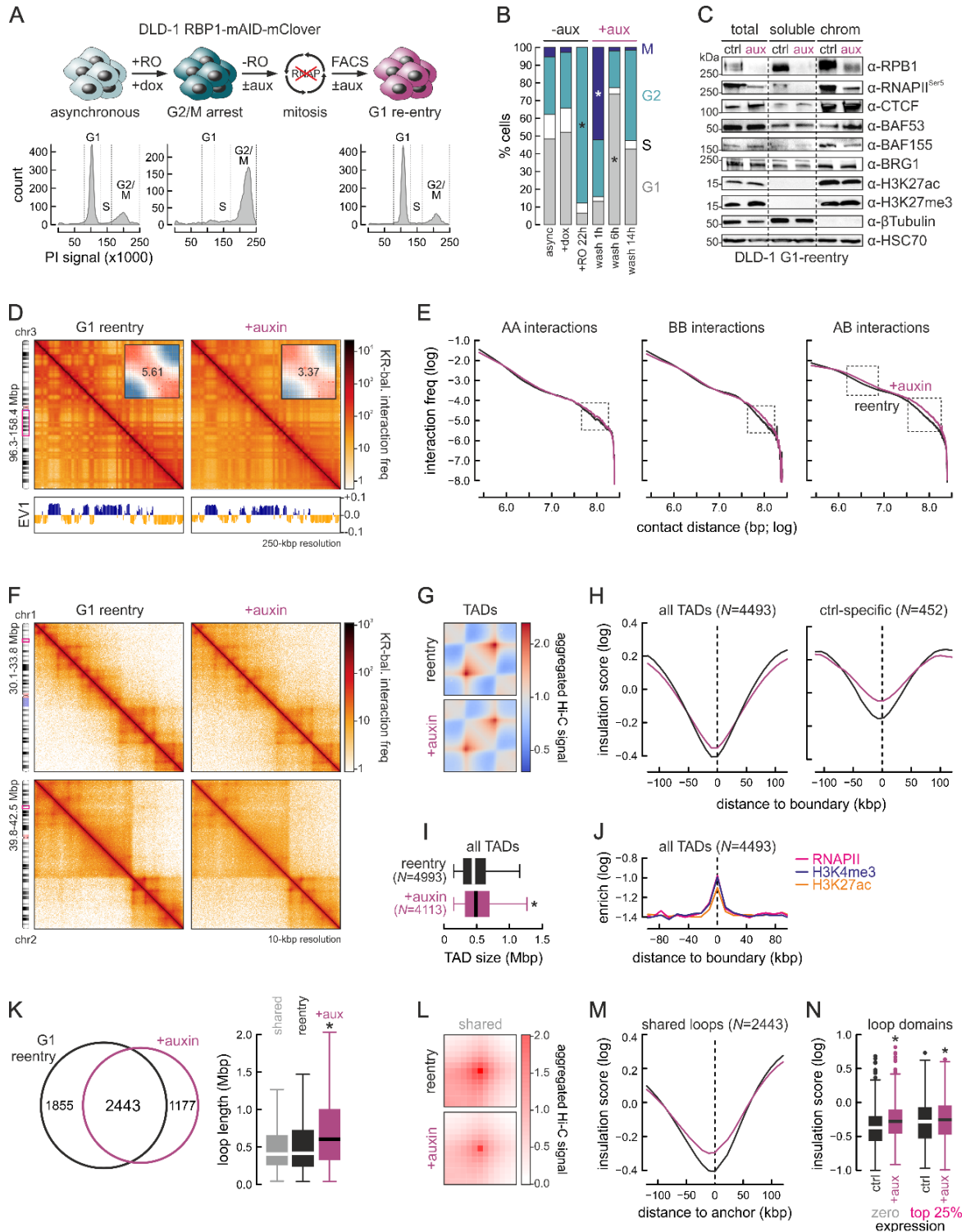


Figure 1.2. RNA Pol II contribution to genome refolding following exit from mitosis.

(A) Top: Experimental scheme for DLD1-mAID-RBP1 cell synchronization. Bottom: Propidium iodide FACS profiles at each step. (B) Bar plots quantifying the percent of cells in each cell cycle phase from the procedure in panel A. *: significantly different; $P < 0.01$, Fischer's exact test. (C) Fractionation western blots showing changes in chromatin-bound RNA Pol II, chromatin remodelers and histone mark levels; HSC70 provides a control. (D) Exemplary Hi-C maps of chr3 from control (*left*) and auxin-treated reentry cells (*right*) at 250-kbp resolution aligned to first eigenvector values (EV1; *below*). Insets: saddle plots showing loss of A/B-compartment insulation. (E) Decay plots showing Hi-C interaction frequency between A- (*left*), B- (*middle*) or A/B-compartments (*right*) as a function of genomic distance (log) in control (*black line*) and auxin-treated cells (*purple line*). Rectangles indicate distances where the two lines deviate the most. (F) Exemplary Hi-C maps of chr1 and 2 subregions from control (*left*) and auxin-treated cells (*right*) at 10-kbp resolution. (G) Heatmaps showing aggregated TAD-level interactions in control (*top*) and auxin-treated cells (*bottom*). (H) Line plots showing mean insulation from control (*black line*) and auxin-treated cells (*purple line*) in the 240 kbp around all (*left*) or control-specific TAD boundaries (*right*). The number of TAD boundaries queried (N) is indicated. (I) Boxplots showing TAD sizes in control (*black*) and auxin-treated cells (*purple*). *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (J) Line plot showing average RNA Pol II (*magenta*; GSM2769059), H3K4me3 (*blue*; GSM2283764) and H3K27ac ChIP-seq signal (*orange*; GSM2037784) in the 100 kbp around TAD boundaries from control cells. (K) *Left*: Venn diagram showing shared and unique loops between control (*black*) and auxin-treated cells. *Right*: Loop lengths displayed as boxplots. *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (L) Aggregate plots showing decreasing Hi-C signal for shared loops from panel J. (M) As in panel H, but for loop anchors shared between control and auxin-treated cells. (N) Boxplots showing changes in insulation between control (*black*) and auxin-treated cells (*purple*) at loop domains harboring zero or top-quantile nascent RNA levels. *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test.

At the loop level, ~1,900 were essentially lost from RNA Pol II-depleted cells concomitantly with the emergence >1,150 new and significantly longer loops (**Fig. 1.2K**). Curiously, these longer loops emerged at sites of existing insulation, whereas loops weakened in RNAP-depleted cells showed reduced insulation (**Appendix Fig. 1.S3L,M**). The 2,443 loops detected in both control and auxin-treated reentry cells were weakened in the absence of RNA Pol II, while also displaying reduced insulation at their anchors (**Fig. 1.2K-M**). Again, we compared loop domains of no gene expression to those harboring top-quantile nascent RNA levels. We found that insulation was significantly weakened in both cases (**Fig. 1.2N**). Finally, looking at stripe formation, we found both accentuated and dissolved ones in RNAP-depleted cells (**Appendix Fig. 1.S3N,O**).

In summary, our data suggest that RNA Pol II is implicated in reestablishing both higher-order and fine-scale chromatin folding following exit from mitosis, and its depletion compromises loop formation. Critically, the folding changes that follow polymerase depletion do not simply reflect structures of an earlier G1 time-point (e.g., compared to data from ref. (Zhang *et al*, 2019), but rather compromised refolding.

Topoisomerase II depletion does not affect G1-reentry chromatin folding

Transcription enforces supercoiling onto DNA and topoisomerase I (TOP1) is stimulated by RNA Pol II to resolve supercoils during elongation. However, TOP1 binding alongside initiating polymerases at TSSs was not matched by high TOP1 activity (Baranello *et al*, 2016). In contrast, topoisomerase II has been linked to chromatin organization along the cell cycle and to transcription, with TOP2A affecting RNA Pol II kinetics (Thakurela *et al*, 2013) and marking its pausing sites (Singh *et al*, 2020). Moreover, TOP2B flanks TAD boundaries in human cells alongside CTCF/cohesin complexes to confine RNA Pol II (Uusküla-Reimand *et al*, 2016) and preserves domain boundaries in yeast (Achar *et al*, 2020).

Given that no elongating RNAPs remain in auxin-treated cells, and that TOP2A-mAID cells prolong, but do not conclude mitosis (Nielsen *et al*, 2020), we asked whether TOP2 depletion from G1-reentry cells explains the effects we observe in RNA Pol II-depleted cells. To this end, we exploited another colorectal cancer line, HCT116, carrying or not a full knockout of the *TOP2B* gene and homozygously expressing mAID-tagged TOP2A. We verified >70% auxin-induced depletion of TOP2A in TOP2B-knockout cells, and applied the same synchronization and FACS sorting scheme as before to obtain G1-reentry cells (**Appendix Fig. 1.S4A,B**). Using “factory” RNA-seq to compare wild-type cells with those lacking both TOP2A and -B, <400 genes, mostly linked to cell cycle control and cell morphogenesis, were affected by TOP2 elimination (**Appendix Fig. 1.S4C**).

Hi-C performed on G1-reentry cells carrying or not TOP2A/B activity revealed marginal changes across all scales of chromatin organization. Compartments were not affected, interactions remained unchanged irrespective of distance, and no increase in *trans* contacts was seen (**Appendix Fig. 1.S4D-F**). Negligible changes to TAD boundary insulation were observed, and the mean size of TOP2A/B-depleted TADs did not differ from that in control cells (**Appendix Fig. 1.S4G-J**). Finally, although ~600 loops were lost or gained upon TOP2-depletion and condition-specific loops were again larger (**Appendix Fig. 1.S3K**), the increase/reduction in Hi-C signal at these loops was significantly less than that recorded upon RNA Pol II depletion and not followed by changes in insulation at their anchors (**Appendix Fig. 1.S4L,M**). In summary, these data suggest that the effects inflicted on chromatin refolding by RNA Pol II degradation cannot be recapitulated by TOP2A/B depletion, so must rather be polymerase-centric.

RNA Pol II removal compromises cohesin chromatin reloading and loop formation

Hi-C data from G1-reentry cells depleted of RNA Pol II clearly demonstrate A/B-compartment mixing, TAD erosion, and differential loss/gain of loops. Given that loop formation relies on the loading and DNA

extrusion by cohesin complexes ending up at CTCF-marked anchors (Nora *et al*, 2019; Rao *et al*, 2017; Schwarzer *et al*, 2017; Haarhuis *et al*, 2017; Li *et al*, 2020), we examined how the levels of CTCF/cohesin subunits change in reentry cells following auxin treatment. Fractionation western blots showed little fluctuation in CTCF, SMC1A or Rad21 levels on chromatin, which was confirmed by quantification of RAD21 levels in individual cells using immunofluorescence (**Fig. 1.3A,B**). However, the chromatin-bound levels of the two cohesin loaders, NIPBL and MAU2, were markedly reduced, as were the levels of the factor responsible for cohesin unloading, WAPL (concomitant with an increased in its soluble pool titers; **Fig. 1.3A**). Critically, this is not due to general downregulation of these proteins, because western blots showed that TIR1 activation to degrade RNA Pol II does not change their abundance in G1-reentry or G2/M-arrested cells (**Appendix Fig. 1.S5A,B**).

To understand whether our findings are due to changes in NIPBL and CTCF nuclear distribution, we performed super-resolution localizations of these factors. Dual-color *d*STORM in control and auxin-treated G1-reentry cells led to the following observations. First, NIPBL localizes in clusters of smaller average size upon RNA Pol II depletion. At the same time, we also observed more localizations in extended and deformed clusters (**Fig. 1.3C**), exemplified by the shift in “eccentricity” of NIPBL clusters from 0.54 in control to 0.69 in auxin-treated cells (eccentricity of 0 refers to a perfect circle, while eccentricity of 1 to a line). Second, CTCF clusters do not change as regards their mean size, but 50% of all CTCF clusters in control cells lie $<129 \text{ px}^2$, while in auxin-treated cells 50% lie $<82 \text{ px}^2$ (**Fig. 1.3C**). Such an increased population of smaller CTCF clusters was also observed via *d*STORM upon cohesin removal in Rad21-mAID cells (Casa *et al*, 2020) occurring as a result of loop collapse. Last, NIPBL distribution relative to CTCF also changed significantly in the absence of RNA Pol II. The median separation between NIPBL and its nearest CTCF cluster was reduced from 353 to 281 nm, with the largest recorded distance dropping from >1200 to 825 nm (**Fig. 1.3C**).

These results, and the changed NIPBL/MAU2/WAPL levels on chromatin, hint to aberrant cohesin loading to (and most likely unloading from) chromatin in the absence of RNA Pol II, and predict that less cohesin will end up at CTCF-bound sites. To test this prediction, we generated SMC1A and Rad21 CUT&Tag data in control and auxin-treated reentry cells and indeed found cohesin signal significantly reduced at CTCF sites genome-wide (**Fig. 1.3D**). We quantified the fraction of scaled CUT&Tag signal falling into CTCF-bound regions, active or inactive TSSs, and non-RNA Pol II-associated intergenic space. We found that the reduced fraction of reads at CTCF sites was accompanied by a 20-30% increase in signal mapping to inactive TSSs in auxin-treated reentry cells, and by a $>10\%$ increase in widespread intergenic signal (**Fig. 1.3E**). It follows that such general reduction of cohesin occupancy at CTCF sites impairs loop formation genome-

wide (Fig. 1.3E). Notably, these changes occurred despite no reduction in CTCF-proximal accessibility as judged by ATAC-seq (Fig. 1.3F).

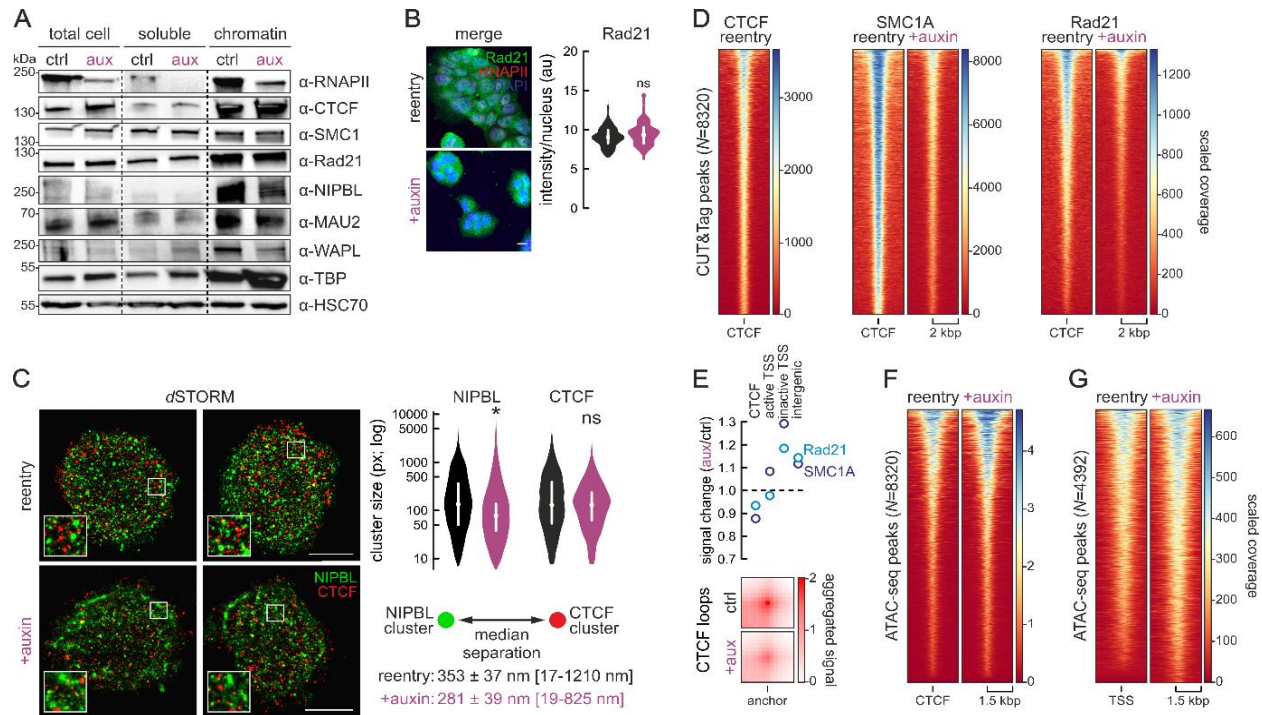


Figure 1.3. RNA Pol II degradation affects CTCF/NIPBL distribution and cohesin loading following mitosis.

(A) Fractionation blots showing changes in chromatin-bound RNA Pol II, cohesin loaders NIPBL and MAU2, the WAPL unloader, and TBP; HSC70 provides a loading control. (B) RAD21 and RNA Pol II immunofluorescence in untreated (*top*) or auxin-treated reentry cells (*bottom*) and signal quantification (*bean plots*). Bar: 5 μm. (C) Left: Rendering of 3D-STORM localizations for NIPBL and CTCF from control (*top row*) and auxin-treated reentry cells (*bottom row*). Bar: 5 μm. Right: Bean plots showing changes in NIPBL and CTCF cluster sizes. *Bottom right*: Changes in separation between the nearest NIPBL/CTCF clusters (smallest and largest distances shown in square brackets). *: significantly different; P<0.01, Wilcoxon-Mann-Whitney test. (D) Heatmaps showing SMC1A (*middle*) and Rad21 CUT&Tag signal (*right*) in control and auxin-treated G1-reentry cells at CTCF positions (*left*). (E) *Top*: Plot showing changes in SMC1A (*dark blue circles*) and Rad21 CUT&Tag signal (*light blue circles*) assigned to CTCF-bound, active/inactive TSSs or intergenic regions in auxin-treated compared to control cells. *Bottom*: Plots showing aggregate Hi-C signal for loop categories anchored at the CTCF sites in panel D. (F) As in panel D, but using ATAC-seq around CTCF-proximal SMC1A/Rad21-bound positions. (G) As in panel F, but for active TSSs.

Finally, one could assume that cohesin loading at TSSs (Busslinger *et al*, 2017; Pelham-Webb *et al*, 2021) simply results from them being rendered accessible by active RNA Pol II. Hence, reduced accessibility would readily explain compromised loading. To our surprise, ATAC-seq signal in RNAP-depleted TSSs rather increased (Fig. 1.3G), as did TBP levels on chromatin (Fig. 1.3A). Looking at CUT&Tag signal at >4,000 RNA Pol II/SMC1A co-occupied TSS, it is drastically reduced upon auxin treatment of reentry cells (Appendix Fig.

1.55C,D). This argues in favor of RNAPs recruiting cohesin loaders and unloaders to these sites, while setting up TSS architecture in G1-reentry cells likely relies on pioneer factors preceding the polymerase (like TBP). In fact, we used co-immunoprecipitations to show that RNA Pol II directly interacts with WAPL (**Appendix Fig. 1.55E**), and that NIPBL co-purifies with RNA Pol II in G1-reentry cells (**Appendix Fig. 1.55F**).

Modeling dissects RNAP contribution to loop extrusion

To dissect the connection between RNA Pol II and cohesin reloading onto chromatin, we turned to *in silico* modeling of chromatin folding. This allowed us to test scenarios that would be challenging to address experimentally. First, we performed 3D chromatin folding simulations using the established HiP-HoP model (Buckle *et al*, 2018) that accounts for the heteromorphic nature of chromatin, and incorporates transcription factor binding and loop extrusion. We modeled a 10-Mbp region from HUVEC chr14 for which gene expression, histone mark, and CTCF positioning ENCODE data are available (www.encodeproject.org).

Control chromatin folding was simulated by assuming that most cohesin loading (90%) occurs at RNA Pol II-occupied TSSs (with 10% loading randomly). Experimentally-defined cohesin residence times on DNA (~20 min; see ref. Hansen *et al*, 2017) were incorporated into the model. Following multiple iterations, our model produced a mean contact map resembling Hi-C data (**Fig. 1.4A,B**). To simulate chromatin folding following RNA Pol II degradation, we eliminated loading at promoters, and only allowed random loading (consistent with low efficacy NIPBL-independent cohesin loading *in vitro*) (Davidson *et al*, 2016). As a result, four major effects were observed. First, a general weakening of interactions and domain insulation across the 10 Mbp modeled (**Fig. 1.4A,B**), similar to what we saw using Hi-C (**Fig. 1.2F**). Second, individual models of the fiber displayed obvious unfolding (**Fig. 1.4B**), likely consistent with the increase in *trans* interactions in our data (**Appendix Fig. 1.53F-H**). Third, reduced cohesin occupancy at CTCF sites (**Fig. 1.4A**), consistent with our CUT&Tag data (**Fig. 1.3D** and **Fig. 1.55C,D**). Fourth, markedly weakened loop formation, but with larger loop sizes (**Fig. 1.4C,D**) matching our experimental results (**Figs 1.2K,L** and **1.3E**). Thus, our modeling suggests that inability to load cohesin via RNA Pol II-bound sites suffices for explaining the major chromatin folding differences observed experimentally.

To interrogate the interplay between RNA Pol II and loop-extruding cohesin directly, we performed 1D simulations of minimal composition. We modeled a 3-Mbp region of HUVEC chr14 (ch14:53-56 Mbp, hg19) as a coarse-grained fiber carrying CTCF at the appropriate positions, as well as RNAPs transcribing genes in the correct orientation. As before, cohesin was predominantly loaded at RNA Pol II-bound TSSs in the control scenario, but only randomly in the “degron” model. First, we observed formation of loops and intricate domain compartmentalization under control settings, despite having only two activities operating

on the fiber (Fig. 1.4E). Notably, transcription affects cohesin deposition and loop formation in our model, as exemplified by simulations in which all genes in this 3-Mbp segment were modeled as tandemly transcribed (Appendix Fig. 1.S6). RNAP depletion eliminated compartmentalization and the frequency of looping was again drastically reduced. Cohesin occupancy was decreased at most loop anchors, with loops again becoming larger (Fig. 1.4E,F). This parsimonious model allows us to deduce that the effects observed in experiments can be explained by a simple relationship between cohesin loading at RNAP-occupied sites and the interplay between active transcription and loop extrusion.

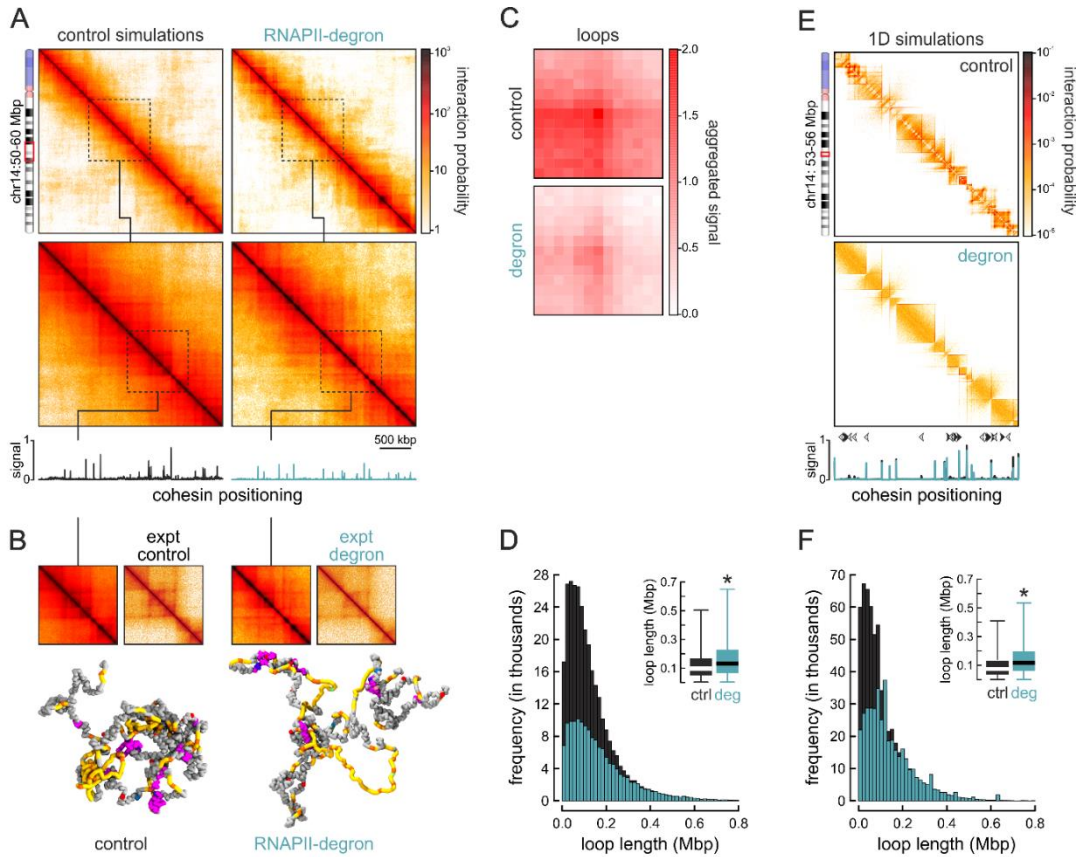


Figure 1.4. Computational modeling of RNA Pol II depletion effects on loop extrusion.

(A) *Top*: Heatmaps rendered from simulations of wild-type (*left*) or RNA Pol II-depleted models (*right*) of HUVEC chr14:50-60 Mbp. *Bottom*: 10 kbp-resolution heatmaps in the chr14:53-56 Mbp region. Cohesin positioning tracks are aligned below each heatmap. (B) Exemplary 3D chromatin folding models of the chr14: 54.5-55.5 Mbp subregion shown in comparison to Hi-C data. (C) APA plots showing weakened loops in RNA Pol II-depleted models. (D) Histogram showing looping frequency in the absence of RNA Pol II (*turquoise*) compared to wild-type models (*black*). *Inset*: Boxplots showing larger loops in RNA Pol II-depleted models (*blue*). *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (E) Heatmaps rendered from 1D simulations representing wild-type (*top*) or RNA Pol II-depleted models (*bottom*) of the chr14 segment from panel A at 3-kbp resolution. Cohesin positioning (*overlaid tracks*) and TSS orientation (*arrows*) are aligned below. (F) As in panel D, but using data from the 1D simulations in panel E.

Discussion

Following cell division, chromosomes refold in order to establish interphase architecture. At this point, cohesin also needs to be reloaded. Its reloading coincides with the extrusion of CTCF-anchored loops and TAD reestablishment (Abramo *et al*, 2019; Zhang *et al*, 2019). However, A/B-compartments, driven by homotypic chromatin interactions, reemerge more rapidly, as do contacts amongst *cis*-regulatory elements. Interestingly, the latter display rates that exceed those of extruded loops (Zhang *et al*, 2019). In parallel, the general transcription factor TBP bookmarks mitotic chromatin to facilitate gene reactivation (Teves *et al*, 2018), and as transcription reinitiates in late telophase, a strong activity spike occurs at most genes and enhancers (Hsiung *et al*, 2016). These observations suggest that RNAP activity may play a central role in reestablishing interphase chromatin organization following mitosis.

Here, using an RPB1 “degron” line (Nagashima *et al*, 2019), we show that RNA Pol II presence on chromatin is necessary for both the establishment of compartments and the formation of loops early in G1. The former is intuitively justified by the homotypic interactions that build the “active” A-compartment, and the recent finding that chromatin acetylation can drive compartment formation (Rosencrance *et al*, 2020). In our RNA Pol II-depleted G1-reentry cells, H3K27ac levels are reduced while H3K27me3 levels increased, and this imbalance most probably underlies many of the compartment-level changes, although alleviating H3K27ac bookmarking had little effect on 3D refolding in mESCs (Pelham-Webb *et al*, 2021).

The latter effect is more perplexing, but agrees with NIPBL binding at active gene promoters (Busslinger *et al*, 2017; Zuin *et al*, 2014; Zhu *et al*, 2021) and, thus, raises two key questions. First, how does RNAP depletion promote formation of hundreds of *de novo* loops that are also longer? According to our simulations, and depending on the direction of elongation, RNAPs can reel DNA such that it counters extrusion while also acting as physical blockades to it. This is reminiscent of the condensin-polymerase antagonism reported for bacteria (Brandão *et al*, 2019) and flies (Rowley *et al*, 2019), and inferred by super-resolution imaging in mESCs (Gu *et al*, 2020). Some of these newly emerging loops form on the basis of strengthened Polycomb interactions, justified by the increase in H3K27me3 levels upon RNA Pol II degradation. Recent work showing that cohesin removal also resulted in enhancer looping of Polycomb-bound regions (Vian *et al*, 2018) suggests that RNA Pol II depletion crosstalks with cohesin loading. As regards increased loop lengths, our simulations argue that this is a consequence of reduced cohesin loading rates to chromatin, as well as of the different loading patterns in control versus RNAP-depleted cells. These

translate into fewer cohesin rings acting locally at any given time, and such reduced “crowding” leads to fewer extrusion conflicts allowing longer loops to form. However, when compared to data from *Wapl*- and/or *Scc4*-knockout cells showing loop enlargement (Davidson *et al*, 2019), our Hi-C data differ (**Appendix Fig. 1.S7A**). But, when compared to Hi-C data from cells undergoing hyperosmotic shock, which affects RNAP as well as cohesin binding to chromatin (Amat *et al*, 2019), there exists more similarities in the interaction patterns (**Appendix Fig. 1.S7B**). These comparisons suggest a distinct and more generalized disruption of the cohesin loading-unloading cycle following RNA Pol II depletion.

Second, why are the effects of RNA Pol II depletion more obvious upon G1 reentry? We believe that this is due to a combination of effects. On the one hand, cell synchronization counters the inherent heterogeneity of contacts in individual cells. On the other, early chromatin refolding and transcription bursts in the mitosis-to-G1 transition suggest that RNAPs preempt a central role in establishing a loop-based chromosomal architecture by instructing cohesin loading and setting up compartments.

In summary, we uncovered a dependency of loop extrusion on RNA Pol II that predominates genome reorganization following exit from mitosis. This dependency may be speaking of the significance of mitotic bookmarking, as transcription factor association with mitotic chromatin could dictate RNAP positioning and, in a next step, cohesin loading and loop extrusion. Nonetheless, the precise interplay between polymerases, transcription factors and cohesin subunits during this transition remains to be elucidated.

Materials and methods

Cell synchronization and sorting

mAID-POLR2A(RPB1)-mClover DLD-1 (Nagashima *et al*, 2019) and TO2B/--TOP2A-mAID HCT116 cells (Roukos lab) were grown in RPMI1640 medium supplemented with 10% FBS under 5% CO₂. Inducible depletion of RPB1 or TOP2A initiated via treatment with doxycycline for 24 h to induce *TIR1* expression, before addition of 500 μM indole-3-acetic acid solution (“auxin”, Sigma-Aldrich) for different times to induce RPB1 degradation. For cell synchronization, G2/M arrest was achieved by the addition of 10 μM RO-3306 inhibitor for 21 h. Following this incubation time, cells were washed with PBS, and auxin-supplemented medium was added for up to 6 h to allow cells to quantitatively enter G1. At this point, synchronized or asynchronous cells treated with auxin for up to 14 h were harvested, where applicable resuspended in 1 μg/ml propidium iodide to counterstain DNA, and sorted to isolate G1 cells using a FACS Canto II flow cytometer (Becton Dickinson).

***In situ* Hi-C and data analysis**

All *in situ* Hi-C was performed using the Hi-C+ kit (Arima Genomics) as per manufacturer's instructions. The resulting Hi-C libraries were paired-end sequenced on a NovaSeq6000 platform (Illumina) to >450 million read pairs per replicate. Reads were separately aligned to the reference build of the human genome (hg38) using BWA and Juicer (v. 1.11.09) to generate .hic files (Durand *et al*, 2016). Only reads with MAPQ >30 were considered for further analysis, and bin-to-bin interactions were extracted from KR-balanced matrices in .hic files using the Juicer "dump" utility at different resolutions. A-/B-compartment stratification was performed using the "eigenvector" Juicer utility on 250 kbp-resolution matrices, with both gene and H3K27ac ChIP-seq signal density used to deduce A-compartments. Saddle plots were generated as described previously (Zhang *et al*, 2019). For topologically-associating domains (TADs), KR-balanced matrices we processed via a combination of "directionality index" plus HMM tools at 10 kbp-resolution and in 500-kbp windows. TADs smaller than 150 kbp or found in centromeric regions were filtered out. For a TAD to be considered "shared" between two datasets, boundary positions should not shift by >60 kbp, and their coordinates should overlap should at least 90%. Insulation scores at TAD boundaries were calculated using a sliding 120 kbp x 120 kbp window along the matrix diagonal at 10-kbp resolution as previously described (Crane *et al*, 2015); squares with a sum of interactions <12 were filtered out. Aggregate TAD plots were generated using *Coolpuppy* (Flyamer *et al*, 2020). For loop detection, we used SIP (Rowley *et al*, 2020) and standard parameters: `-res 10000 -mat 2000 -g 2 -d 3 -fdr 0.01 -nbZero 4 -cpu 1 -factor 1 -max 2 -min 2 -sat 0.01 -t 2800 -norm KR -del true`, and an FDR <0.01 to filter the resulting loop lists. Loops specific to a given condition were determined using *pgltools* (with `-d 29999` due to the smallest allowable loop size of 30 kbp) (<https://github.com/billgreenwald/pgltools>). Aggregate peak plots were generated via the APA utility in Juicer using standard parameters (`-r 10000 -k KR -q 3 -w 6 -n 15 -u`), before scaling between 0-2 to facilitate comparison. Last, architectural stripes in Hi-C data were detected using *stripenn* (<https://github.com/ysora/stripenn>). Visualizations and plots were performed using data from merged Hi-C replicates (except for TOP2A-mAID HCT1166 and 2-h RPB1-mAID data). All code is available at: <https://github.com/shuzhangcourage/HiC-data-analysis>.

High throughput 3D-DNA fluorescence in situ hybridization (FISH)

Dual color DNA FISH was performed using the BAC probes targeting different chromosomes and labeled with Alexa488-dUTP, Alexa568-dUTP or Alexa647-dUTP by nick translation on G1-sorted control and auxin-

treated DLD-1 reentry cells seeded on glass slides. Images were acquired using an Opera Phenix High Content Screening System (PerkinElmer), equipped with four laser lines (405 nm, 488 nm, 568 nm, and 640 nm) and two 16-bit CMOS cameras. Images for 3D and radial distances were acquired in confocal mode using a 40X water objective (NA 1.1), and analyzed as described previously to also quantify DNA content and infer cell cycle phase stratification (Roukos *et al*, 2015).

Chromatin immunoprecipitation (ChIP) coupled to sequencing

DLD-1 cells cultured to 80% confluence in 15-cm dishes were crosslinked in 1% PFA/PBS 10 min at room temperature. Cells were processed using the NEXSON ChIP protocol (Arrigoni *et al*, 2016). In brief, nuclei were isolated via sonication using a Bioruptor Pico (Diagenode; 9 cycles of 10 sec *on* and 30 sec *off*). Chromatin was then sheared in the recommended shearing buffer (27-30 cycles, 30 sec *on* and 30 sec *off*) to a range of 200-500 bp-long fragments, and immunoprecipitation was performed using 4 µg of the appropriate antibody (anti-CTCF: 61311, Active Motif; anti-RAD21: ab88572, Abcam; anti-GFP: ab290, Abcam). Paired-end sequencing was performed on a NovaSeq6000 platform (Illumina) yielding >25 million reads per sample. Raw reads were processed for mapping and peak calling using the ENCODE Data Coordinating Center pipeline (DCC v1.5.0; <https://github.com/ENCODE-DCC>). Coverage plots and heatmaps were generated via Deeptools (Ramírez *et al*, 2014).

Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq)

Tn5 transposed chromatin was isolated from human DLD-1 mAID-RPB1 cells according to the standard ATAC-seq protocol with a modification (Corces *et al*, 2017) aiming at quantitative scaling of the resulting data. In brief, 10⁵ DLD-1 cells per replicate were “spiked” with 200 *D. melanogaster* S2 cells, washed in 1x PBS and added to lysis buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% NP-40, 0.1% Tween-20, and 0.01% digitonin) for 3 min to isolate nuclei. Nuclei were next washed in washing buffer (10 mM Tris-HCl pH 7.4, 10 mM NaCl, 3 mM MgCl₂, 0.1% Tween-20) and pelleted by centrifugation. Isolated nuclei were resuspended in transposase reaction mix (25 µl 2x TD buffer, 16.5 µl 1x PBS, 0.5 µl 10% Tween-20, 0.5 µl 1% digitonin, 2.5 µl Tn5, and 5 µl nuclease-free H₂O) and incubated at 37°C for 30 min under constant shaking at 1000 rpm. The transposition reaction was terminated by the addition of stop buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS), and purified using the DNA Clean & Concentrator kit (Zymo Research). Following standard library generation, samples were sequenced to >40 million reads on a NovaSeq6000 platform (Illumina). Read pairs were mapped to the hg38 and dm6 reference genome builds for human and

Drosophila, respectively, using Bowtie 2 (Langmead & Salzberg, 2012). Unmapped, duplicate, and mitochondrial reads were removed before merging replicates. Finally, ChIPseqSpike was utilized for calculating scaling factors (<https://bioconductor.org/packages/release/bioc/html/ChIPseqSpike.html>) to produce RPKM-normalized and scaled coverage.

Cleavage Under Targets and tagmentation (CUT&Tag)

Following lifting from plates using accutase and FACS sorting, 0.5 million G1-phase DLD-1 cells were processed according to manufacturer's instructions (Active Motif). Samples were paired-end sequenced to obtain at least 10^7 reads. Reads were processed according to the standard CUT&Tag pipeline (https://yezhengstat.github.io/CUTTag_tutorial/). Briefly, paired-end reads were trimmed for adapter removal and mapped to the human (hg38) and *E. coli* (ASM584v2) reference genomes using Bowtie 2 (Langmead & Salzberg, 2012). *E. coli* mapped reads were quantified and used for calibrating human-mapped reads. Peak calling was performed with SEACR (Meers *et al*, 2019) and using IgG controls for thresholding peak calling. For stringency, cohesin-bound sites were considered those shared by both the SMC1A and Rad21 control datasets, while CTCF-bound those shared by the CUT&Tag and publicly-available CTCF ChIP-seq data (<http://chip-atlas.org/view?id=DRX013180>). All heatmaps were generated using Deeptools (Ramírez *et al*, 2014).

Immunofluorescence and image quantification

DLD-1 cells grown on coverslips were fixed in 4% PFA/PBS for 10 minutes at room temperature. After washing once in PBS, cells were permeabilized with 0.5% Triton-X/PBS for 5 min at room temperature, washed three times in PBS, blocked using 1% BSA for 1 h, and incubated with the appropriate primary antibody for 1 h at room temperature (anti-RNA Pol II: 1:500, 61086, Active Motif; anti-H3K27ac: 1:500, 39133, Active Motif; anti-H3K27me3: 1:500, 39155, Active Motif; anti-Rad21: 1:800, ab992, Abcam; anti-Fibrillin: 1:100, sc-393968, Santa Cruz). For visualizing nascent transcripts, cells were pre-incubated with 3 mM 5-ethynyl uridine (EU) for 30 min at 37°C in their growth medium, fixed and processed with the Click-iT EdU chemistry kit (Invitrogen). Images were acquired on a Leica dmi8 microscope using the LASX software. Quantification of nuclear fluorescence were performed by drawing a mask based on DAPI staining, and then calculating the mean intensity per area falling under this mask. Colocalization was assessed using the ImageJ plugin, JACoP (<https://imagej.nih.gov/ij/plugins/track/jacop.html>).

Chromatin fractionation and western blotting

For assessing protein abundance in different sample preparations, approx. 10^6 cells were gently scraped off 15-cm dishes, and pelleted for 5 min at $600 \times g$ at room temperature, supernatants were discarded, and pellets resuspended in 100 μ l of ice-cold RIPA lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 1% NP-40, 1% sodium deoxycholate) containing 1x protease inhibitor cocktail (Roche). Next, lysates were incubated for 20 min on ice and centrifuged for 15 min at $>20,000 \times g$ to pellet cell debris to collect the supernatants. The concentration of each protein extract was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific). For fractionation, the protocol previously described was used (Watrin *et al*, 2006). Following protein separation on precast SDS-PAGE gels (BioRad), proteins were detected using different primary antibodies, and visualized using the Pierce SuperSignal West Pico ECL kit (Thermo Fisher Scientific).

Factory RNA sequencing and data analysis

Nascent RNA from ~ 10 million mAID-RPB1-mCLOver DLD-1 or TO2B^{-/-}-TOP2A-mAID HCT116 cells was isolated according to “factory-seq” protocol (Caudron-Herger *et al*, 2015). Briefly, cells were gently scraped and lysed in isotonic “physiological buffer” supplemented with 0.5% NP40 buffer. After assessing lysis and nuclei integrity on a hemocytometer microscopy, nuclei were treated with DNase I (Worthington) for 30 min at 33°C, washed, and nuclei were lysed in “native lysis buffer” and treated with caspase group III enzyme mix (PromoKine), pelleted by centrifugation, before the supernatant holding nascent RNA was collected in TRIzol (Invitrogen) and purified using the Direct-Zol RNA purification kit (Zymo). Following standard strand-specific cDNA library preparation using the TruSeq kit (Illumina), sequencing was performed on a NovaSeq6000 platform (Illumina) to >40 million paired-end reads. Raw reads were mapped to human genome (build hg38) using STAR (Dobin *et al*, 2013), quantified using iRNAseq (Madsen *et al*, 2015) and the *-gene* option, before RUVseq normalization (Risso *et al*, 2014) and differential gene expression analysis. For gene set enrichment, GSEA (<http://www.gsea-msigdb.org/gsea/index.jsp>) was run on the significantly changing genes ($P_{adj} \leq 0.05$).

Dual-color super-resolution dSTORM imaging and analysis

DLD-1 control and auxin-treated reentry cells were seeded onto coverslips, fixed, stained, and imaged as described previously (Casa *et al*, 2020). In brief, fixed and immunostained cells for NIPBL and CTCF (as

described above) were mounted to an Attofluor cell chamber (Thermo Fisher Scientific) and in 1 mL of *d*STORM buffer (25 mM MEA, glucose oxidase, 50 mM NaCl, and 10% glucose in 10 mM Tris-HCl pH 8.0). The chamber is then sealed with a coverslip and left on the microscope at room temperature for 30 min prior to imaging, to minimize drift. Imaging was performed on a Zeiss Elyra PS1 system fitted with an Andor iXon DU 897, 512 × 512 EMCCD camera. Images were made using a 100× 1.49NA TIRF objective in HiLo mode. Movies of 12,000 frames were recorded with an exposure time of 33 msec. Multichannel images were acquired sequentially from high wavelength to lower wavelengths. *d*STORM movies for each protein target were analyzed via the Zeiss ZEN 2012 software, and any localizations with a precision of >50 nm were discarded. All remaining localizations were drift-corrected using a model-based approach. All additional analysis was done in R (<https://www.R-project.org/>), localizations from individual nuclei were clustered based on their density using a kernel density estimation (KDE)-based clustering algorithm with the threshold set to 0.05 for all channels. The areas of CTCF or NIPBL clusters were measured using the KDE binary image, and distances between closest neighbors calculated.

Computational modeling

For the 3D simulations, we used our previously described HiP-HoP model (Buckle *et al*, 2018), extended to account for interactions in inactive regions. This model combines our initial “transcription factory” model (Brackley *et al*, 2016) with loop extrusion (Fudenberg *et al*, 2016), while also accounting for the heteromorphic nature of the chromatin fiber, which means that the local compaction (in DNA base pairs per nanometer) varies along the fiber. Here, we modeled a 10-Mbp region of HUVEC chr14 as a bead-spring polymer containing N beads of diameter σ , each representing 1 kbp of chromatin. We allowed beads in the polymer to interact via three potentials: (i) a Weeks-Chandler-Andersen (WCA) potential, which provides excluded volume interactions; (ii) a finitely-extensible-nonlinear-elastic (FENE) potential accounting for chain connectivity; and (iii) a Kratky-Porod potential describing the flexibility of the chain with parameters set to give a persistence length of 4-5 kbp (in line with that of chromatin *in vivo*). To model the heteromorphic nature of the chromatin fiber in a simple way, we included additional springs (with constants of $200 k_B T / \sigma^2$) between next-to-neighbor chromatin beads along the chain (i.e., beads i and $i+2$) which are not associated with H3K27ac marks. As H3K27ac marks correlate with active euchromatin regions, these springs cause a local crumpling of the polymer in inactive chromatin fragments, or equivalently a swelling in active regions accounting for their generally more open conformation. Transcription factors (TFs) were simulated as diffusing beads interacting with each other via steric repulsion, again modeled via WCA potentials. We considered three types of TFs: (i) generic active TFs bind

strongly (potential depth $7.9 k_B T$) to chromatin beads associated with accessible chromatin (defined using ENCODE DHS-seq data), and weakly to beads associated with H3K27ac (potential depth $3.4 k_B T$); (ii) HP1-like inactive TFs bind to beads enriched in H3K9me3 marks (potential depth $3.4 k_B T$); and (iii) Polycomb-like TFs bind to beads enriched in H3K27me3 marks (potential depth $7.9 k_B T$). All of these interactions were modeled via a truncated-and-shifted Lennard-Jones (LJ) potential. To account for post-translational modifications, we allowed each TF type to switch between a binding and a non-binding state at a defined rate ($k_{sw} = 10^{-3} \tau^{-1}$, where τ is the simulation time unit). The binding state was characterized by the aforementioned interaction strengths, whereas the non-binding state only by steric interactions with chromatin beads (via WCA potentials). We considered 250 active TFs, 625 HP1-like TFs, and 125 Polycomb-like TFs in the wild-type simulations, whereas the RNA Pol II-degron simulations were run without active TFs. We also implemented non-specific interactions between inactive chromatin beads (via a truncated-and-shifted LJ potential with depth $0.45 k_B T$) to account for the generic “phase separation” between eu- and hetero-chromatin. Finally, loop extrusion was modeled by representing cohesin dimers as further additional springs. Loop extrusion dynamics were determined by the number density of cohesins ($n_c = 0.01/\text{kbp}$) on the chromatin fiber and two rates: the unbinding rate ($k_{off} = 2.5 \times 10^{-5} \tau^{-1}$) and the extrusion rate ($v = 4 \times 10^{-3} \text{kbp}/\tau$). Upon binding of cohesin, we introduced an additional spring between two nearby beads along the fiber (i and $i+3$, since crumpling springs already link i and $i+2$); the equilibrium length and the spring constant of cohesin bonds were set to 1.5σ and $40 k_B T/\sigma^2$, respectively. When cohesins were removed from chromatin, they were instantly repositioned along the fiber. Wild-type conditions were simulated by background random loading (with 10% probability) but with predominant loading at DHS beads (with 90% probability); the RNA Pol II-degron was simulated by only considering random loading. Finally, a cohesin halted either upon colliding with another extruding complex or upon reaching a CTCF site whose direction was against the direction of extrusion (as shown experimentally; ref. Rao *et al*, 2017). Note that CTCF sites and orientation were obtained by ENCODE tracks, taking care to include in our simulations only sequences overlapping cohesin (RAD21) peaks – this procedure singles those CTCF binding sites that are relevant to looping. All constituents of the system (chromatin beads and TFs) were allowed to diffuse, and their dynamics were governed by a Langevin equation as described before (Buckle *et al*, 2018) and implemented using Python and the LAMMPS molecular dynamics software package (Plimpton, 1995) as a library.

For the 1D simulations, we considered a 3 Mbp-long chromatin fiber coarse-grained into segments of 1 kbp. Again, we modeled data from a specific subregion of HUVEC chr14. We simulated the dynamics of cohesin rings (total number $N_{rings} = 30$, comparable to what was used in ref. Fudenberg *et al*, 2016), each

of which could be in one of two states: either bound (i.e., on the fiber) or unbound (i.e., in the diffuse pool). Binding and unbinding were modeled as stochastic processes with rates k_{on} and k_{off} , respectively. When on the fiber, a cohesin molecule was modeled as a dimer, with each monomer undergoing active extrusion at speed v . Each monomer could proceed until it hit a CTCF site with orientation conflicting with its direction of travel, at which point it became immobile (Fudenberg *et al*, 2016). If a cohesin complex was halted on one side, its other side could continue to move independently. When both monomers in a cohesin dimer became stuck at convergent CTCF sites, the unbinding rate of the dimer was decreased by a factor of 10 to model CTCF-mediated stabilization of extruded chromatin loops, in line with results in ref. Fudenberg *et al*, 2016. We let monomers in a cohesin dimer interact with each other via steric exclusion so that extrusion would be halted temporarily if another monomer was in their way. Wild-type conditions were simulated by assuming that cohesin was loaded as described above, with 10% background random loading and 90% at DHS sites, whereas RNA Pol II-degron conditions were simulated by only retaining the random background loading. To simulate feedback of transcription on extrusion, we assumed that the speed of extrusion was reduced by a factor f when the direction of extrusion and that of transcriptional elongation of an active gene were conflicting. Parameters in the simulations were set to $k_{\text{on}} = 2 \times 10^{-2} \tau^{-1}$, $k_{\text{off}} = 10^{-3} \tau^{-1}$, $v = 0.16 \text{ kbt}/\tau$ for the wild-type, and to $k_{\text{on}} = 2 \times 10^{-3} \tau^{-1}$, $k_{\text{off}} = 10^{-3} \tau^{-1}$, $v = 0.16 \text{ kbt}/\tau$ for the RNA Pol II-degron. These values can be mapped to $k_{\text{on}} = 1 \text{ min}^{-1}$, $k_{\text{off}} = 0.05 \text{ min}^{-1}$, $v = 0.133 \text{ kbp/s}$ (wild-type, corresponding to a residence time on chromatin $\sim 20 \text{ min}$) or $k_{\text{on}} = 0.067 \text{ min}^{-1}$, $k_{\text{off}} = 0.033 \text{ min}^{-1}$, $v = 0.089 \text{ kbp/s}$ (RNA Pol II-degron, corresponding to a residence time on chromatin $\sim 30 \text{ min}$). The k_{off} value in simulation units ($k_{\text{off}} = 10^{-3} \tau^{-1}$) was chosen to be sufficiently small to allow study of extrusion effects. To align simulation times to real times, we consider a residence time on chromatin of ~ 20 and 30 min for the wild-type and “degron” simulations, respectively. The wild-type value is comparable to that reported previously (Hansen *et al*, 2017). We also set $k_{\text{on}} = 1 \text{ min}^{-1}$ (wild-type; while there is no accurate measurements of this rate, it needs to be significantly larger than k_{off}) and reduced its value by a factor 10 in RNA Pol II-degron simulations in order to have the same rate of random loading both scenarios. Finally, extrusion speed v was chosen such that the chromatin length explored during an extrusion event, $\lambda = v/k_{\text{off}}$, was the same for wild-type and “degron”, and comparable to that used previously (Fudenberg *et al*, 2016). For simulations including the feedback of transcription on extrusion, we varied f (which is unknown experimentally) between 0.1 and 0.9 to simulate a variety of scenarios, and presented the case with $f = 0.1$, which led to the most pronounced effects.

Statistical analyses

P-values associated with the Student's *t*-tests and Fisher's exact tests were calculated using GraphPad (<http://graphpad.com/>), those associated with the Wilcoxon-Mann-Whitney test using R. Unless otherwise stated, *P*-values <0.01 were deemed significant.

Data availability

NGS data generated in this study are available via the NCBI Gene Expression Omnibus (GEO) repository under accession number GSE160321. Any other data related to this paper may be requested from the corresponding author.

Author contributions

S.Z., N.Ü., and N.J. performed experiments. S.Z., N.J., and E.G.G. performed bioinformatics analyses. G.F., M.C., and D.M. performed computational modeling. H.G. and V.R. performed 3D DNA FISH. J.S., A.B.H., and K.S.W. performed super-resolution STORM imaging. C.B. and J.A. performed all next-generation sequencing. A.P. conceived and supervised the study, and compiled the manuscript with input from all coauthors.

Appendix – Supplementary figures 1-6

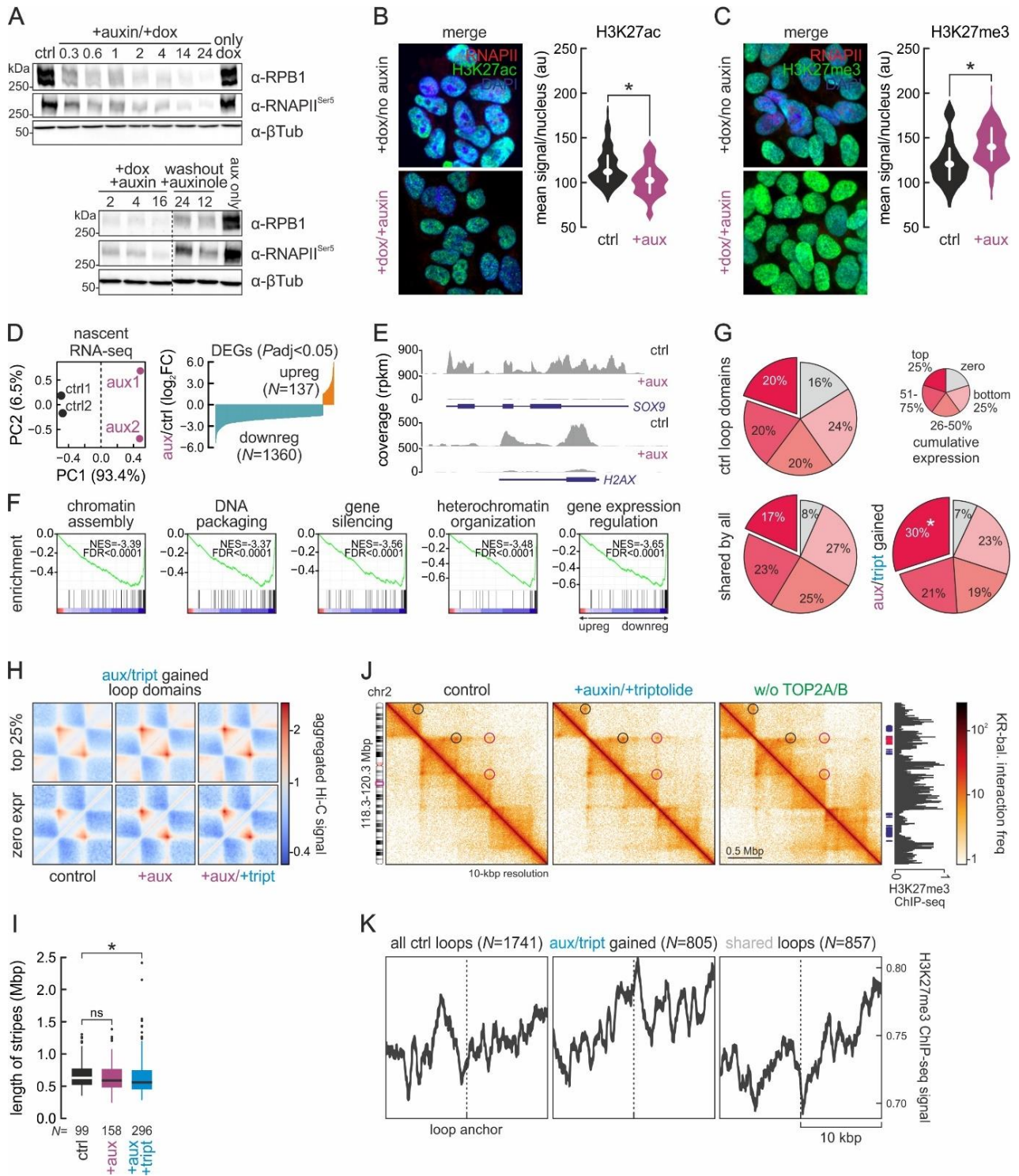


Figure 1.S1. RNA Pol II degradation, recovery and its effects on interphase chromatin folding.

(A) Western blots showing depletion of total cell RPB1 or phospho-Ser5-RNA Pol II on increasing exposure to doxycycline plus auxin (*top*) or RNA Pol II recovery following auxin washout in the presence of auxinole (*bottom*); β -tubulin provides a loading control. (B) Representative H3K27ac immunofluorescence from untreated (*top*) or 14-h auxin-treated cells (*bottom*) and signal quantification (*bean plots*). *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. Bar: 5 μ M. (C) As in panel B, but for H3K27me3 levels. (D) *Left*: PCA plot for G1-sorted control (*black*) and 14-h auxin-treated nascent RNA-seq data (*purple*). *Right*: Nascent RNA changes (\log_2 fold-change compared to control; $P_{\text{adj}} < 0.05$) in 1497 genes upon auxin treatment. (E) Genome browser examples of nascent RNA reduction at two typical gene loci. (F) Gene set enrichment analysis of data in panel D; top five enriched pathways. (G) Pie charts showing distribution of all (*top*), shared (*bottom left*) and auxin/triptolide-shared loops (*bottom right*) according to their cumulative gene expression levels (no expression – *grey*; four nonoverlapping quantiles – *shades of red*). *: significantly different; $P < 0.05$, Fisher's exact test. (H) Heatmaps showing mean loop domain interactions in control (*top*), auxin- (*middle*) and auxin-/triptolide-treated cells (*bottom*) for the auxin/triptolide-shared loops of panel G. (I) Boxplots showing changes in the length of stripes detected in Hi-C data from control (*black*), auxin-treated (purple) or auxin/triptolide-treated Hi-C data (*blue*). *: significantly different to control; $P < 0.05$, Wilcoxon-Mann-Whitney test. (J) Hi-C maps from control (*left*), auxin/triptolide-treated (*middle*) or TOP2A/B-depleted cells (*right*) in the chr2 subregion encompassing the *HOXD* gene cluster. H3K27me3 ChIP-data from control cells are aligned to the maps, and emerging H3K27me3-anchored loops are denoted (*red circles*). (K) Line plots showing mean H3K27me3 ChIP-seq signal enrichment in the 20 kbp around all (*left*), auxin-/triptolide-gained loops (*middle*) or loops shared between RNA Pol II-depleted and control cells (*right*). The number of loops in each group (N) is indicated

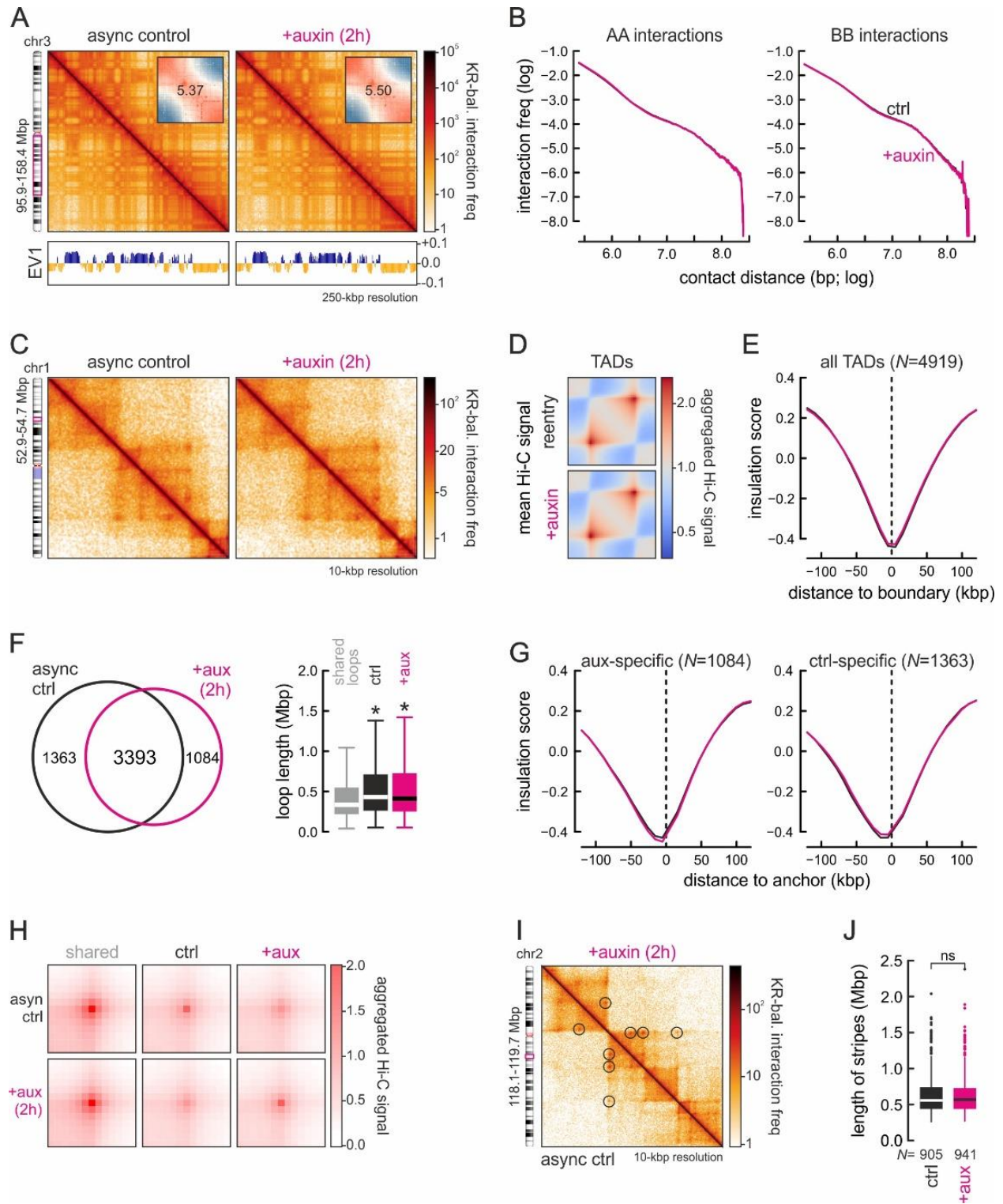


Figure 1.S2. Short-term RNA Pol II degradation does not affect G1-phase chromatin folding. (A) Exemplary Hi-C maps of a subregion of chr3 from asynchronous control (left) and 2-h auxin-treated DLD1-mAID-RPB1 cells (right) at 250-kbp resolution aligned to first eigenvector values (EV1; below). *Insets*: saddle plots showing no change in A/B-compartment insulation. (B) Decay plots showing Hi-C interaction frequency between A- (left) or B-compartments (right) as a function of genomic distance (log) in control (black line) and 2-h auxin-treated cells (magenta line). (C) Exemplary Hi-C maps of a subregion of chr1 from control (left) and auxin-treated cells (right) at 10-kbp resolution. (D) Heatmaps showing aggregated TAD-level interactions in control (top) and auxin-treated cells (bottom). (E) Line plots showing mean insulation score from control (black line) and auxin-treated cells (magenta line) in the 240 kbp around all TAD boundaries in control cells. The number of TAD boundaries queried (N) is indicated. (F) *Left*: Venn diagram showing shared and unique loops in control (black) and auxin-treated Hi-C data (magenta). *Right*: Loop lengths displayed as boxplots. *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (G) As in panel E, but for the anchors of control- (left) and degron-specific loops (right; from panels H,I). (H) Aggregate plots showing mean Hi-C signal at shared (left), control- (middle), and degron-specific loops (right) from panel F. (I) Composite Hi-C map showing little change in loop emergence (circles) between control (bottom half) and auxin-treated cells (top half). (J) Boxplots showing no significant change in the length of stripes detected in Hi-C data from control (black) or auxin-treated cells (magenta).

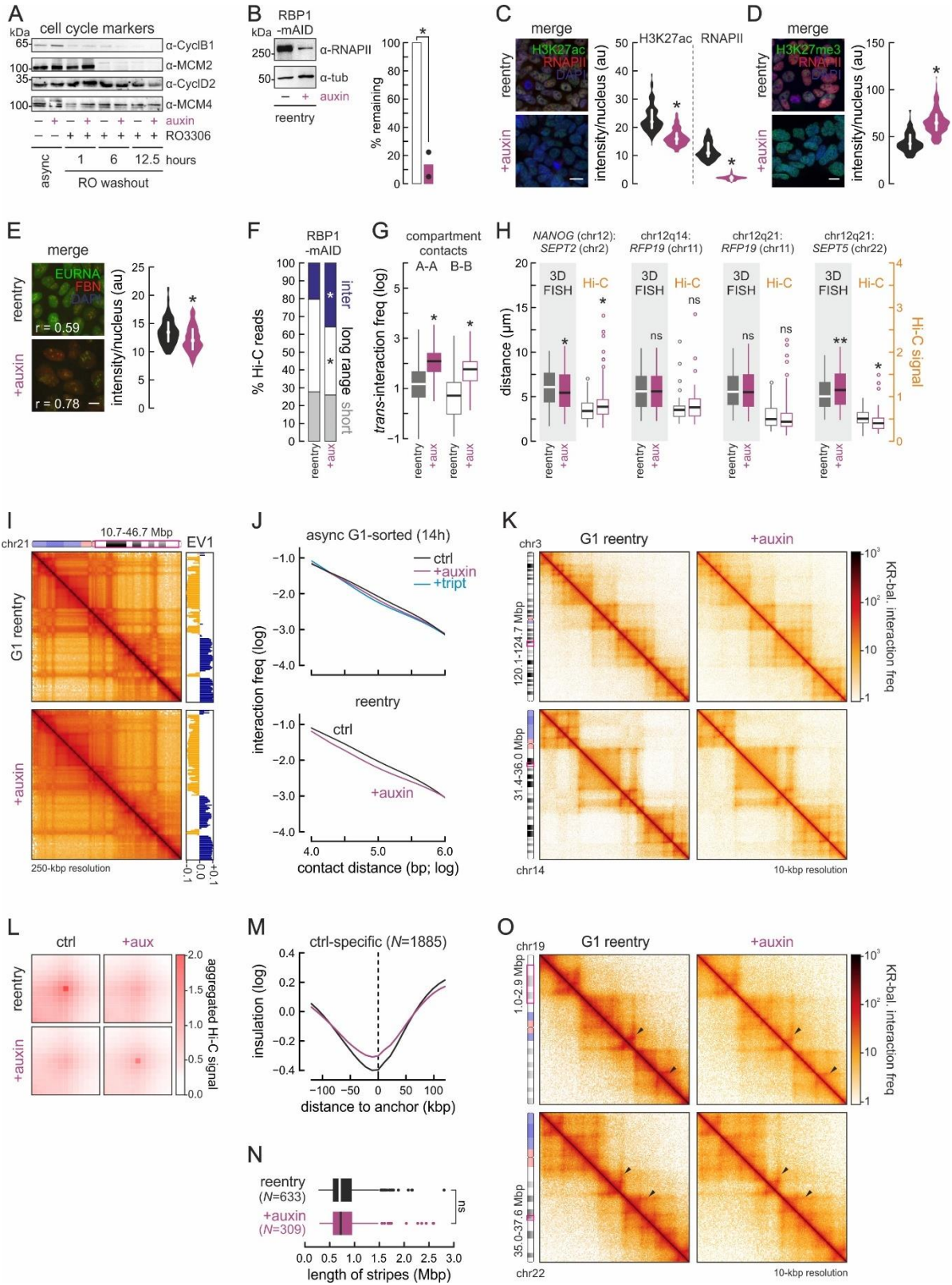


Figure 1.S3. RNA Pol II degradation affects chromatin refolding in cis and in trans following mitosis. (A) Western blots of selected cell cycle markers in cells treated or not with auxin at different times after release from the G2/M block. (B) Left: Western blot of RNA Pol II (RPB1) in G1-reentry cells treated or not with auxin; β -tubulin provides a loading control. Right: Quantification of such western blot data from two independent replicates. *: significantly different mean; $P < 0.01$, unpaired two-tailed Student's t-test. (C) Left: Exemplary widefield immunofluorescence images of DLD1-mAID-RPB1 G1-reentry cells treated with doxycycline plus auxin (bottom) or not (top) and stained for H3K27ac and RNA Pol II (RPB1); nuclei were counterstained with DAPI. Right: Bean plots showing mean fluorescence intensity per nucleus. *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (D) As in panel C, but stained for RNA Pol II and H3K27me3, and quantifying H3K27me3 levels. (E) As in panel C, but stained for fibrillin and EU-labeled nascent RNA, and quantifying EU-RNA levels. The correlation of signal from the two fluorescent channels (r) is also indicated. (F) Bar plots showing percent of Hi-C reads in inter- (blue), long-range (>20 kbp) or short-range intra-chromosomal contacts (white) across datasets. *: significantly different; $P < 0.01$, Fisher's exact test. (G) Boxplots showing inter-chromosomal interactions between A-A and B-B compartments in auxin-treated versus control reentry cells. *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (H) Boxplots comparing interchromosomal distance changes for the loci indicated assessed using high throughput 3D-DNA FISH (grey background) or Hi-C data at 0.5-Mbp resolution. *: significantly different; $P < 0.01$, Wilcoxon-Mann-Whitney test. (I) Additional Hi-C examples of a subregion of chr3 from control (top) and auxin-treated reentry cells (bottom) at 250-kbp resolution aligned first eigenvector values (right). (J) Decay plots showing Hi-C interaction frequency as a function of genomic distance (log) at the scale of TADs (0.01-1 Mbp) in control (black line) or auxin-treated reentry cells (purple/blue lines). (K) Additional Hi-C examples of subregions in chr3 and 14 from control (left) and auxin-treated reentry cells (right) at 10-kbp resolution. (L) Plots showing aggregate Hi-C signal for loops lost/gained in control (left) and auxin-treated reentry cells (right). (M) Line plots showing mean insulation scores in the 240 kbp around control- (top) or degran-specific loops (bottom) from control (black line) and auxin-treated cells (purple line). The number of anchors queried (N) is indicated. (N) Boxplots showing no significant change in the length of stripes detected in Hi-C data from control (black) or auxin-treated cells (purple). (O) Hi-C maps showing typical changes in "stripes" (arrowheads) between control (left) and auxin-treated reentry cells (right) at 10-kbp resolution.

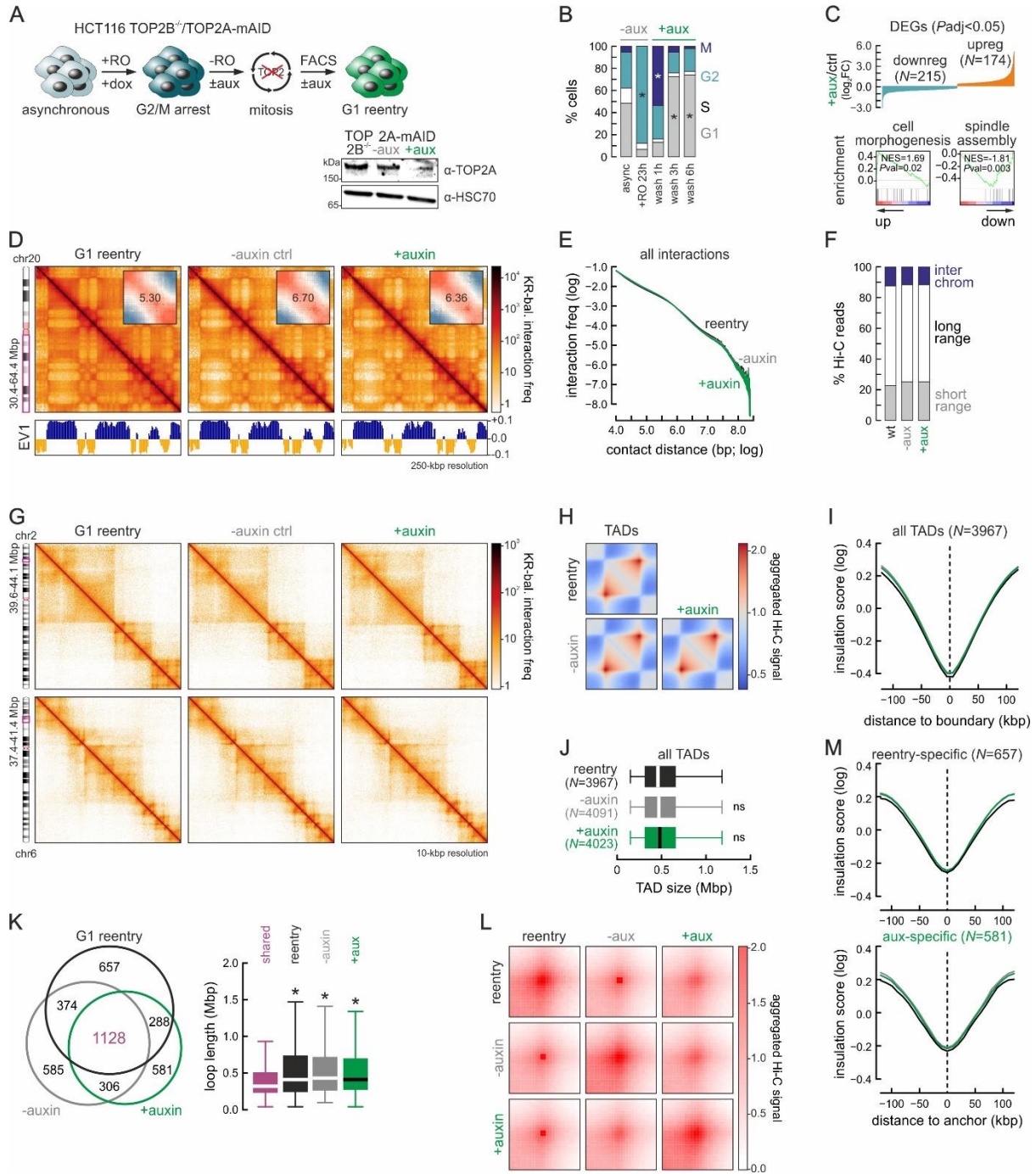


Figure 1.S4. Topoisomerase II depletion marginally affects chromatin refolding following mitosis. (A) Top: Overview of the experimental scheme for HCT116-TOP2B^{-/-}-TOP2A-mAID cell synchronization and release. Bottom right: Western blots showing auxin-mediated TOP2A degradation; HSC70 provides a loading control. (B) Bar plots showing the percent of cells in each phase from panel A. *: significantly different; P<0.01, Fisher's exact test. (C) Top: Graph showing nascent transcription changes (\log_2 fold-change compared to control, $P_{\text{adj}}<0.05$) upon TOP2-depletion. Bottom: Gene set enrichment analysis. (D) Exemplary Hi-C maps of a subregion of chr20 from control (left), TOP2B^{-/-}-untreated and TOP2B^{-/-}-auxin-treated G1-reentry cells (right) at 250-kbp resolution aligned to first eigenvector values (below). Insets: saddle plots showing compartment insulation. (E) Decay plots showing interaction frequency as a function of genomic distance (log) in the Hi-C data from panel D. (F) Bar plots showing the percentage of Hi-C reads in interchromosomal (blue), long- (>20 kbp; white) and short-range intra-chromosomal contacts (grey) in the Hi-C data from panel D. (G) Exemplary 10-kbp resolution Hi-C maps of subregions in chr1 and 6 from the same conditions as in panel D. (H) Heatmaps showing aggregated TAD-level interactions in control (top), TOP2B^{-/-}-untreated (bottom left) and TOP2B^{-/-}-auxin-treated reentry cells (bottom right). (I) Line plots showing mean insulation scores in the 240 kbp around TAD boundaries from panel H. The number of TADs queried (N) is indicated. (J) Boxplots showing size changes in the TAD groups from panel H. (K) Left: Venn diagram showing shared and unique loops between the Hi-C datasets from panel D. Right: Loop lengths displayed as boxplots. *: significantly different to shared loops; P<0.01, Wilcoxon-Mann-Whitney test. (L) Plots showing aggregate Hi-C signal for the loop categories from panel K. (M) As in panel I, but for the anchors of loop from panels K,L

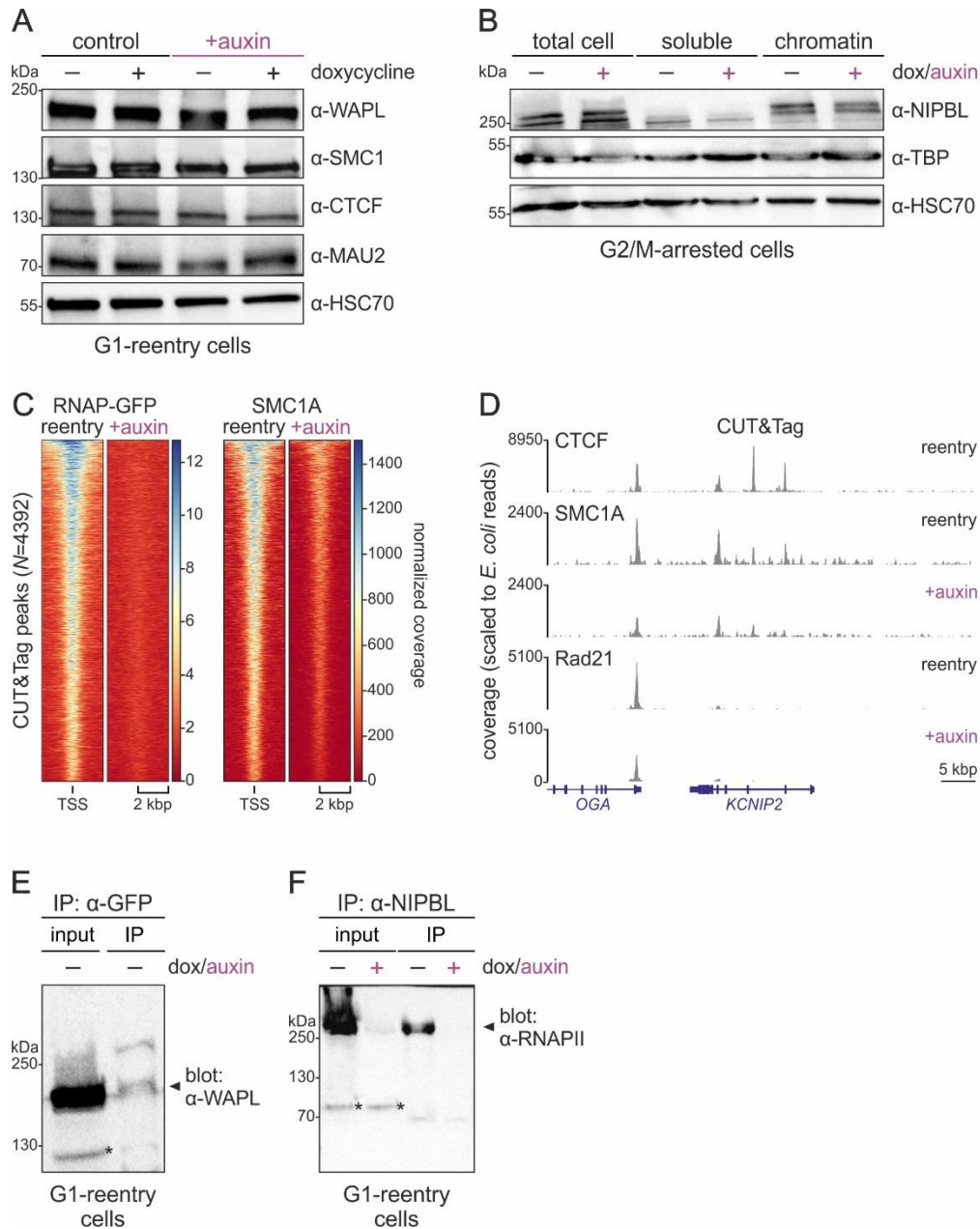


Figure 1.S5. RNAPII interacts with cohesin complex subunits and its depletion affects their binding to DNA. (A) Western blots showing unchanged cohesin components (SMC1A, MAU2, WAPL) and CTCF levels upon RNAPII degradation in G1-reentry cells; HSC70 provides a control. (B) Fractionation blots for NIPBL and TBP in G2/M-arrested cells treated or not with auxin; HSC70 provides a control. (C) Heatmaps showing scaled RNAPII (left) and SMC1A (right) CUT&Tag signal from control and auxin-treated G1-reentry cells in the 4 kbp around active TSSs. (D) Typical browser views of CTCF, SMC1A, and Rad21 CUT&Tag data from G1-reentry cells treated or not with auxin. (E) Western blot showing WAPL co-immunoprecipitating with mClover-mAID-RPB1 in G1-reentry cells. (F) As in panel E, but showing RNAPII (RPB1) co-immunoprecipitating with NIPBL only in untreated G1-reentry cells. *: unspecific bands.

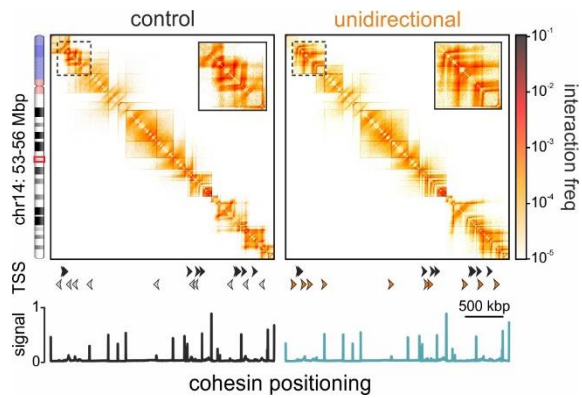


Figure 1.S6. The direction of transcription can affect loop extrusion. Heatmaps rendered from loop extrusion 1D simulations representing wild-type chromatin (left) or chromatin where all TSS are transcribed in the same direction (right) in the HUVEC chr14:53-56 Mbp segment. Profiles of cohesin positioning and TSS orientations are aligned below each heatmap (reoriented TSSs are indicated in orange).

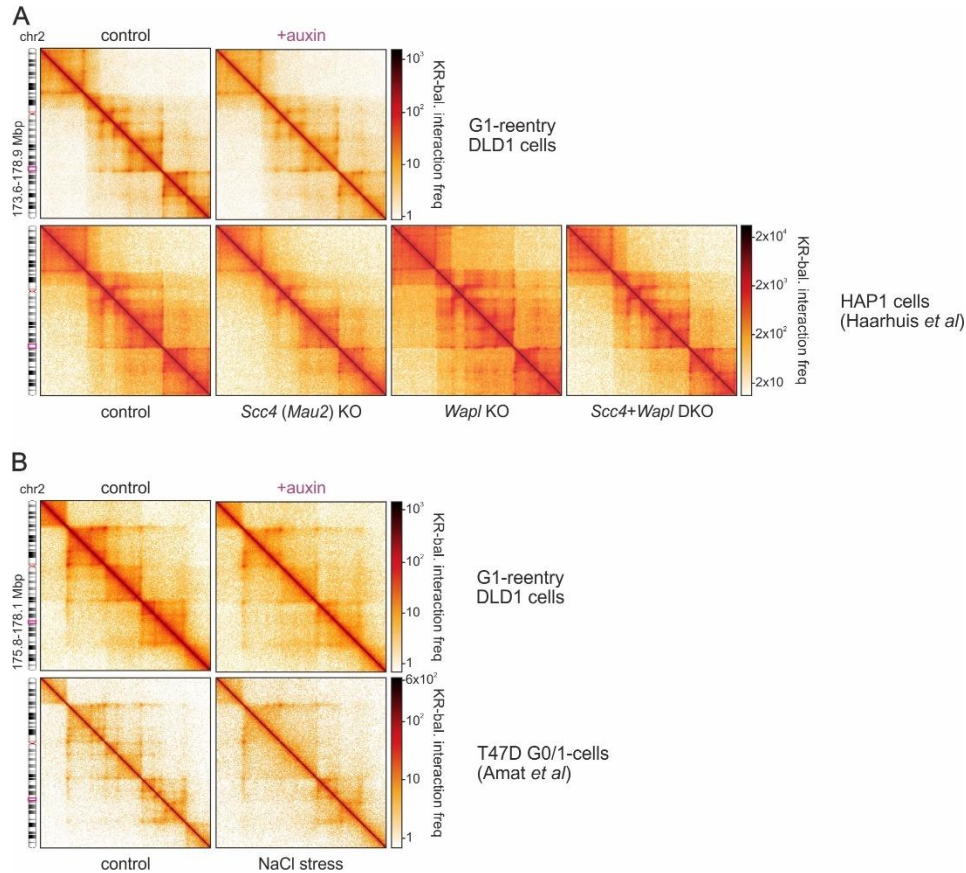


Figure 1.S7. Comparison of RNAPII-depletion to Wapl/Mau2-depletion or hyperosmotic stress Hi-C data. (A) Exemplary Hi-C heatmaps from a subregion for chr2 (ideogram) derived from G1-reentry DLD-1 (top row) compared to heatmaps from wild-type, *Scc4*-KO, *Wapl*-KO, and *Scc4/Wapl*-DKO HAP1 cells (bottom row; data from Haarhuis *et al*, 2017). (B) As in panel A, but comparing Hi-C from G1-reentry DLD-1 (top row) to data from T47D cells before and after NaCl stress (bottom row; data from Amat *et al*, 2019).

Chapter 2 - HMGB1 coordinates SASP-related chromatin folding and RNA homeostasis on the path to senescence

My contribution is reflected in the following experimental/bioinformatic analysis as outlined below:

- Culturing of IMR90 proliferative and senescent cells and validation of senescence in IMR90 cells by qPCR and immunofluorescence staining for CHIP-seq.
- Part of CHIP-seq experiments for senescent IMR90 together with Dr. K. Sofiadis and N. Übelmesser.
- RNA immunoprecipitation experiment for sCLIP-seq HMGB1 validation.
- Bioinformatic analysis of all CHIP-seq and RNA-seq datasets included in the manuscript.
- Part of bioinformatic analysis of Hi-C data from proliferating and senescent IMR90 and HUVEC cells.

The following figure panels were prepared by Dr. A. Papantonis, I. Varamogianni or myself with the data input from experiments prepared and/or analyzed by me:

- Figure 2.1B. Changes in gene expression levels upon senescence entry in IMR90 and HUVEC samples prepared by Dr. A. Zirkel.
- Figure 2.2A. and Figure 2.S2A. Representative HMGB1 CHIP-seq coverage in IMR90 and HUVEC.
- Figure 2.2B-C. Genomic distribution and overlap of HMGB1 peaks in IMR90 and HUVEC CHIP-seq.
- Figure 2.2E-F. and Figure 2.S2B-C HMGB1 and CTCF CHIP-seq signal across TADs and TAD boundaries. Hi-C experiments were performed by Dr. A. Zirkel and Hi-C analysis of TADs performed by Dr. M. Nikolic.
- Figure 2.2G. Hi-C insulation scores at the HMGB1 CHIP-seq peaks.
- Figure 2.2H. and Figure 2.S2D Aggregate Hi-C signal plots of proliferating and senescent loop domains.
- Figure 2.2I. and Figure 2.S2E. Distribution of HMGB1 and Rad21 or CTCF CHIP-seq signal across loop domains.
- Figure 2.2J. and Figure 2.S2F. Differentially expressed genes in senescent versus proliferating IMR90s across loop domains.
- Figure 2.3B-C and Figure 2.S4B-D Panels constructed by I. Varamogianni with the input from TiLO analysis by Dr. M. Nikolic and HMGB1 CHIP-seq data analyzed by me.

- Figure 2.3D-E and Figure 2.S4G. Differentially expressed genes in senescent versus proliferating IMR90s across specific TAD categories and in SASP.
- Figure 2.4B-D Overview of differentially expressed genes upon HMGB1 knock-down performed by Dr. A. Zirkel, GO term enrichment and GSEA analysis of pathway enrichment.
- Figure 2.5E-I Analysis of alternative splicing in senescent versus proliferating and HMGB1 knock-down RNA-seq data.
- Figure 2.S2B. ChIP-seq profiles of H3K27ac and CTCF at HMGB1 ChIP-seq peak positions.
- Figure 2.S2D. and Figure 2.S2F. Motif and KEGG pathway enrichment analysis of HMGB1 ChIP-seq
- Figure 2.S2G-H. Differentially expressed genes in proliferating HUVEC and IMR90 associated with HMGB1 binding.
- Figure 2.S4A. Distribution of TADs in TiLO data and HMGB1 ChIP-seq signal enrichment.
- Figure 2.S5G-H. GO term enrichment of differentially expressed senescent versus proliferating HMGB1 bound genes and correlation of senescent/proliferating and HMGB1 knock-down RNA-seq datasets.
- Figure 2.S6D and 2.S6F-H. GO term enrichment of RNA-seq differentially expressed genes overlapping with sCLIP-seq data. sCLIP-seq performed and analyzed by Dr. Y. Kargapolova.
- Figure 2.S6N. RIP-qPCR experiment for validation of sCLIP-seq HMGB1 bound mRNAs.

HMGB1 coordinates SASP-related chromatin folding and RNA homeostasis on the path to senescence

Konstantinos Sofiadis^{1,#}, Natasa Josipovic^{1,#}, Milos Nikolic^{2,#}, Yulia Kargapolova^{2,6}, Nadine Übelmesser¹, Iliia Varamogianni¹, Anne Zirkel², Ioanna Papadionysiou¹, Gary Loughran³, James Keane^{3,4}, Audrey Michel³, Eduardo G. Gusmao¹, Christian Becker⁵, Janine Altmüller⁵, Theodore Georgomanolis^{2,5}, Athanasia Mizi¹ & Argyris Papantonis^{1,2,*}

¹Institute of Pathology, University Medical Center Göttingen, 37075 Göttingen, Germany

²Center for Molecular Medicine Cologne, University of Cologne, 50931 Cologne, Germany

³Ribomaps, Western Gateway Building, T12 XF62 Cork, Ireland

⁴Cork Institute of Technology, T12 P928 Cork, Ireland

⁵Cologne Center for Genomics, University of Cologne, 50931 Cologne, Germany

⁶Present address: Heart Center, University Hospital Cologne, 50937 Cologne, Germany

#These authors contributed equally to this work

*Corresponding author: AP

Abstract

Spatial organization and gene expression of mammalian chromosomes are maintained and regulated in conjunction with cell cycle progression. This is perturbed once cells enter senescence and the highly abundant HMGB1 protein is depleted from nuclei to act as an extracellular pro-inflammatory stimulus. Despite its physiological importance, we know little about the positioning of HMGB1 on chromatin and its nuclear roles. To address this, we mapped HMGB1 binding genome-wide in two primary cell lines. We integrated ChIP-seq and Hi-C with graph theory to uncover clustering of HMGB1-marked topological domains that harbor genes involved in paracrine senescence. Using sCLIP and functional tests, we show that HMGB1 is also a *bona fide* RNA-binding protein (RBP) binding hundreds of mRNAs. It presents an interactome rich in RBPs implicated in senescence regulation. The mRNAs of many of these RBPs are directly bound by HMGB1 and regulate availability of SASP-relevant transcripts. Our findings reveal a broader than hitherto assumed role for HMGB1 in coordinating chromatin folding and RNA homeostasis as part of a regulatory loop controlling cell-autonomous and paracrine senescence

Introduction

The high-mobility group box 1 (HMGB1) protein, a member of the highly conserved non-histone DNA binding HMG protein family, was named after its characteristically rapid electrophoretic mobility (Stros, 2010). HMGB1 is the most abundant non-histone protein in mammalian nuclei, with 1 HMGB1 molecule per every ~10 nucleosomes (Thomas & Stott, 2012). Despite its high abundance and conservation, HMGB1 has been predominantly studied as an extracellular signaling factor, hence its characterization as an “alarmin” (Lohani & Rajeswari, 2016; Bianchi *et al*, 2017).

To function as an alarmin, HMGB1 is actively secreted by cells like activated monocytes and macrophages or passively released by necrotic and damaged cells. Once received by other cells in the niche, HMGB1 is recognized by RAGE receptors to potently signal inflammation (Scaffidi *et al*, 2002; Bonaldi *et al*, 2003; Orlova *et al*, 2007). In cells entering senescence, HMGB1 translocates from the nucleus to the cytoplasm and is then secreted to stimulate NF- κ B activity via Toll-like receptor signaling. This relocalization and secretion controls the senescence-associated secretory phenotype (SASP) of cells, thus representing a major paracrine contributor both *in vitro* and *in vivo* (Salminen *et al*, 2012; Acosta *et al*, 2013; Davalos *et al*, 2013).

Inside proliferating cell nuclei, HMGB1 has been studied in some detail for its contribution to DNA repair (Ito *et al*, 2015; Mukherjee & Vasquez, 2016), V(D)J recombination (Little *et al*, 2013; Zagelbaum *et al*, 2016) or chromatin assembly (Bonaldi *et al*, 2002), but far less for its transcriptional role (Calogero *et al*, 1999; Mitsouras *et al*, 2002). Cells lacking HMGB1 contain reduced numbers of nucleosomes, rendering chromatin more susceptible to DNA damage, spurious transcription, and inflammatory activation (Giavara *et al*, 2005; El Gazzar *et al*, 2009; Celona *et al*, 2011; Toma *et al*, 2014). As regards its association with chromatin, HMGB1 is thought to bind in a nonspecific manner via its two HMGB-box domains. This allows it to bend and contort DNA and, thus, facilitate recruitment of transcription factors like p53 (Stros, 2010; Rowell *et al*, 2012). HMGB1 associates with cognate DNA sites via characteristically high “on/off” rates, while its acidic tail is important for stabilizing binding (Pallier *et al*, 2003; Ueda *et al*, 2004; Stros, 2010; Blair *et al*, 2016). However, HMG-box DNA-binding domains are particularly insensitive to standard fixatives like formaldehyde (Pallier *et al*, 2003; Teves *et al*, 2016). Thus, capturing HMGBs on chromatin is challenging, and there exist no datasets describing HMGB1 binding in mammalian cells (see <http://chip-atlas.org>). As a result, our appreciation of its on-chromatin roles remains vague.

To address this, we employed a tailored approach that previously allowed us to efficiently map HMGB2 binding genome-wide (Zirkel *et al*, 2018). We can now show that HMGB1 binding in primary endothelial and lung fibroblast cells is far from nonspecific, while also disparate to that by HMGB2. Following integration of its binding positions with genome-wide chromosome conformation capture data (Hi-C), we found that HMGB1 demarcates the boundaries of a subset of topologically-associating domains (TADs; Dixon *et al*, 2012; Nora *et al*, 2017) and loop domains (Rao *et al*, 2014). This topological contribution is eliminated upon senescence entry, and knockdown/overexpression experiments show that HMGB1 controls the expression of genes that are central to the senescent program and embedded in these domains. Critically, as HMGB1 was proposed to have RNA-binding capacity (Castello *et al*, 2016), we used sCLIP (Kargapolova *et al*, 2017) to show it also influences the availability of senescence-relevant mRNAs. This occurs via a network of RNA-binding factors interacting with HMGB1. In summary, using replicative senescence as a model, we characterize the multiple roles of HMGB1 that converge on the coordination of chromatin and RNA control for SASP regulation.

Results

Senescence entry is marked by HMGB1 nuclear loss and secretion

To investigate the nuclear roles of HMGB1 across cellular contexts, we used primary human umbilical vein endothelial cells (HUVECs) and fetal lung fibroblasts (IMR90) that are of distinct developmental origins and have disparate gene expression programs. We defined an early-passage (proliferative) state and a late-passage (senescent) state by combining β -galactosidase staining, cell cycle staging by FACS, and MTT proliferation assays (**Fig. 2.1A**), as well as a “senescence clock” based on the methylation of six CpGs (Zirkel *et al*, 2018). Next, we used RNA-seq data from proliferating and senescent HUVEC and IMR90 (from two different donors/isolates) to look into changing mRNA levels of chromatin factors. Significant and convergent changes between the two cell types involved strong suppression of histone chaperones, heterochromatin-, Polycomb- and lamina-associated proteins, centromere components, cohesin and condensin complexes, as well as all HMGB/N-family proteins (**Fig. 2.1B**). Many of these factors were consistently suppressed also at the protein level (**Fig. 2.1C**; Davalos *et al*, 2013; Shah *et al*, 2013; Rai *et al*, 2014; Zirkel *et al*, 2018). We focused on HMGB1 due to its conservation, high nuclear abundance (Thomas & Stott, 2012, **Fig. 2.1D**) and key role in SASP induction (Davalos *et al*, 2013), but also mostly due to its elusive role on chromatin, especially in respect to spatial chromosome organization.

Immunodetection in early- and late-passage cells documented a >50% decrease in HMGB1 nuclear levels in the heterogeneous senescence-entry populations of HUVEC or IMR90 (**Fig. 2.1E**). HMGB1 nuclear depletion was most dramatic in the enlarged senescent nuclei of either cell type, while smaller nuclei remained largely unaffected. FACS-sorting IMR90 based on light scattering allowed enrichment for cell populations with enlarged nuclei (i.e., ~70% of cells had larger than average nuclei, with >35% being >1.5-fold larger; **Appendix Fig. 2.S1A**). This showed that enlarged nuclei lacking HMGB1 almost invariably represent senescent cells and harbor reduced levels of H3K27me3, a mark of facultative heterochromatin – an effect which would otherwise be masked (**Fig. 2.1C** and **Appendix Fig. 2.S1B,C**). Last, we showed that it is these larger cells that secrete HMGB1, but not HMGB2, into the growth medium to presumably contribute to paracrine senescence (**Appendix Fig. 2.S1D**).

The senescence program is predominantly transcriptionally-driven

Despite strong changes documented by RNA-seq, it is still not known to which extent the senescent program is implemented via changes at the transcriptional or the translational level. To address this, we generated matching mRNA-seq, Ribo-seq and whole-cell proteomics data from proliferating and senescent IMR90 in biological triplicates. Comparative analysis of mRNA-seq and Ribo-seq data showed that essentially all significant changes at the level of mRNA translation were underlied by equivalent changes in transcript availability (**Fig. 2.1F**). Few transcripts (~800) showed increased translation that counteracted transcriptional suppression (e.g., *TNFRSF19*, *HMGN2*, *LMNB2*) or the converse (e.g., *IL12A*, *CDKN1A*, *CDKN2B*). Gene set enrichment analysis of these two subgroups (with a “buffer” ratio of at least $\log_2 0.6$) showed that mRNAs translationally upregulated while transcriptionally suppressed are linked to the formation and secretion of endosomal vesicles (and thus to HMGB1 release into the extracellular milieu; **Appendix Fig. 2.S1E**), as well as to cell cycle regulation. On the other hand, transcripts translationally downregulated but transcriptionally upregulated associate with ribosome complex formation and translation, and surprisingly also with RNA binding and mRNA catabolism (**Appendix Fig. 2.S1F**).

Similar analysis of mRNA-seq against whole-cell proteome data verified that the vast majority of proteins with significantly altered levels in senescence were similarly regulated transcriptionally (**2.Fig. 1G**). Genes linked to the hallmark GO terms and pathways of senescence entry were convergently up- (e.g., ECM organization, lysosome) or downregulated at both the mRNA and protein levels (e.g., cell cycle, DNA conformation change, RNA processing; **2.Fig. 1G,H**). Curiously, several processes relevant to the SASP (e.g., TNF α /NF- κ B signaling) appeared regulated by a combination of higher transcription and diminished protein availability (**2.Fig. 1H** and **Appendix Fig. 2.S1G**). Our findings correlate with Ribo-seq analyses in a model of

oncogene-induced senescence (Loayza-Puch *et al*, 2013), which also highlighted that cell cycle arrest is mainly driven by transcriptional changes.

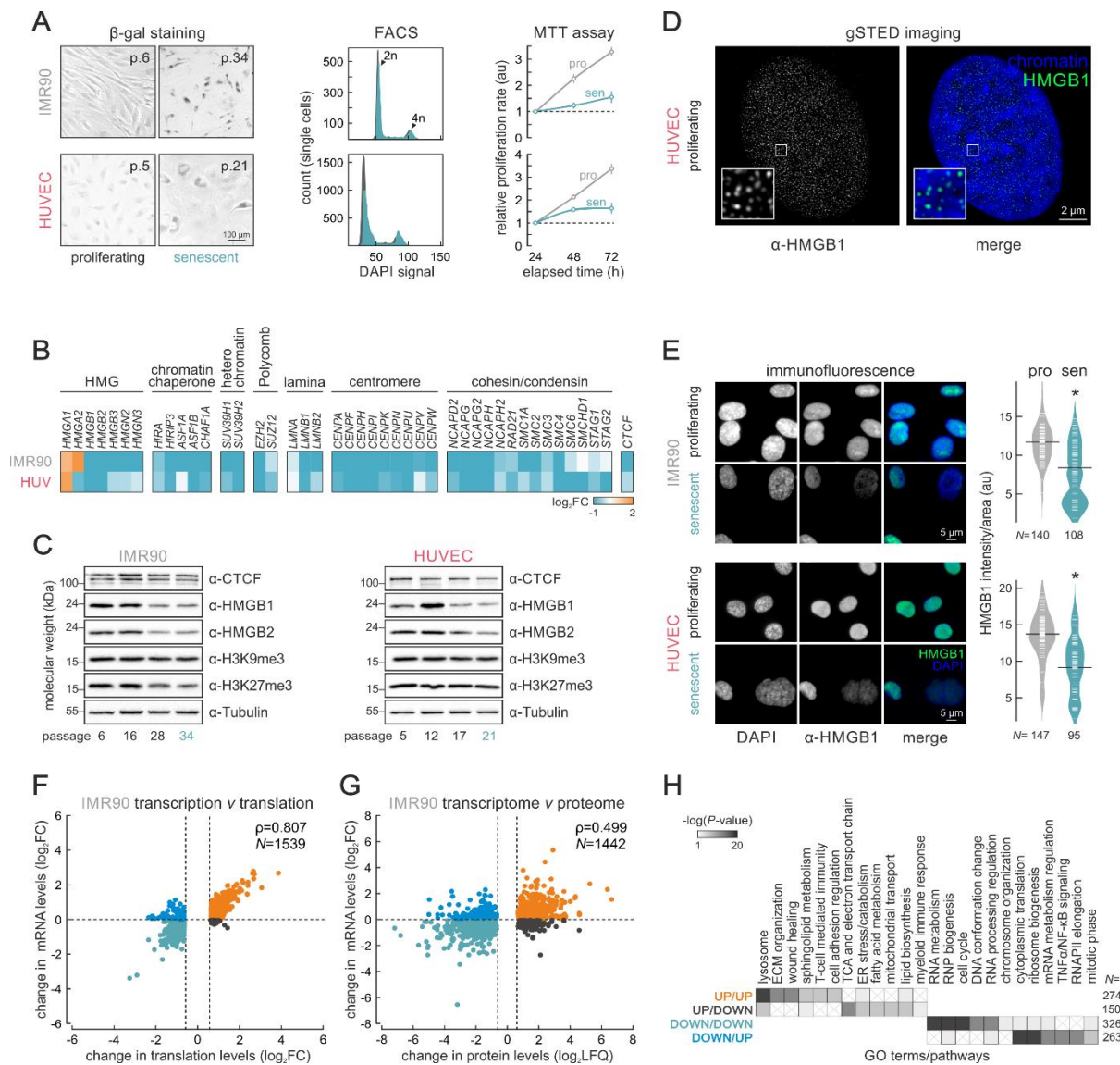


Figure 2.12. Senescence entry by human cells is mostly transcriptionally driven.

(A) Proliferating and senescent IMR90 and HUVECs assayed for β -galactosidase activity (*left*), cell cycle profiling via FACS (*middle*) and proliferation via MTT assays (*right*). (B) Heatmaps showing changes in gene expression levels upon senescence (\log_2FC) of genes encoding selected chromatin-associated factors. For each gene shown, statistically significant expression changes were recorded in at least one cell type. (C) Western blots showing changing protein levels on the path to senescence in IMR90 and HUVECs. Passage 6 cells represent the proliferating state, and passages 21 and 34 the senescent state for HUVECs and IMR90, respectively.

(D) Super-resolution (gSTED) imaging of HMGB1 distribution in proliferating HUVEC nuclei counterstained with DAPI. Bar: 2 μ m. (E) Representative immunofluorescence images of IMR90 and HUVECs (*left*) showing reduced HMGB1 levels in senescent nuclei; bean plots quantify this reduction (*right*; N is the number of cells analyzed per each condition/cell type). Bars: 5 μ m. *: $P < 0.01$; Wilcoxon-Mann-Whitney test. (F) Scatter plots showing correlation between RNA-seq (transcription) and Ribo-seq (translation; \log_2) in proliferating and senescent IMR90. Pearson's correlation values (ρ) and the number of genes in each plot (N) are also shown. (G) As in panel F, but correlating RNA-seq and whole-cell proteome changes. (H) Heatmap showing GO terms/pathways associated with the gene subgroups from panel G (color-coded the same way). The

HMGB1 binds active chromatin loci in a cell type-specific manner

Capturing HMGB proteins on chromatin has proven challenging, because their HMG-box DNA-binding domains are not compatible with standard formaldehyde fixation (Pallier *et al*, 2003; Teves *et al*, 2016). Here, we employed a tailored dual-crosslinking ChIP strategy to efficiently capture HMGB1 bound to its cognate sites genome-wide in both proliferating HUVECs and IMR90 (Fig. 2.2A and Appendix Fig. 2.S2A; see **Methods** for details). HMGB1 binding is restricted to regions marked by H3K27ac (Appendix Fig. 2.S2B), the vast majority of which are promoters and gene bodies of active genes (92% and 73% of 2038 and 1611 peaks in HUVECs and IMR90, respectively; Fig. 2.2B). The two cell types share >550 HMGB1 peaks, which is more than would be expected by chance, but still >1000 cell type-specific peaks remain that can be justified by the very different gene expression programs of HUVECs and IMR90 (Fig. 2.2C). We also used senescent IMR90 to ask whether residual HMGB1 (<30% that of proliferating cells; Appendix Fig. 2.S2C) is redirected to particular loci. The fact that we discovered just 44 peaks in our ChIP-seq replicates argue against this possibility (Fig. 2.2A-C).

Although HMGBs are assumed to bind chromatin in a non-specific manner (Stros, 2010), our *de novo* motif analysis in DNase I hypersensitive “footprints” (ENCODE Project Consortium, 2012) under HMGB1 peaks revealed that HMGB1 prefers similar G/A-rich motifs in IMR90 and HUVECs (Appendix Fig. 2.S2D). We also surveyed these accessible footprints for known transcription factor (TF) motifs to infer co-binding complexes or TFs that might replace HMGB1 on senescent chromatin. Of the many motifs discovered, we focused on those from factors differentially-regulated upon senescence entry. In both HUVECs and IMR90, E2F-family motifs were enriched and these cell cycle-regulating TFs (e.g., E2F2 and E2F6) are strongly downregulated in senescence. In contrast, motifs of pro-inflammatory and senescence-activated TFs (e.g., NFKB2 and STAT1) were enriched therein and could presumably take over these positions to facilitate inflammatory stimulation and the SASP (Appendix Fig. 2.S2E).

To identify genes regulated by HMGB1, we crossed ChIP-seq with RNA-seq data in both cell types. Looking at genes directly bound by HMGB1 (i.e., excluding intergenic peaks >10 kbp from TSSs), a first observation was that these are involved in various signaling cascades (most notably the TNF α one), as well as in cellular senescence (**Appendix Fig. 2.S2F**). Interestingly, HMGB1 consistently associated with more up- rather than downregulated genes in both HUVECs and IMR90. Of those, downregulated genes were involved cell cycle transitions, while upregulated ones were related to ECM organization, cell adhesion, and inflammatory signaling (**Appendix Fig. 2.S2G,H**). Together, this analysis demonstrates that HMGB1 binds active loci relevant for the induction of the senescence gene expression program.

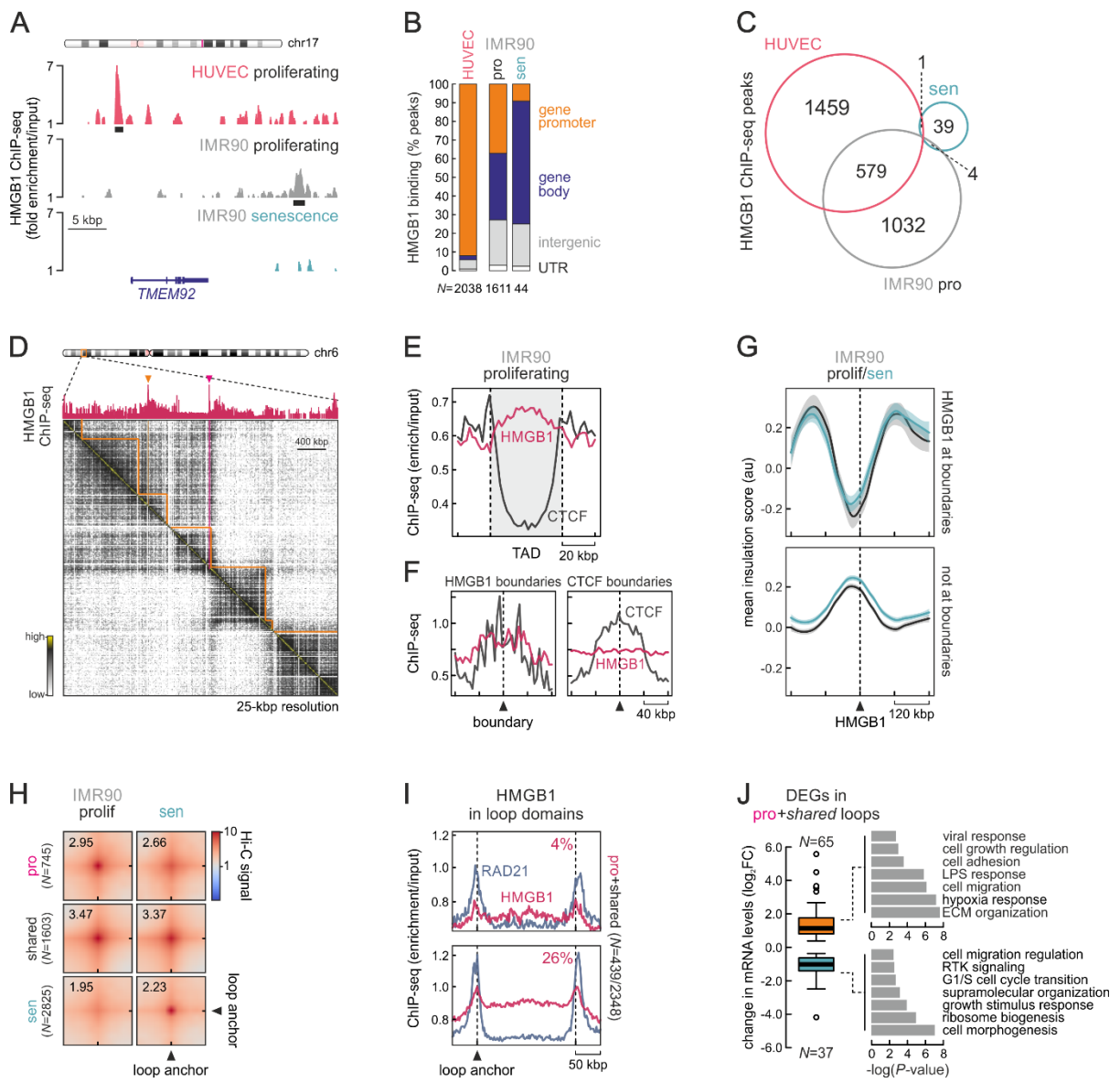


Figure 2.13. HMGB1 binds active genomic loci and demarcates a subset of TAD and loop domains.

(A) Genome browser view of HMGB1 ChIP-seq (normalized to input; mean from two replicates) from proliferating HUVEC (*red*) or proliferating/senescent IMR90 (*grey/green*) in the *TMEM92* locus. (B) Bar graphs showing the genomic distribution of HMGB1 ChIP-seq peaks in HUVEC and IMR90. The number of peaks (*N*) analyzed per each cell type is indicated. (C) Venn diagram showing HMGB1 ChIP-seq peaks shared between proliferating HUVEC (*red*) or proliferating/senescent IMR90 (*grey/green*) data. *: $P < 0.001$; more than expected by chance, hypergeometric test. (D) Exemplary Hi-C heatmap for a subregion of IMR90 chr6 aligned to HMGB1 ChIP-seq; peaks at TAD boundaries (*orange lines*) are indicated (*red arrowheads*). (E) Line plots showing normalized HMGB1 (*red*) and CTCF ChIP-seq signal (*dark grey*) along TADs ± 20 kbp from proliferating IMR90. (F) Line plots showing normalized HMGB1 (*red*) and CTCF ChIP-seq signal (*dark grey*) in the 160 kbp around HMGB1- (*left*) or CTCF-marked TAD boundaries (*right*). (G) Line plots showing mean normalized insulation scores calculated from proliferating (*black*) and senescent IMR90 Hi-C data (*blue*) in the 240 kbp around HMGB1 peaks residing at TAD boundaries (*top*) or not (*bottom*). (H) Aggregate plots showing 20 kbp-resolution Hi-C signal for proliferating (*top*), senescent-specific (*bottom*) or shared loops (*middle*). The number of loops in each category (*N*) is indicated. (I) Line plots showing normalized HMGB1 (*red*) and Rad21 ChIP-seq signal (*grey*) along proliferating (*top*) or shared loop domains (*bottom*) ± 50 kbp from proliferating IMR90. The percentage of HMGB1 peaks residing at loop anchors is indicated. (J) Box plot (*left*) showing significantly up/downregulated IMR90 genes harbored inside the 607 loops from panel H. The number of genes in each group (*N*) is indicated; center lines represent medians, box limits indicate the 25th and 75th percentile, whiskers extend 1.5 times the 25th-75th interquartile range, and outliers are represented by dots. Bar plots (*right*) show Go terms associated with these up/downregulated genes. The number of genes in each subgroup (*N*) are indicated.

HMGB1 marks a subset of senescence-regulated domain boundaries

TADs are often considered the building blocks of chromosomes (Beagan & Phillips-Cremins, 2020), their boundaries representing sites of local insulation for spatial chromatin interactions. TAD boundaries are often marked by the presence of bound CTCF and/or active gene promoters (Dixon *et al*, 2012; Nora *et al*, 2017). We recently showed that a considerable number of TAD boundaries in proliferating human cells are marked by HMGB2, and that these boundaries are remodeled upon senescence entry and nuclear HMGB2 loss (Zirkel *et al*, 2018). We now reasoned that HMGB1 may also function similarly. To test this, we used Hi-C data from proliferating IMR90 and found that the majority of HMGB1 ChIP-seq signal (normalized to input) resides inside TADs (called at 40-kbp resolution; **Fig. 2.2D,E**). Nevertheless, $\sim 10\%$ of HMGB1 peaks reside at TAD boundaries at positions not overlapping CTCF. Conversely, CTCF-marked TAD boundaries show no enrichment for HMGB1 (**Fig. 2.2F**). Upon senescence entry, TAD boundaries that lose HMGB1 demarcation exhibit a reduction in their insulation score indicative of 3D interaction reshuffling (**Fig. 2.2G**), without losing their insulatory character completely though.

High-resolution Hi-C studies showed that human chromosomes are populated with sub-TAD loop domains anchored at CTCF/cohesin-bound sites (Rao *et al*, 2014). We mapped loop domains in our proliferating and senescent Hi-C data to discover that a considerable number of loops either weaken ($N=745$) or emerge upon IMR90 senescence entry ($N=2825$), while 1603 loops remain invariant (**Fig. 2.2H**). Remarkably, 4% of HMGB1 peaks mark the anchors of proliferating-specific loops and 26% mark invariant loops anchors; this constitutes almost 1/5 of all loop domains in proliferating IMR90 (**Fig. 2.2I**). Given the loss of HMGB1 in senescence, we reasoned that genes within these loop domains would be deregulated. Indeed, these loops include 105 significantly-regulated genes, 2/3 of which are upregulated in senescence, consistent with HMGB1-based regulation (see **Appendix Fig. 2.S3A**), and associate with pathways central to senescence induction, especially inflammatory activation and cell cycle arrest (**Fig. 2.2J**). These effects held true also when analyzing HMGB1 ChIP-seq and Hi-C data from HUVECs (**Appendix Fig. 2.S3A-F**). Our analyses suggest that the topological contribution of HMGB1 is relevant for gene regulation on the path to senescence.

Spatial TAD co-association reveals functional specialization of chromosome domains

Given that HMGB1-marked boundaries are not fully abolished upon its senescence-induced loss, but chromosomes undergo large-scale changes upon replicative senescence entry (Zirkel *et al*, 2018), which are further accentuated in “deep” senescence (Criscione *et al*, 2016), we looked into how TADs in each chromosome associate with one another in higher-order “metaTAD”-like conformations (Fraser *et al*, 2015). We employed an unsupervised approach inspired from “topologically-intrinsic lexicographic ordering” (TiLO; Johnson, 2012). Here, TADs are treated as nodes in a clustered spatial network inferred on the basis of inter-TAD Hi-C interactions that is then tested for robustness via iterative network slicing (**Fig. 2.3A**; see **Methods** for details).

We applied TiLO to TADs derived from proliferating and senescent HUVEC and IMR90 Hi-C data and visualized the final output using Chrom3D (Paulsen *et al*, 2018) that also takes into account inter-chromosomal contacts and interactions with the lamina to render chromosome models (for an example see **Fig. 2.3B** and **Appendix Fig. 2.S4B**). Although the fraction of smaller and larger TAD clusters emerging upon senescence differs between HUVECs and IMR90 (**Appendix Fig. 2.S4A**), there is an overall trend for larger and fewer clusters in senescent chromosomes (**Fig. 2.3C** and **Appendix Fig. 2.S4C,D**). This is consistent with the general spatial chromatin compaction observed upon senescence entry (Criscione *et al*, 2016; Zirkel *et al*, 2018) and the fact that 43% of HUVEC and 26% of IMR90 TADs were found to merge into larger

ones in Hi-C data (Zirkel *et al*, 2018); see example in **Appendix Fig. 2.S4E**), and our TiLO-Chrom3D combination recapitulates these organization changes (**Appendix Fig. 2.S4F**).

Closer inspection of the distribution of HMGB1 peaks along TiLO clusters (**Appendix Fig. 2.S4A**) and of Hi-C/ChIP-seq data showing HMGB1-rich domains appearing depleted of CTCF loops (**Fig. 2.3A**), led us to query the relative position of HMGB1-marked TiLO clusters in our models. We grouped clusters by the number of HMGB1 peaks they carry as follows: “nonHMGB1” clusters carrying zero peaks, “low” carrying 1-2, “medium” carrying 3-4, and “high” carrying 5 or more peaks; most IMR90 and HUVEC chromosomes do carry such HMGB1 “high” hotspot clusters (e.g., IMR90 chr20 does not). Strikingly, in both HUVEC and IMR90 chromosomes, clusters with increasing numbers of HMGB1 peaks are positioned progressively closer together. “High” HMGB1 hotspot clusters are closer to one another than to any other cluster (**Fig. 2.3C** and **Appendix Fig. 2.S4C,D**). This specialized spatial clustering is reshuffled in senescent chromosomes, although former hotspot clusters remain on average closer together (**Fig. 2.3C** and **Appendix Fig. 2.S4C,D**). To find out whether this spatial clustering also correlates with some special functional output, we characterized the differentially-expressed genes they harbor. Using IMR90 as an example, TiLO TAD clusters not marked by HMGB1 and HMGB1-hotspot ones contain a large fraction of all genes differentially-expressed upon senescence entry (i.e., in total ~50% of up- and >20% of downregulated genes). However, nonHMGB1 clusters uniquely harbor genes involved in replication, RNA localization, and p53 signaling, where HMGB1-hotspot clusters are notably rich in genes relevant to the SASP and pro-inflammatory signaling (**Fig. 2.3D**). Of the 313 upregulated genes in these hotspot clusters, >60 were linked to SASP production and 46 were not only upregulated in senescent IMR90 but also discovered as secreted SASP factors in fibroblasts via proteomics (Basisty *et al*, 2020); more than half of these genes directly relied on HMGB1 loss for induction (**Fig. 2.3E**). This also held true when HUVEC HMGB1 “hotspot” TiLO clusters were analyzed (**Appendix Fig. 2.S4G**). In summary, TiLO has the power to identify spatial co-associations of TADs, explaining the functional specialization of different genomic domains demarcated by HMGB1.

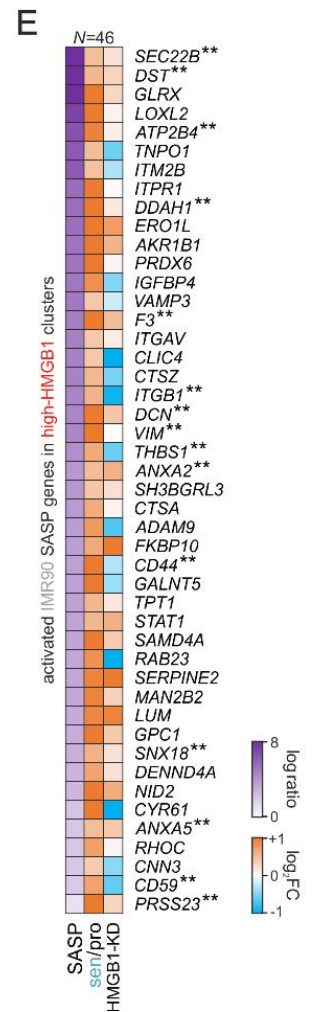
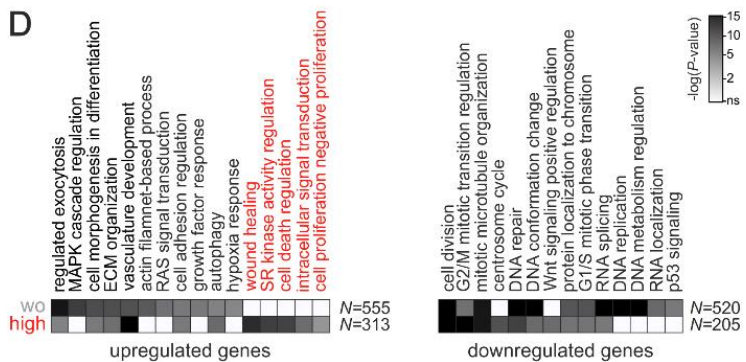
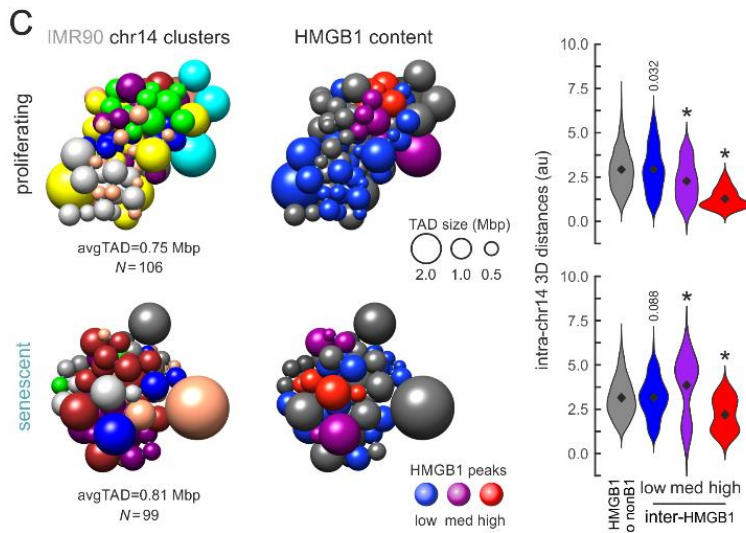
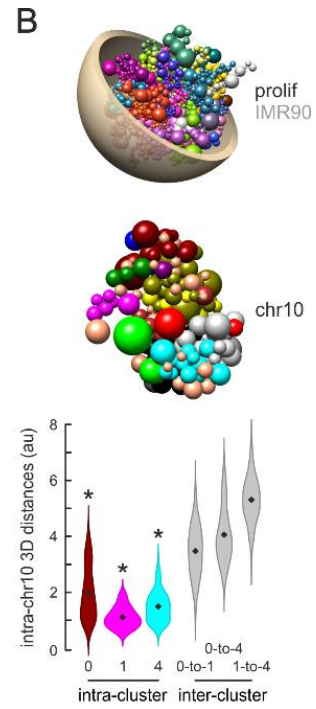
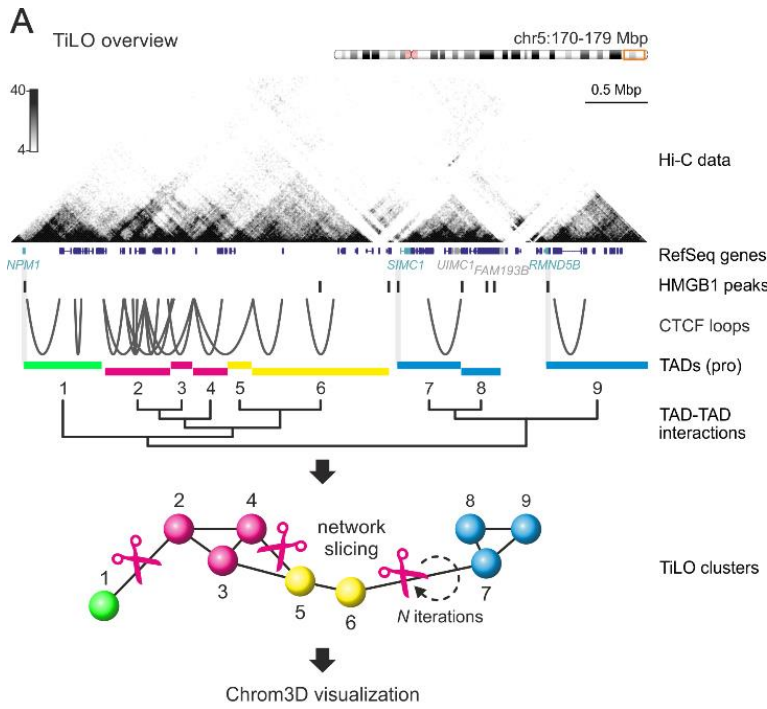


Figure 2.14. HMGB1 marks co-associating TAD clusters harbouring SASP-related genes.

(A) Overview of TiLO using a subregion of HUVEC chr5 as an example. TADs along each chromosome are treated as nodes in an interaction network and inter-TAD Hi-C signal is used to infer network connections. Inferred connections are then sliced and network robustness is assessed iteratively to obtain the final clustering. TAD clusters are visualized in the space of the nucleus using Chrom3D. (B) Chrom3D visualisation of the whole genome (*top*) and of TAD clustering in chr10 from proliferating IMR90 (*middle*); each sphere represents one TAD. Violin plots (*bottom*) show 3D distances amongst TADs in three randomly-selected clusters (0, 1, and 4) or between TADs from different clusters. *: significantly different to inter-cluster distances, Wilcoxon-Mann-Whitney test. (C) Chrom3D visualisation of TAD clustering in chr14 from proliferating (*top*) and senescent IMR90 (*bottom*). TADs (*spheres*) are coloured by the cluster they belong to (*left*) or according to their HMGB1 ChIP-seq content (*middle*; grey – zero peaks, blue – 1 or 2 peaks, purple – 3 or 4 peaks, red – 5 or more peaks). Violin plots (*right*) show 3D distances amongst TADs in each subgroup or between HMGB1-containing and non-containing TADs. *: significantly different to inter-cluster distances, Wilcoxon-Mann-Whitney test. (D) Heatmaps showing GO terms associated with differentially-expressed genes in two TAD groups from panel C; SASP-related GO terms are highlighted. The number of genes (*N*) behind each heatmap are indicated. (E) Heatmap showing protein (SASP, from <http://www.saspatlas.com/>; log fold ratio) and gene expression levels (\log_2FC) of IMR90 SASP-related genes embedded in high-HMGB1 (red) TADs like those in panel (C) **: genes bound by HMGB1 in ChIP-seq data are more than expected by chance; $P > 0.001$, hypergeometric test.

HMGB1 depletion underlies induction of the senescence program

It was previously shown that transduction of WI-38 human fibroblasts with shRNAs against *HMGB1* suffices for senescence induction (Davalos *et al*, 2013). Here, we treated HUVECs with self-delivering siRNAs targeting *HMGB1*. This led to a ~2-fold reduction of HMGB1 protein and RNA levels within 72 h (**Appendix Fig. 2.S5A**) accompanied by a doubling of β -galactosidase and p21-positive cells in knockdown populations, but by only small changes in nuclear size and BrdU incorporation (**Appendix Fig. 2.S5B-E**). To obtain stronger effects, we turned to IMR90 where standard siRNA transfections allowed for a >10-fold decrease in HMGB1 protein and RNA levels, as well as to changing expression of senescence markers like *CDKN1B* and *HMGA1* (without affecting *HMGB2*; **Fig. 2.4A**). Analysis of RNA-seq data from siRNA-treated and control IMR90 returned ~900 up- and >950 downregulated genes upon *HMGB1*-knockdown (**Fig. 2.4B**). GO term and gene set enrichment analyses showed that the upregulated genes could be linked to pro-inflammatory signalling, while downregulated ones associated with changes in chromatin organization, transcriptional silencing, and the p53 pathway, all hallmarks of senescence entry (**Fig. 2.4C,D**). Looking for direct HMGB1 targets in knockdown data, we identified 104 up- and 121 downregulated genes bound by HMGB1; they showed mean ChIP-seq signal enrichment at their 5' or 3' ends. Reassuringly, upregulated genes were linked to NF- κ B and p38 signaling, as well as exocytosis; downregulated ones were linked to growth signalling, chromatin reorganization, and cell cycle arrest (**Appendix Fig. 2.S5F**). Interestingly, comparison of significant gene

expression changes upon *HMGB1*-knockdown to those from senescence entry cells revealed poor correlation (**Appendix Fig. 2.S5G**). Genes downregulated in both knockdown and senescent cells were notably linked to RNA splicing, processing, and cleavage. Genes downregulated in senescence and upregulated upon *HMGB1*-knockdown were relevant to RNA metabolism, but also stress response and apoptosis (**Appendix Fig. 2.S5G**). Thus, *HMGB1* appears to control a specific leg of the senescence gene expression program.

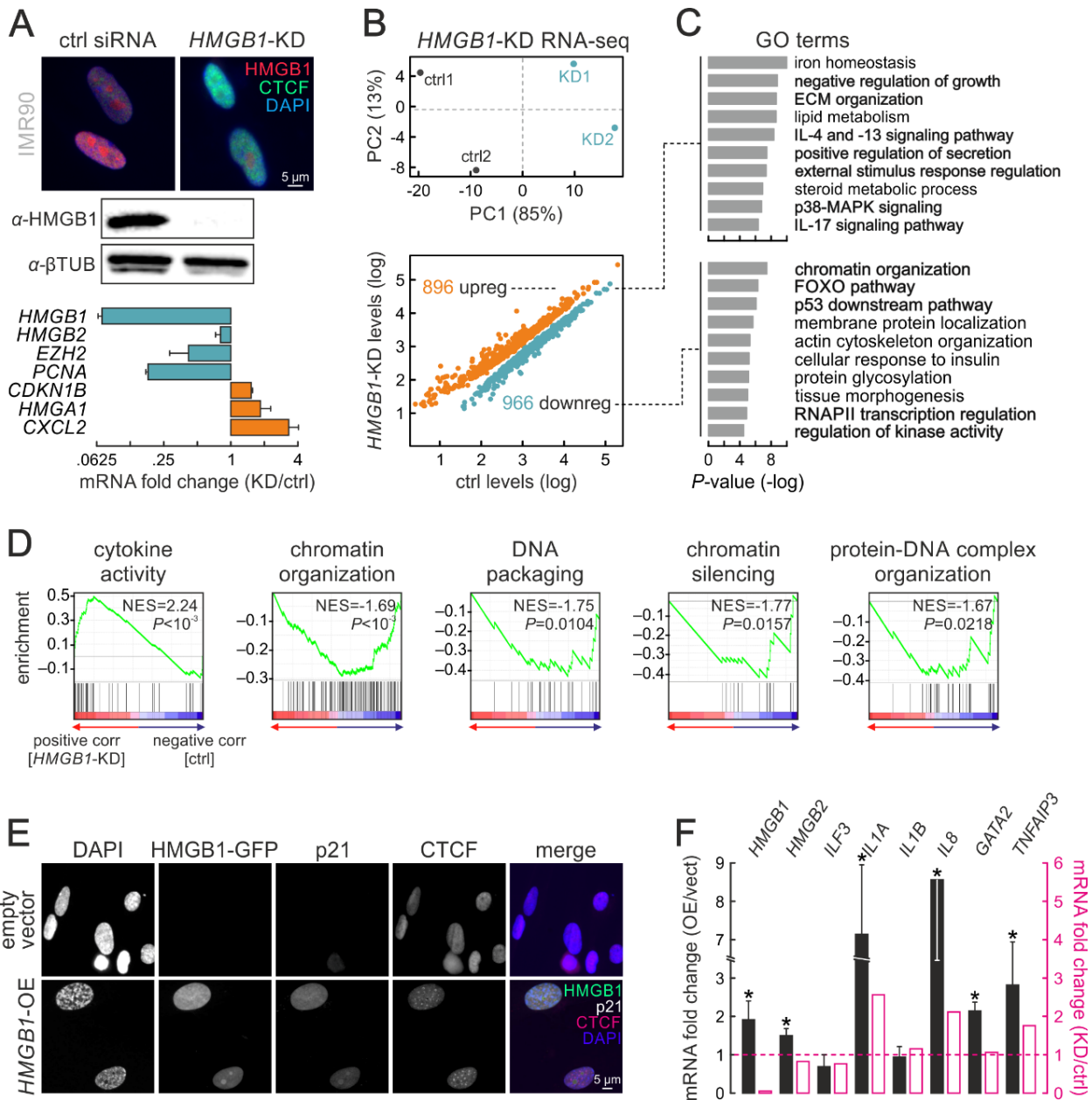


Figure 2.15. Modulating HMGB1 expression induces senescence-specific gene expression changes.

(A) Immunofluorescence (*top*), western blot (*middle*) and RT-qPCR analyses (*bottom*; mean fold-change \pm S.D. from two biological replicates) confirm *HMGB1* knockdown in IMR90. (B) PCA analysis plot (*top*) of control (*black*) and *HMGB1*-knockdown replicates (*blue*). Scatter plot (*bottom*) showing significantly up- (>0.6 log₂-fold change; *orange*) or downregulated genes (<-0.6 log₂-fold change; *green*) upon *HMGB1* knockdown. (C) Bar plots showing GO terms associated with the up/downregulated genes from panel B and their enrichment *P*-values ($-\log$; *right*). (D) Gene set enrichment analysis of *HMGB1*-KD data. Normalized enrichment scores (NES) and associated *P*-values for each gene set are shown. (E) Representative images of IMR90 overexpressing *HMGB1*-GFP, immunostained for p21 and CTCF, and counterstained with DAPI. IMR90 transfected with empty vectors provide a control. Bar: 5 μ m. (F) Bar plots showing RT-qPCR data (mean mRNA fold-change \pm S.D., from two biological replicates) for selected genes in *HMGB1*-overexpressing compared to control IMR90. The mean fold-change for each mRNA from *HMGB1*-knockdown RNA-seq data is also shown for comparison (*magenta bars*). *: $P < 0.01$; unpaired two-tailed Student's *t*-test.

We complemented our knockdown experiments with *HMGB1* overexpression. We expressed an *HMGB1*-GFP fusion protein in IMR90 using a doxycycline-inducible PiggyBac system (**Fig. 2.4E**). We selected for transfected cells using antibiotics, but refrained from generating single cell-derived populations to gauge heterogeneity arising from differences in integration sites. Within <10 h of overexpression induction, nuclear accumulation of *HMGB1* in a subset of the population led to a strong increase in p21 signal, to the emergence of characteristic DAPI-dense foci and, in some cases, to the formation of senescence-induced CTCF clusters (Zirkel *et al*, 2018; **Fig. 2.4E**). All these are hallmarks of senescence entry, and agree with changes in SASP gene expression following *HMGB1* overexpression (**Fig. 2.4F**). Thus, increased *HMGB1* levels can also drive senescence entry, most likely via reinforced paracrine signaling.

HMGB1 binds and regulates a discrete set of senescence-related mRNAs

The number of protein-coding loci bound by *HMGB1* and regulated upon senescence entry and *HMGB1* knockdown does not explain the full extent of the senescence program (**Appendix Figs 2.S2F,G and 2.S5F**). To address this disparity, we pursued the idea that *HMGB1* also acts as an RNA-binding protein, as was suggested by recent classifications of the human proteome (Castello *et al*, 2016; Trendel *et al*, 2019). This idea was reinforced by our analysis of Ribo- and RNA-seq data (**Appendix Figs 2.S1F and 2.S5G**), and by our cataloguing of *HMGB1* protein partners in proliferating IMR90. Mass spectrometry revealed a broad range of RBPs and splicing regulators co-immunoprecipitating with *HMGB1*, in addition to the expected chromatin-associated proteins (**Fig. 2.5A**) and despite co-IPs being performed after RNase A treatment of samples to avoid indirect interactions. In the end, $\sim 40\%$ of *HMGB1* interactors qualify as RBPs (**Fig. 2.5B**).

To study *HMGB1* as an RNA binder, we applied sCLIP (Kargapolova *et al*, 2017) to proliferating IMR90 (**Appendix Fig. 2.S6A,B**). Analysis of two well-correlated replicates (**Appendix Fig. 2.S6C**) provided a

set of 1,773 binding peaks on 866 different transcripts (**Fig. 2.5C**). Reassuringly, HMGB1-bound mRNAs display <14% overlap to HMGB1-bound genes in ChIP-seq (**Appendix Fig. 2.S6D**). Thus, crosslinking of HMGB1 to RNA cannot simply be the byproduct of its binding to transcriptionally-active chromatin loci.

On RNA, HMGB1 mostly binds exons and 5'/3' UTRs, but also a substantial number of non-coding RNAs (**Fig. 2.5C,D**). HMGB1-bound sites present the same hexameric 5'-NMWGRA-3' (M=A/C, W=A/T, R=A/G) motif irrespective of the predicted folding of the underlying RNA (**Fig. 2.5C** and **Appendix Fig. 2.S6E**). Much like what we observed in ChIP-seq, HMGB1 binds ~3-fold more transcripts that are up- rather than downregulated upon senescence. Upregulated mRNAs associated with senescence-related GO terms like ECM organization, wound healing and negative regulation of cell proliferation, while downregulated ones mostly with processes like RNA splicing, RNA-/miRNA-mediated gene silencing, or histone remodeling and deacetylation (**Appendix Fig. 2.S6F**). After crossing sCLIP with RNA-seq data from *HMGB1*-knockdown IMR90, 56 up- and 97 downregulated mRNAs were also bound by HMGB1. Curiously, upregulated transcripts showed a slight bias for HMGB1 binding in their 5' ends, while downregulated ones showed stronger 3' end binding (**Appendix Fig. 2.S6G,H**). Consistent with all previous observations, upregulated mRNAs could be linked to processes like ECM organization and inflammatory activation, while downregulated mRNAs to splicing and chromatin reorganization (**Appendix Fig. 2.S6H**).

As RBPs are known to affect the splicing patterns of their target transcripts, we examined how splicing is altered upon senescence entry by IMR90 using Whippet (Sterne-Weiler *et al*, 2018). We documented ~4,000 significant changes in mRNA splicing, the majority of which concerned alternative usage of transcription start and polyadenylation sites (>80% of cases; **Fig. 2.5E**), consistent with recent observations in senescing HUVECs (Shen *et al*, 2019). This trend remained essentially invariable when we interrogated the kind of splicing changes occurring upon *HMGB1* knockdown or to HMGB1-bound and differentially-spliced mRNAs (**Fig. 2.5E**). A large fraction (~26%) of splicing events in mRNAs bound by HMGB1 in sCLIP overlap events seen upon both senescent and HMGB1-KD IMR90 (and >95% overlap the events seen in at least one of the two conditions; **Fig. 2.5F**). Differentially-spliced mRNAs shared by senescence entry and *HMGB1*-knockdown encode factors linked to senescence-regulated processes like cell cycle and cell growth regulation, and the p53 pathway (**Fig. 2.5G**), while those additionally bound by HMGB1 show a bias towards processes linked to RNA splicing and biogenesis (**Fig. 2.5H**). Thus, the nuclear loss of HMGB1 correlates with changes to the cell's transcriptome processing.

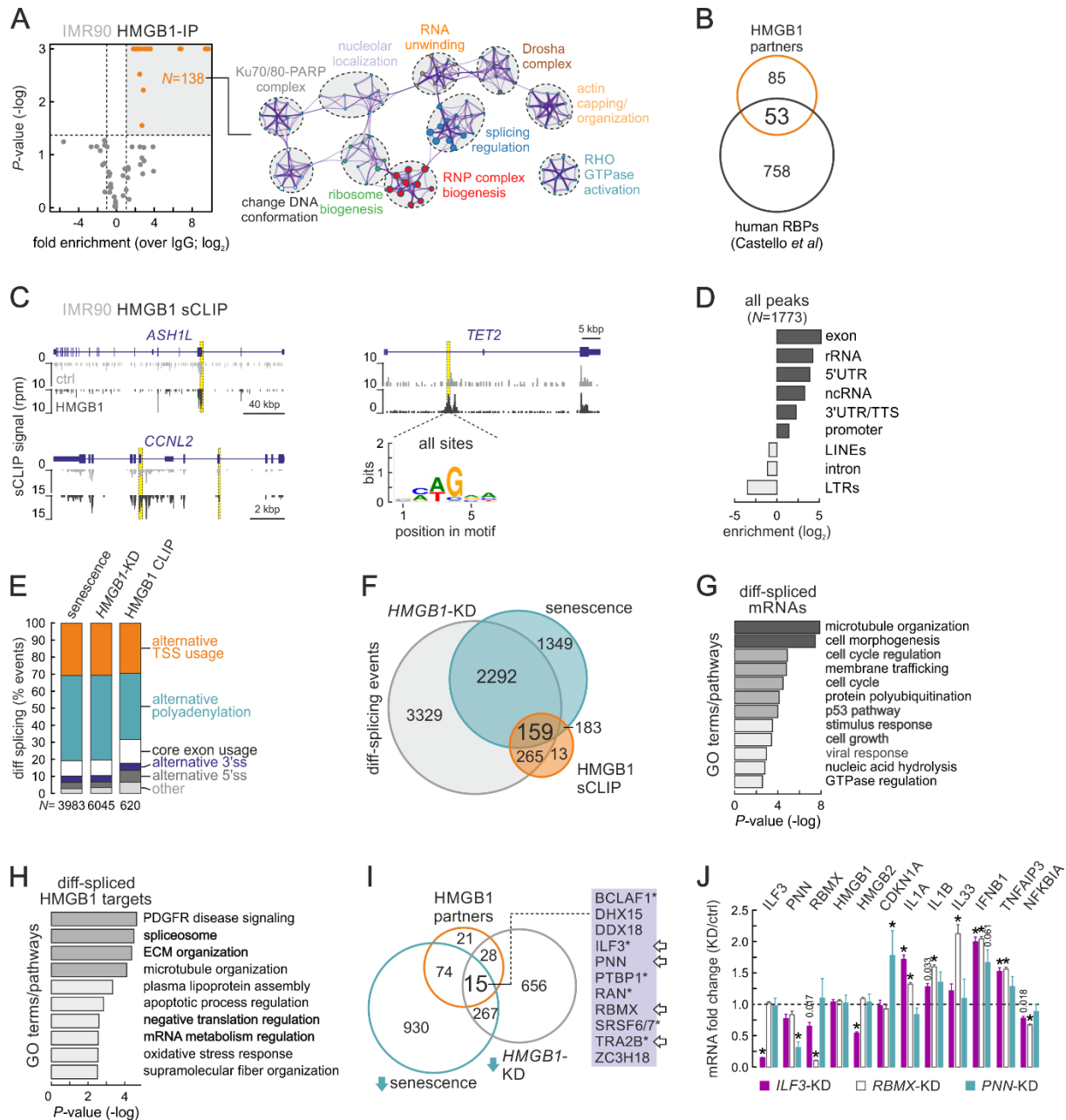


Figure 2.16. HMGB1 binds specific mRNAs and its loss affects splicing.

(A) Volcano plot (*left*) showing mass-spec data for proteins co-immunoprecipitating with HMGB1. Statistically-enriched HMGB1 interactors (*orange dots*) associate with the GO terms/pathways illustrated in the network analysis (*right*); node size reflects the number of proteins it includes; proteins are listed in Dataset EV4). (B) Venn diagram showing 1/3 of HMGB1 interactors classifying as RNA-binding proteins (according to Castello *et al*, 2016). (C) Genome browser views showing HMGB1 sCLIP data (*black*) along the *ASH1L*, *CCNL2* and *TET2* loci; input tracks (*grey*) provide background levels. *: significantly-enriched peaks. The consensus motif for HMGB1 binding on RNA is also shown (*bottom right*). (D) Bar graphs showing genomic distribution of HMGB1 RNA-bound peaks (\log_2 enrichment).

(E) Bar plots showing relative occurrence of differential-splicing events in IMR90 undergoing senescence (*left*), in HMGB1-knockdown IMR90 (*middle*) or in HMGB1-bound mRNAs (*right*). The number of bound mRNAs (*N*) analyzed is indicated below each bar. (F) Venn diagram showing differential-splicing events shared between conditions from panel F. (G) Heatmaps showing GO terms/pathways associated with differentially-spliced mRNAs shared between senescence entry and HMGB1-knockdown from panel F. (H) As in panel G, but for the 75 mRNAs also bound by HMGB1 in sCLIP data. (I) Venn diagram (*left*) showing 15 HMGB1 interacting proteins from panel A are also downregulated upon both senescence entry and *HMGB1*-knockdown in IMR90. Of these, 12 are RBPs, 6 have been implicated in senescence (*asterisks*), and 4 are bound by HMGB1 in sCLIP data (*arrows*). (J) Bar graphs showing mean fold-change of selected mRNAs (over \pm S.D. from two biological replicates) from *ILF3*- (*purple*), *RBMX*- (*white*) or *PNN*-knockdown experiments (*blue*) in proliferating IMR90. *: significantly different to siRNA controls; $P < 0.01$, unpaired two-tailed Student's t-test. Dashed line indicates no change in expression.

In light of this unforeseen role, we revisited the HMGB1 interactome (**Fig. 2.5A**) and found that 15 HMGB1 protein partners were downregulated in senescence, but also upon *HMGB1*-knockdown. Of these, 12 qualified as RBPs (Castello *et al*, 2016), with BCLAF1 (Shao *et al*, 2016), ILF3 (Wu *et al*, 2015; Tominaga-Yamanaka *et al*, 2012), PTBP1 (Georgilis *et al*, 2018), RAN (Cekan *et al*, 2016; Gu *et al*, 2016; Sobuz *et al*, 2019), SRSF7 (Chen *et al*, 2017), and TRA2B (Chen *et al*, 2018a) recently implicated in senescence and SASP regulation (**Fig. 2.5I**, *starred*), while ZC3H18 was identified as a mediator of NF- κ B activation (Gewurz *et al*, 2012). Strikingly, the mRNAs of *ILF3*, *PNN*, *RBMX* and *TRA2B* are also directly bound by HMGB1 in sCLIP data (**Fig. 2.5I**, *arrows*). Query of the STRING database (<https://string-db.org>; Szklarczyk *et al*, 2019) links all these RBPs in a single network relevant to RNA processing (**Appendix Fig. 2.S6I**).

To validate this, we first performed co-immunoprecipitation experiments showing that HMGB1 and ILF3 physically interact (**Appendix Fig. 2.S6J**). We next showed that, although ILF3 is downregulated upon *HMGB1*-knockdown, its protein levels increase upon HMGB1 overexpression. This increase also manifests in a paracrine manner, since cells not carrying the HMGB1-overexpression cassette still show elevated ILF3 (**Appendix Fig. 2.S6K**). In addition, the ILF3 increase coincides with NF κ B translocation into cell nuclei signifying inflammatory activation of cells (**Appendix Fig. 2.S6L**). In line with what was reported for oncogene-induced senescence (Tominaga-Yamanaka *et al*, 2012), ILF3 binds SASP-relevant mRNAs in proliferating IMR90 (**Appendix Fig. 2.S6M**) and its senescence-induced loss (**Appendix Fig. 2.S6N**) can lead to their stabilization for translation. Notably, in our model ILF3 also binds HMGB2 mRNAs (**Appendix Fig. 2.S6M**) and may thus be implicated in a feedforward regulatory loop of this factor. Finally, we went on to knock down *ILF3*, *RBMX*, and *PNN* individually in proliferating IMR90. With the exception of p21 induction, *PNN*-knockdown did not affect *HMGB1/2* or SASP-related mRNA levels. In contrast, *ILF3* or *RBMX* knockdown led to the upregulation of interleukins, *IFNB1* or *TNFAIP3* (**Fig. 2.5J**), all related to inflammatory

activation. Interestingly, *HMGB1* levels did not change in any of these experimental setups, while *ILF3*-knockdown did suppress *HMGB2*. On this basis, we infer that HMGB1 is an upstream regulator of this cascade (given its binding on the mRNAs of all three RBPs; **Fig. 2.5I**), while ILF3 can specifically modulate HMGB2 (which does not occur upon *HMGB1*-knockdown; **Fig. 2.4A**). Thus, HMGB1 is central to a regulatory circuit comprising RBP cofactors that regulate one another, as well as cell-autonomous and paracrine senescence.

Discussion

Unlike the well-documented extracellular role of HMGB1 as a pro-inflammatory stimulus, its positioning along mammalian chromosomes and the transcriptional control it exerts are poorly understood. Here, we assign a multifaceted role to HMGB1 — first as an on-chromatin regulator of active gene loci, and then as a *bona fide* RNA-binding protein regulating a distinct subset of mRNAs. Together, we deduce that HMGB1 acts to “buffer” gene expression levels, its loss from senescent cell nuclei mostly triggering upregulation of target loci and mRNAs. In addition, a subset of HMGB1-bound positions mark TAD/loop domains boundaries. Domains rich in HMGB1 binding peaks preferentially co-associate in 3D nuclear space and harbor SASP-related genes induced upon both senescence entry and *HMGB1* knockdown. This suggests that HMGB1 loss also affects chromatin topology in a manner linked to gene expression changes. Interestingly, 3D chromatin domains rich in HMGB1 are generally depleted of CTCF loops. This implies that these spatial conformations might be incompatible and form by different mechanisms.

In its novel role as a direct RNA-binding regulator, HMGB1 is part of RNP complexes that affect transcript splicing and processing. In fact, a number of HMGB1 partners are RBPs implicated in the regulation of senescence induction and the SASP (Tominaga-Yamanaka *et al*, 2012; Wu *et al*, 2015; Cekan *et al*, 2016; Gu *et al*, 2016; Shao *et al*, 2016; Chen *et al*, 2017; Georgilis *et al*, 2018; Sobuz *et al*, 2019). In addition, the mRNAs of four of these RBPs (*ILF3*, *PNN*, *RBMX* and *TRA2B*) are also direct HMGB1 targets. This renders HMGB1 a central player also in RNA homeostasis. Once HMGB1 is depleted from nuclei, mRNAs necessary for paracrine senescence can become stabilized (Tominaga-Yamanaka *et al*, 2012). This constitutes a remarkable example of a regulatory circuit, where deregulation of genes, transcripts and topology in one cellular compartment (the nucleus) is in direct and quantitative crosstalk with signaling deployed in another (the extracellular space where HMGB1 is secreted). Our observations come to substantiate previous hypotheses of low nuclear HMGB1 titers being necessary for the full deployment of

the SASP (Davalos *et al*, 2013). Thus, the cascade regulating senescence entry has a strong, almost hierarchical, dependency on the nuclear events preceding SASP induction.

Recently, we characterized the function of the sister protein to HMGB1, HMGB2, as regards entry into replicative senescence (Zirkel *et al*, 2018). The loss of HMGB2 appears to precede that of HMGB1, and drives formation of prominent senescence-induced CTCF clusters (SICCs). This affects the spatial architecture of chromosomes, and concomitantly gene expression. Intriguingly, HMGB2 target loci are also usually upregulated once relieved of HMGB2 binding; however, this is the only similarity between the functions of HMGB1 and HMGB2. The loss of HMGB1 does not trigger SICC formation, the same way that the loss of HMGB2 does not trigger p21 activation or SASP induction. Also, HMGB1 and HMGB2 bind non-overlapping genomic loci and demarcate TADs in distinct modes – HMGB2 marks the extremities of TADs that shift one boundary upon senescence entry, while HMGB1 is mostly found at invariable TAD/loop domains boundaries enriched for SASP-related genes. Critically, knockdown of *HMGB1* does not reduce *HMGB2* levels in our cells. Conversely, knocking down *HMGB2* does not affect *HMGB1* levels (Zirkel *et al*, 2018), meaning that the pathways these two factors control do not overlap, but are rather deployed in parallel. However, the nuclear loss of either HMGB coincides with changes in lamins, cohesin, and CTCF levels, which will also contribute to the effects observed in senescence.

Finally, *HMGB1*-knockdown in primary lung fibroblasts leads to gene expression changes that are partially inverted upon senescence entry of the same cells (e.g., the negative regulation of RNAPII transcription is suppressed in the knockdown, but not in senescence; *MYC* activation is upregulated in the knockdown, but suppressed upon senescence entry). This may be interpreted as a coordinated counter-regulation of HMGB1-driven effects on the path to senescence, and can be explained by the fact that the nuclear presence of HMGB1 is linked to favorable proautophagic effects that enhance cell survival and limit programmed cell death (Tang *et al*, 2010). This might also be a simple way to explain the strong overexpression of *HMGB1* in various cancer types (Tang *et al*, 2010; Li *et al*, 2014). This overexpression, although highly deleterious for normal cells, seems to favor increased cell proliferation (Kang *et al*, 2013; Li *et al*, 2014). Thus, the nuclear abundance of HMGB1 (and likely also of HMGB2) can be seen as a marker for proliferative capacity: senescent cells have essentially no nuclear HMGBs, while continuously dividing cancer cells display levels even higher than those seen in normal tissue. In a next step, deciphering the functional implications behind this “readout” may potentially help us understand how a subset of cells escape senescence to acquire a malignant identity.

Materials and methods

Primary cell culture and senescence markers

HUVECs from single, apparently healthy, donors (passage 2-3; Lonza) were continuously passaged to replicative exhaustion in complete Endopan-2 supplemented with 2% FBS under 5% CO₂. Cells were constantly seeded at ~10,000 cells/cm², except for late passages when they were seeded at ~20,000 cells/cm². Single IMR90 isolates (I90-10 and -79, passage 5; Coriell Biorepository) were continuously passaged to replicative exhaustion in MEM (M4655, Sigma-Aldrich) supplemented with non-essential amino acids and 10% FBS under 5% CO₂. Senescence-associated β -galactosidase assay (Cell Signaling) was performed according to the manufacturer's instructions to evaluate the fraction of positively-stained cells at different passages. Cell proliferation was monitored by MTT assays at different passages. In brief, ~5,000 cells are seeded in 96-well format plates in quadruplicates. On the next day, the medium is replaced with 100 μ l fresh medium plus 10 μ l of a 12 mM MTT stock solution (Invitrogen), and cells are incubated at 37°C for 4 h. Subsequently, all but 25 μ l of the medium is removed from the wells, and formazan dissolved in 50 μ l DMSO, mixed thoroughly and incubated at 37°C for 10 min. Samples are then mixed again and absorbance read at 530 nm. Measurements are taken at 24, 48 and 72 h post-seeding, background subtracted, and normalized to the 24 h time point. Finally, nascent DNA synthesis was monitored by EdU incorporation and subsequent labelling with Alexa 488 fluors (Click-iT EdU Imaging Kit; Invitrogen). In brief, cells were incubated in 10 mM EdU for 7 h, fixed using 3.7% PFA/PBS for 15 min at room temperature, permeabilized, and labelled as per manufacturer's instructions, before imaging on a widefield Leica microscope.

Immunofluorescence and image analysis

Proliferating and senescent cells were grown on coverslips from the stage indicated and were fixed in 4% PFA/PBS for 15 min at room temperature. After washing once in PBS, cells were permeabilized in 0.5% Triton-X/PBS for 5 min at room temperature. Blocking with 1% BSA/PBS for 1h was followed by incubation with the following primary antibodies for 1-2 h at the indicated dilution: mouse monoclonal anti-HMGB1 (1:1000; Abcam ab190377-1F3); rabbit polyclonal anti-HMGB2 (1:1000; Abcam ab67282); mouse monoclonal anti-HMGB1/2 (1:1000; Sigma-Aldrich 12248-3D2); rabbit polyclonal anti-CTCF (1:500; Active motif 61311); rabbit polyclonal anti-H3K27me3 (1:1000; Diagenode C15410069); mouse monoclonal anti-p21 (1:500; Abcam ab184640-GT1032); rabbit polyclonal anti-lamin B1 (1:2000; Abcam ab16048); mouse monoclonal anti- β -tubulin (1:1000; Sigma-Aldrich T0198-D66). Following immunodetection, cells were

washed twice with PBS for 5 min before incubating with secondary antibodies for 1 h at room temperature. Nuclei were stained with DAPI (Sigma-Aldrich) for 5 min, washed, and coverslips mounted onto slides in Prolong Gold Antifade (Invitrogen). Note that for gSTED microscopy only, the 2C Pack STED 775 secondary antibodies (1:2000; Abberior 2-0032-052-6) were used. For image acquisition, a widefield Leica DMI 6000B with an HCX PL APO 63x/1.40 (Oil) objective was used; confocal and super-resolution images were acquired on a Leica TCS SP8 gSTED microscope with a 100x/1.40 (Oil) STED Orange objective. For immunofluorescence image analysis, the NuclearParticleDetector2D of the MiToBo plugin (ver. 1.4.3; available at http://mitobo.informatik.uni-halle.de/index.php/Main_Page) was used. Measurements of nuclear immunofluorescence signal were automatically generated using a mask drawn on DAPI staining to define nuclear bounds. Background subtractions were then implemented to precisely determine the mean intensity per area of each immunodetected protein. Deconvolution of super-resolution images was performed using the default settings of the Huygens software (Scientific Volume Imaging).

Whole-cell protein extraction, western blotting, and mass spectrometry

For assessing protein abundance at different passages, $\sim 4 \times 10^6$ cells were gently scraped off 15-cm dishes, and pelleted for 5 min at 600 x *g*. The supernatant was discarded, and the pellet resuspended in 100 mL of ice-cold RIPA lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM EDTA pH 8.0, 1 mM EGTA pH 8.0, 1% NP-40, 1% sodium deoxycholate) containing 1x protease inhibitor cocktail (Roche), incubated for 20 min on ice, and centrifuged for 15 min at $>20,000 \times g$ to pellet cell debris and collect the supernatant. The concentration of nuclear extracts was determined using the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific), before extracts were aliquoted and stored at -70°C to be used for western blotting. For fractionations, the protocol previously described was used (Watrin *et al*, 2006) with addition of 1.5 μM EGS (10 min, room temperature) to stabilize HMGB1 on chromatin. Resolved proteins were detected using the antisera mentioned above, plus a mouse monoclonal anti-H3K9me3 (1:200; Active motif 39286). For whole-cell proteomics, extracts in RIPA buffer were analyzed by the CECAD proteomics core facility in biological triplicates on a Q-Exactive Plus Orbitrap platform (Thermo Scientific) coupled to an EASY nLc 1000 UPLC system with column lengths of up to 50 cm.

HMGB1 sCLIP and analysis

sCLIP was performed on ~ 25 million UV-crosslinked nuclei from proliferating IMR90 as previously described (Kargapolova *et al*, 2017) using the same the monoclonal HMGB1 antiserum (DSHB; PCR-P-HMGB1-4F10) as for CHIP. Following sequencing of strand-specific libraries on a HiSeq4000 platform (Illumina), raw reads

were mapped to the human reference genome (hg19). Consistent peaks were identified by overlapping intervals of peaks with a P -value <0.05 from 2 biological replicates to obtain 1773 peaks. This peak annotation was used to count reads uniquely aligned to each peak region using HTSeq, HMGB1-bound transcript coordinates were retrieved via Ensembl (GRCh37) and annotated using HOMER (<http://homer.ucsd.edu>), and Gene Ontology analysis was performed using Metascape (www.metascape.org). Finally, the final merged peak list was used for *de novo* motif analysis using ssHMM (Heller *et al*, 2017) and significantly enriched motifs were compared to existing RBP motifs via Tomtom (<http://meme-suite.org/tools/tomtom>).

Chromatin immunoprecipitation (ChIP) sequencing and analysis

For each batch of ChIP experiments ~25 million proliferating cells, cultured to $> 80\%$ confluence in 15-cm dishes, were crosslinked in 1.5 mM EGS/PBS (ethylene-glycol-bis-succinimidyl-succinate; Thermo) for 20 min at room temperature, followed by fixation for 40 min at 4°C in 1%PFA. From this point onward, cells were processed via the ChIP-IT High Sensitivity kit (Active motif) as per manufacturer's instructions. In brief, chromatin was sheared to 200-500 bp fragments on a Bioruptor Plus (Diagenode; 2x 9 cycles of 30 sec *on* and 30 sec *off* at the highest power setting), and immunoprecipitation was carried out by adding 4 μg of a monoclonal HMGB1 antiserum (Developmental Studies Hybridoma Bank; PCRP-HMGB1-4F10) to ~30 μg of chromatin and rotating overnight at 4°C in the presence of protease inhibitors. Following addition of protein A/G agarose beads and washing, DNA was purified using the ChIP DNA Clean & Concentrator kit (Zymo Research) and used in next-generation sequencing on a HiSeq4000 platform (Illumina) to obtain at least 25 million reads were obtained of both sample and its respective "input". Raw reads (typically 100 bp-long) were processed with Encode ChIP-Seq pipeline (v1.5.1; ENCODE Project Consortium, 2012). Reads were mapped to hg19 human reference genome with BWA (Li, 2013); subsequent processing and filtering steps were performed with default pipeline settings by Picard (<https://broadinstitute.github.io/picard/>), BEDTools (Quinlan & Hall, 2010), Phantompeaktools and SPP peak callers. Aligned reads were used to produce IP/Input signal coverage tracks with MACS2 (Zhang *et al*, 2008). IDR thresholded HMGB1 ChIP-seq peaks (conservative peak calling) per each cell type were annotated using Chipseeker (Yu *et al*, 2015); Signal coverage files representing log₂-fold change of IP/Input signal were used in Deeptools (v3.5.0, Ramírez *et al*, 2016) for plotting signal coverage over particular genomic positions for different conditions/cell types. Finally, transcription factor recognition motif enrichments within DHS footprints under HMGB1 ChIP-seq peaks were calculated using the Regulatory Genomics Toolbox (Gusmao *et al*, 2014). Note that all other ChIP-seq datasets used here come from

previous work (Zirkel *et al*, 2018), with the exception of publicly available CTCF ChIP-Seq datasets for IMR90 (SRR639078) and HUVEC (ENCSR000ALA), RAD21 for IMR90 (ENCSR000EFJ), H3K4me3 for IMR90 (ENCSR431UUY) and HUVEC (ENCF203KHF), and H3K27ac for IMR90 (ENCSR002YRE) and HUVEC (ENCF038HNR). IMR90 CTCF ChIP-seq was processed with Encode ChIP-Seq pipeline to match the rest of used datasets.

Total RNA isolation, sequencing, and analysis

Control and *HMGB1*-knockdown were harvested in Trizol LS (Life Technologies) and total RNA was isolated and DNase I-treated using the DirectZol RNA miniprep kit (Zymo Research). Following selection on poly(dT) beads, barcoded cDNA libraries were generated using the TruSeq RNA library kit (Illumina) and were paired-end sequenced to at least 50 million read pairs on a HiSeq4000 platform (Illumina). Raw reads were mapped to the human reference genome (hg19) using default settings of the STAR aligner (Dobin *et al*, 2013), followed by quantification of unique counts using *featureCounts* (Liao *et al*, 2014). Counts were further normalized via the RUVs function of RUVseq (Risso *et al*, 2014) to estimate factors of unwanted variation using those genes in the replicates for which the covariates of interest remain constant and correct for unwanted variation, before differential gene expression was estimated using DESeq2 (Love *et al*, 2014). Genes with an FDR <0.01 and an absolute (\log_2) fold-change of >0.6 were deemed as differentially-expressed. For splicing analysis, a reference index on the basis of hg19 annotation was first constructed, combined with all splice sites contained in the mapped RNA-seq reads. Raw reads were then aligned using Whippet (Sterne-Weiler *et al*, 2018) to the constructed index in order to quantify and annotate alternative splicing events. Subsequent plots were plotted using BoxPlotR (<http://shiny.chemgrid.org/boxplotr/>), and GO term enrichment bar plots using Metascape (<http://metascape.org/gp/index.html>); (Zhou *et al*, 2019b).

siRNA-mediated *HMGB1* knockdown and overexpression

HUVECs were seeded at ~20,000 cells/cm² the day before transfection. Self-delivering Accell-siRNA pools (Dharmacon) targeting *HMGB1*, plus a non-targeting control (NTC; fluorescently-tagged to allow transfection efficiency to be monitored), were added to the cells at a final concentration of 1 mM. Knockdown efficiency was assessed 72 h after transfection using RT-qPCR and immunofluorescence. For IMR90 cells, transfections using two different siRNAs and RNAiMAX (Invitrogen) were carried out as previously described (Zirkel *et al*, 2018). For *HMGB1*-GFP overexpression using the PiggyBac system details were exactly as described previously for *HMGB2* (Zirkel *et al*, 2018) with the difference that cells only tolerated a <10-h induction of overexpression via doxycycline. In brief, the *HMGB1* open reading frame was

subcloned into the doxycycline-inducible KA0717 expression vector to generate an HMGB1-GFP fusion. The construct was co-transfected into IMR90 together with trans-activator and transposase-encoding vectors (KA0637 and SBI Biosciences #PB200PA-1, respectively) at a DNA mass ratio of 10:1:3 using Fugene HD (Promega) as per manufacturer's instructions. Stable, transgene-positive, proliferating IMR90 were selected using 250 mg/ml G418 (Sigma Aldrich) and driven into senescence before being either harvested in Trizol for downstream RT-qPCR analysis or crosslinked with 4% paraformaldehyde for 5 min at room temperature for immunofluorescence analyses.

RNA immunoprecipitation (RIP) experiments

Approximately 5×10^7 proliferating (p. 16; for HMGB1 and ILF3 IP) and senescent IMR90 (p. 35; for HMGB1 IP) were scraped on ice in PBS (Sigma), pelleted by centrifugation, and lysed in ice-cold Polysome lysis buffer (100 mM KCl, 5 mM MgCl₂, 10 mM HEPES-NaOH pH 7.0, 1 mM DTT, 200 U/ml RNaseIn, 1x PIC, 0.5% NP-40) for 30 min. To ensure complete cell lysis, lysates were homogenized via 10 strokes with a Dounce homogenizer, passed through 27½-gauge needle 4 times and sonicated for 2 x 6 cycles (30 sec on/30 sec off, low input) on a Bioruptor sonicator (Diagenode). Lysates were next treated with 400 U/ml DNase I (Worthington, LS006343) for 30 min and centrifuged at maximum speed for 10 min to collect supernatants. 5% of the cell lysate was saved as input and the rest was subjected to overnight immunoprecipitation at 4°C in the presence of 10 µg HMGB1 (DSHB, 4F10; 68 µg/µl), ILF3 (Abclonal Biotechnology, A2496) or IgG antibody (Milipore, 12-371B; 1 µg/µl). Next day, 30 µl of protein-G beads (Dynabeads, Invitrogen) per IP were pre-washed 3 times with 1 ml of NT2 buffer (1x buffer contains 250 mM Tris-HCl pH 7.4, 750 mM NaCl, 5 mM MgCl₂, 0.25% NP-40) and incubated for 1 h at 4°C with 25 µl/IP of Mouse bridging antibody (Active Motif; 1 µg/ul) with end-to-end rotation. After 3 washes with NT2 buffer, the beads were added to the lysates and incubated at 4°C for 2 h under end-to-end rotation. After incubation, beads were washed 6 times with 1 ml of NT2 buffer for 3 min under end-to-end rotation between washes. Finally, samples were resuspended in TRIzol (Invitrogen) and RNA was purified using the Direct-zol RNA Miniprep kit (Zymo). Reverse transcription was carried out using the SuperScript™ II Reverse Transcriptase according to manufacturer's instructions (Invitrogen) and qPCR reactions using the qPCRBIO SyGreen Mix Separate-ROX (NIPPON).

Co-immunoprecipitation experiments coupled to mass spectrometry

Approx. 6×10^6 proliferating IMR90s were gently scraped and pelleted for 5 min at 600 x *g*, supernatant discarded, and the pellet resuspended in 500 µl of ice-cold lysis buffer (20 mM Tris-HCl pH 8.0, 1% NP-40,

150 mM NaCl, 2 mM EDTA pH 8.0) supplemented with 1x protease inhibitor cocktail (Roche). This mixture was then incubated for 20 min on ice, followed by 3 cycles of sonication (30 sec on, 30 sec off, low input) and RNase A treatment, before centrifugation for 15 min at $>20,000 \times g$ to pellet cell debris and collect the supernatant. While lysates were precleared, 30 μ l protein-G magnetic beads (Active motif) and 10 μ g of HMGB1 antiserum (PCRP-HMGB14F10s; DSHB) were incubated for 3 h at 4°C under rotation. Subsequently, the beads were captured on a magnetic rack (Active motif) and added to the lysates for incubation at 4°C overnight under rotation. Next day, the beads were captured, washed four times with 800 μ l ice-cold wash buffer I (50 mM Tris, 0,05% NP-40 and 50 mM NaCl), two times with 500 μ l of wash buffer II (150 mM NaCl, 50 Mm Tris), recaptured, supernatant discarded and purified proteins were predigested in 50 μ l elution buffer (2 M urea, 50 mM Tris pH 7.5, 1 mM DTT, 50 ng trypsin) for 30 min at room temperature with gentle agitation. Following addition of 50 μ l digestion buffer (2 M Urea dissolved in 50 mM Tris pH 7.5 and 5 mM chloroacetamide) and incubation for 30 min, another 50 μ l of elution buffer supplemented with 50 ng of LysC and 100 ng of trypsin were added to each tube. Proteins were digested overnight at room temperature, the digestion was stopped by adding 1 μ l trifluoroacetic acid, and peptides of each experiment were split in half, purified on two C18 stage tips, and all three replicates were analyzed by the CECAD proteomics core facility as above.

Ribo-seq and analysis

High throughput ribosome profiling (Ribo-seq) on proliferating and senescent IMR90 was performed in collaboration with Ribomaps Ltd (<https://ribomaps.com>) according to an established protocol (Ivanov *et al*, 2018). Three independent replicas of proliferating or senescent IMR90 were grown, harvested in ice-cold polysome isolation buffer supplemented with cyclohexamide, and shipped to Ribomaps for further processing and library preparation. Approx. 15% of each lysate was kept for isolation of RNA and used for RNA-seq of poly(A)-enriched fractions on a HiSeq2500 platform (Illumina). Following sequencing of both Ribo- and mRNA-seq libraries, the per base sequencing quality of each replicate passed the quality threshold, raw read counts were assigned to each protein-coding open reading frame (CDS) for Ribo-seq and to each transcript for mRNA-seq, and replicate correlations were tested. Read length distribution for Ribo-seq datasets fell within the expected range of 25-35 nt, with a peak between 28-32 nt showing strong periodic signals and an enrichment in annotated CDSs. For mRNA-seq, read lengths ranged between 47-51 nt and distributed uniformly across transcripts. For differential gene expression analysis, Anot2seq (Oertlin *et al*, 2019) was used. Changes in Ribo-seq data represent changes in the ribosome occupancy of the annotated protein coding open reading frame (CDS) and, thus, only ribosome-protected fragments that

map to the CDS were used in the analysis. VST normalized counts outputted using DESeq2 (Love *et al*, 2014) and inputted into anota2seq were used for all subsequent downstream analysis. Differences on genes that pass a default false discovery rate (FDR) threshold of 15% were considered regulated. Such significant differences are then categorized into one of the following three modes: (i) translational: changes in Ribo-seq that are not explained by changes in RNA-seq and imply changes at the protein level are due to changes at the translational level; (ii) mRNA abundance: matching changes in RNA-Seq and Ribo-Seq that infer changes at the protein level are predominantly induced by changes at the transcriptional level; (iii) buffering: changes in RNA-seq that are not explained by changes in Ribo-seq and suggest maintenance of constant protein levels induced by changes at the transcriptional level or *vice versa*.

Whole-genome chromosome conformation capture (Hi-C) and TiLO analysis

Hi-C data from proliferating and senescent HUVEC and IMR90 were generated previously (Zirkel *et al*, 2018), and the HiTC Bioconductor package was used to annotate, correct data for biases in genomic features (Servant *et al*, 2012), and visualize 2D heatmaps with a maximum resolution of 25-kbp at which TADs were also called via TADtool (Kruse *et al*, 2016). For plotting insulation and “loop-o-gram” heatmaps, normalized interactions values in the twenty 25-kbp bins around each HMGB1 peak were added up, normalized to the median value in each matrix and plotted provided the local maxima are higher than the third quantile of Hi-C data in the matrix. All R scripts were described previously (Zirkel *et al*, 2018). Loop calling was performed with FAN-C (Kruse *et al*, 2020) set of loops calling and filtering commands at resolution of 20-kbp.

For Topologically-intrinsic Lexicographic Ordering (TiLO), we directly applied an algorithm from mathematical knot theory that makes zero assumptions about the structure, shape or number of clusters in the data (Johnson, 2012). In brief, topologically-intrinsic ordering was used to permute the linear order of TADs (as the starting organization level in the Hi-C matrices) until a certain “robustly irreducible” topological condition is satisfied. Then, the “pinch ratio” algorithm (Heisterkamp & Johnson, 2013) was applied to heuristically slice the network at connections between TADs exhibiting local interaction minima, while also considering noise in the matrices. Finally, this analysis returns a list of TADs grouped into multiple clusters *in cis*, also via its built-in measure for network robustness defining the end-point. For rendering 3D chromosome model of IMR90 and HUVEC TiLO data the Chrom3D interface was used (Paulsen *et al*, 2018). For all chromosomes per cell type, intra-TAD interactions were specified according to TiLO output. Association with LADs were added as described in the Chrom3D manual for each chromosome (<https://github.com/Chrom3D>). LADs for proliferating and senescent IMR90 cells were inferred from

LMNB1 CHIP-seq data (GSM1197635). Reads were aligned to the hg19 reference genome using Bowtie2 v2.3 with default parameters and merged using SAMtools v1.9. The outputs were applied to EDD v1.0 for LAD calling also using default parameters (Lund *et al*, 2014). In the absence of similar data for HUVECs, constitutive LAD positions (cLADs and ciLADs) were downloaded from the LAD atlas (Meuleman *et al*, 2013) and used the same way. In the end, a *.gtrack* file (Chrom3D input) for chromosome visualization was produced using Chrom3D scripts (https://github.com/Chrom3D/preprocess_scripts). Next, a *.BED* file specifying the genomic positions of the TADs (1 TAD = 1 bead) was created, and any gaps between them were filled as described in the Chrom3D manual. TADs belonging to the same contiguous TiLO cluster were grouped and a separate *.gtrack* file was created for each TiLO cluster containing only intrachromosomal interactions. All single-base beads corresponding to gaps between TADs were removed from the final file. Finally, *.gtrack* files corresponding to each cluster were merged and inputted in Chrom3D, using 200000 iterations (-n), a nuclear radius of 5 (-r) and a scale total volume of the beads relative to the volume of the nucleus set to 0.15 (-y). For whole genome visualizations that take into account interchromosomal interactions, Hi-C data were analysed via HiCPro v2.11.4 at 40-kbp and 1-Mbp resolution, before LADs, TADs and Hi-C matrices were used for the production of a diploid *.gtrack* file using default parameters; chromosome Y and M were removed. IDs of beads containing HMGB1 peaks were identified and colored using the script *processing_scripts/color_beads.py* and the *blend* keyword to maintain coloring.

Statistical tests

P-values associated with Student's t-tests and Fischer's exact tests were calculated using GraphPad (<https://graphpad.com/>), those associated with the Wilcoxon-Mann-Whitney test using R, and those with the hypergeometric test using an online tool (http://nemates.org/MA/progs/overlap_stats.html). Unless otherwise stated, *P*-values <0.01 were deemed as statistically significant.

Data availability

All NGS data generated in this study have been deposited to the NCBI Gene Expression Omnibus (GEO) repository as part of the GSE171782 SuperSeries as follows: Hi-C and RNA-seq from proliferating/senescent HUVEC and IMR90 (GSE98448), sCLIP from proliferating IMR90 (GSE146047), Ribo-seq data from proliferating/senescent IMR90 (GSE171780), HMGB1 CHIP-seq data from proliferating/senescent HUVEC and IMR90 (GSE171781) and HMGB1-knockdown RNA-seq data from IMR90 (GSE171779).

Author contributions

SK, AZ, NS, NJ, YK, TG, IP, NÜ, and AM performed experiments; GL, JK, and AM generated and analyzed Ribo-seq data; MN, NJ, YK, IV, and EGG performed computational analyses; CB and JA performed all NGS; AP conceived the study and wrote the manuscript with input from all coauthors.

Appendix – Supplementary figures 1-6

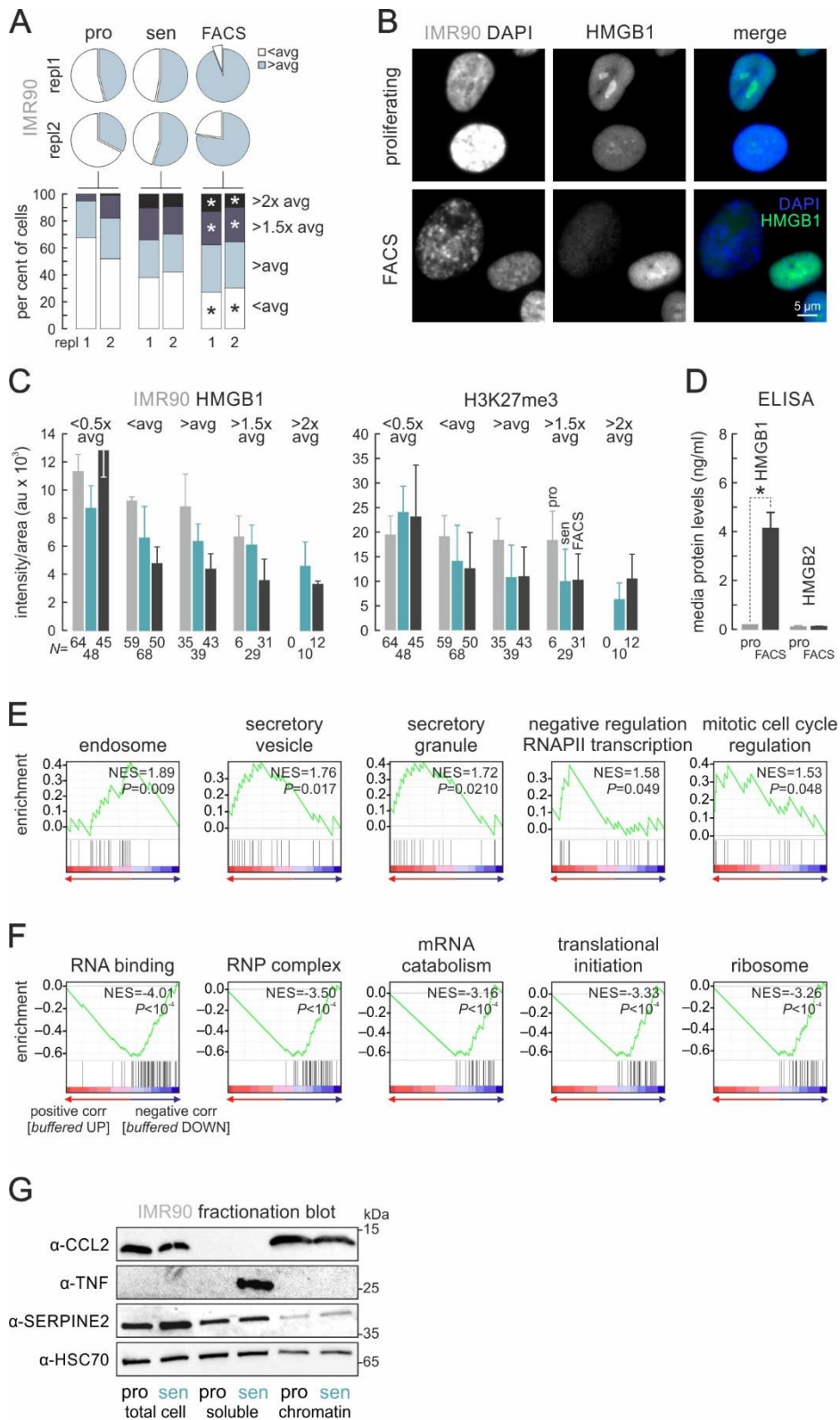


Figure 2.S1. HMGB1 is depleted from enlarged nuclei and accompanied by gene expression changes.

(A) Pie charts (top) of nuclear size in IMR90 populations. Bar graphs (bottom) stratify cells according to increasing nuclear size. *: significantly different to proliferating cells, $P < 0.05$; Fisher's exact test. (B) Representative images of proliferating (top row) and FACS-sorted senescent IMR90 (bottom row) immunostained for HMGB1 and counterstained with DAPI. Bar: 5 μm . (C) Bar graphs showing declining HMGB1 (left) and H3K27me3 levels (mean \pm S.D.; right) in proliferating (grey), senescent (green) and FACS-sorted IMR90 (black) stratified according to nuclear size from images like those in panel B. The number of cells analyzed (N) is indicated. (D) Bar graphs showing HMGB1/B2 levels (mean \pm S.D.) detected in the growth media of proliferating (grey) or FACS-sorted IMR90 (black). *: $P < 0.01$; unpaired two-tailed Student's t-test (N=2). (E) Gene set enrichment analysis of genes translationally upregulated but transcriptionally suppressed in Ribo-seq data. Normalized enrichment scores (NES) and associated P-values are shown. (F) As in panel E, but for genes translationally downregulated but transcriptionally upregulated. (G) Chromatin fractionation western blots of exemplary SASP-relevant factors in proliferating and senescent IMR90; CCL2 provides a negative and HSC70 a loading control.

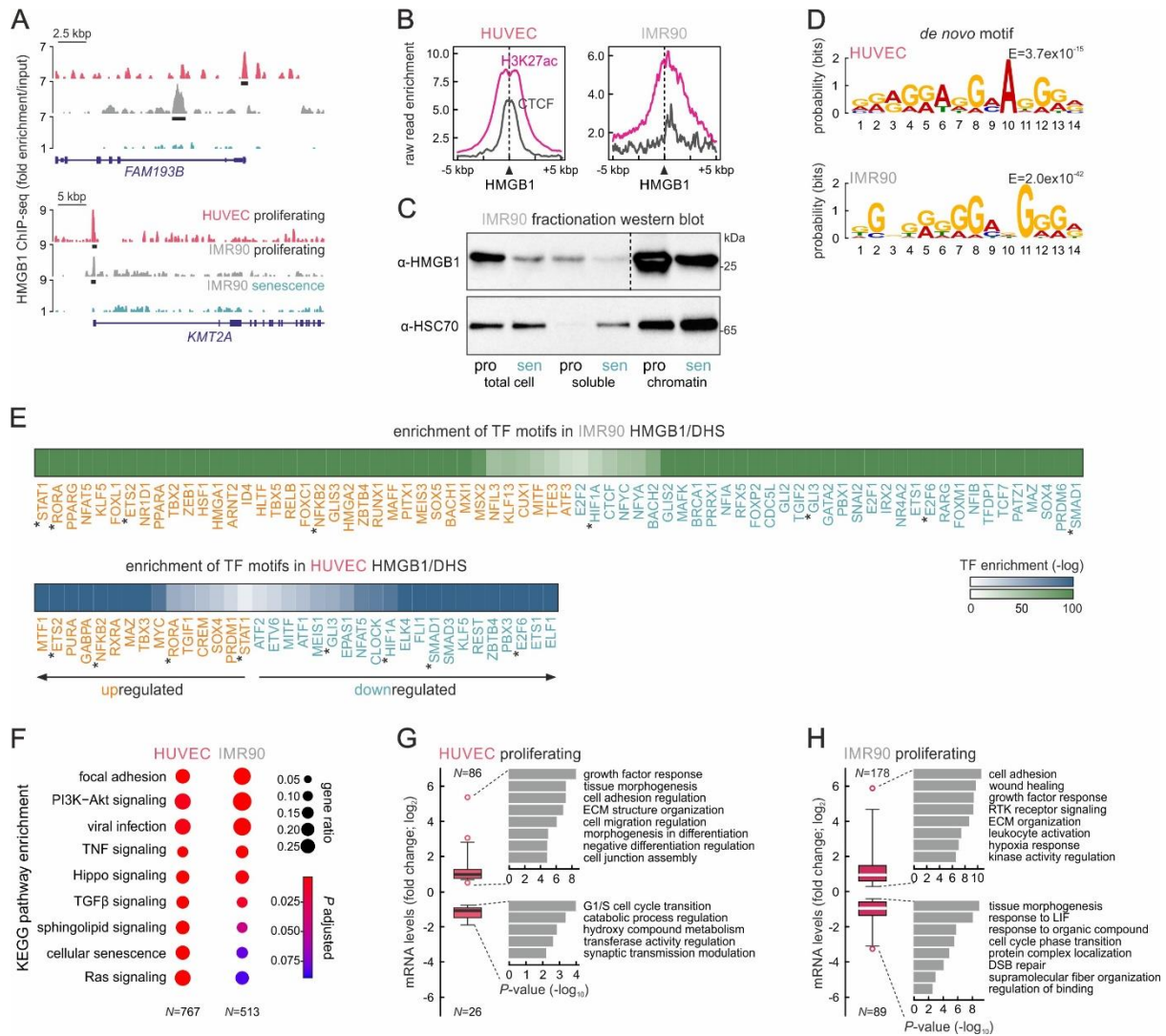


Figure 2.S2. HMGB1 chromatin-binding features in proliferating human cells.

(A) Genome browser views showing normalised HMGB1 ChIP-seq signal at the FAM193B and KMT2A locus from both HUVEC (pink) and proliferating/senescent IMR90 (grey/green). (B) Line plots showing distribution of H3K27ac (magenta) and CTCF ChIP-seq signal (grey) in the 10 kbp around HMGB1 peaks from proliferating HUVECs and IMR90. (C) Fractionation western blot in proliferating or senescent IMR90; HSC70 provides a loading control. The chromatin fraction part of the HMGB1 blot (dashed line) is shown at a shorter exposure (1.0 sec) than the rest of the blot (5.5 sec; see also Source Data). (D) Logos and associated E-values for the most enriched de novo motifs under HMGB1 peaks. (E) Heatmap showing most enriched TF binding motifs in the footprints from panel C; up- (>0.6 log₂-fold change; orange) or downregulated TF genes (<0.6 log₂-fold change; green) are indicated. (F) Plots KEGG pathways associated with HUVEC and IMR90 HMGB1-bound genes. The number of genes in each cell type (N) is indicated. (G) Box plots (left) showing significantly up/downregulated genes from proliferating HUVECs with HMGB1 peaks within 20 kbp of their TSS. GO terms associated with these differentially-regulated genes are plotted (right). The number of genes in each group (N) is indicated. (H) As in panel G, but for HMGB1-bound genes in IMR90.

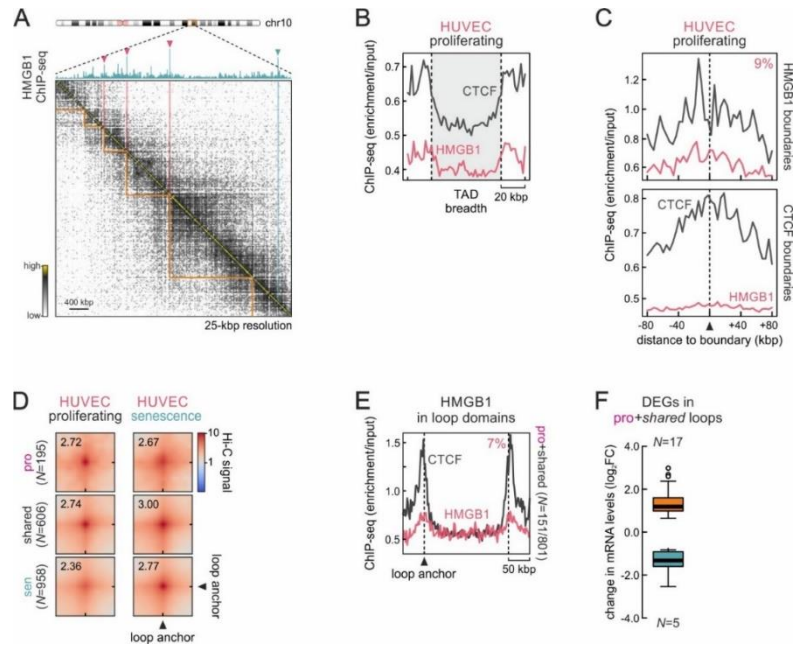


Figure 2.S3. HMGB1 demarcates TAD and loop domains boundaries in proliferating HUVECs.

(A) Exemplary Hi-C heatmap for a subregion in HUVEC chr10 aligned to HMGB1 ChIP-seq; peaks at TAD boundaries are indicated (pink arrowheads). (B) Line plot showing normalised HMGB1 (green) and CTCF ChIP-seq signal (grey) along all 2888 TADs ± 20 kbp from proliferating HUVEC. (C) As in panel B, but in the 160 kbp around HMGB1- (top) or CTCF-marked TAD boundaries (bottom). The percentage of HMGB1 peaks at boundaries is indicated (top right). (D) Aggregate plots showing 20 kbp-resolution Hi-C signal for proliferating (top), senescent-specific (bottom) or shared loops (middle). The number of loops in each category (N) is indicated. (E) Line plot showing normalised HMGB1 (red) and CTCF ChIP-seq signal (grey) along all 801 loop domains TADs ± 50 kbp from proliferating HUVEC. (F) Box plot showing significantly up/downregulated HUVEC genes harboured inside the loops from panel E. The number of genes in each group (N) is indicated.

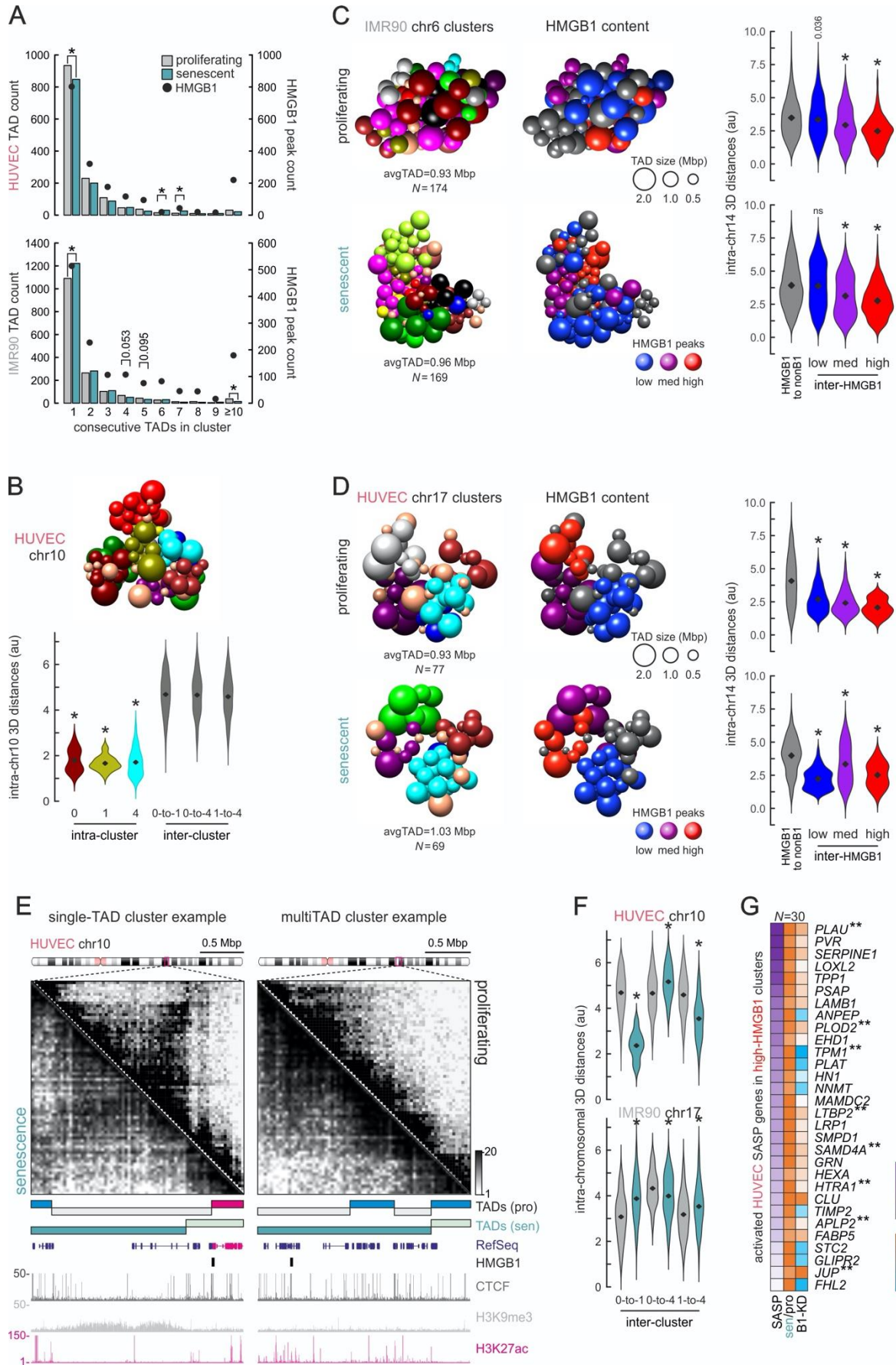


Figure 2.S4. HMGB1 demarcates specific TAD subsets in primary human cells.

(A) Bar plots showing the number of TADs contained in progressively larger clusters derived using TiLO on proliferating (grey) and senescent Hi-C data (green) from HUVEC (top) and IMR90 (bottom). Dots show the absolute number of HMGB1 ChIP-seq peaks in each TiLO cluster category. *: $P < 0.01$; Fisher's exact test. **(B)** Chrom3D visualisation of TAD clustering in chr10 from proliferating HUVEC (top); each sphere represents one TAD. Violin plots (below) show 3D distances amongst TADs in three randomly-selected clusters (0, 1, and 4) or between TADs from different clusters. *: significantly different to inter-cluster distances, Wilcoxon-Mann-Whitney test. **(C)** Chrom3D visualisation of TAD clustering in chr17 from proliferating (top) and senescent HUVEC (bottom). TADs (spheres) are coloured by the cluster they belong to (left) or according to their HMGB1 ChIP-seq content (middle; grey – zero peaks, blue – 1 or 2 peaks, purple – 3 or 4 peaks, red – 5 or more peaks). Violin plots (right) show 3D distances amongst TADs in each subgroup or between HMGB1-containing and non-containing TADs. *: significantly different to inter-cluster distances, Wilcoxon-Mann-Whitney test. **(D)** As in panel C, but for chr6 from proliferating and senescent IMR90. **(E)** Left: Exemplary Hi-C data (40-kbp resolution) around a single-TAD cluster (magenta) in HUVEC chr10 aligned to TAD positions, HMGB1 peaks, and ENCODE ChIP-seq. Right: Exemplary Hi-C data for TADs forming a multi-TAD cluster in senescent HUVEC according to TiLO. **(F)** Violin plots showing interchromosomal distances between TADs in different clusters of HUVEC chr10 (top) or IMR90 chr17 (bottom) based on TiLO/Chrom3D visualisations. *: significantly different to inter-cluster distances, Wilcoxon-Mann-Whitney test. **(G)** Heatmap showing protein (SASP, from <http://www.saspatlas.com/>; log) and gene expression levels (log2FC) of HUVEC SASP-related genes embedded in high-HMGB1 (red) TADs like those in panel E. **: genes bound by HMGB1 in HUVEC ChIP-seq data are more than expected by chance; $P > 0.001$, hypergeometric test.

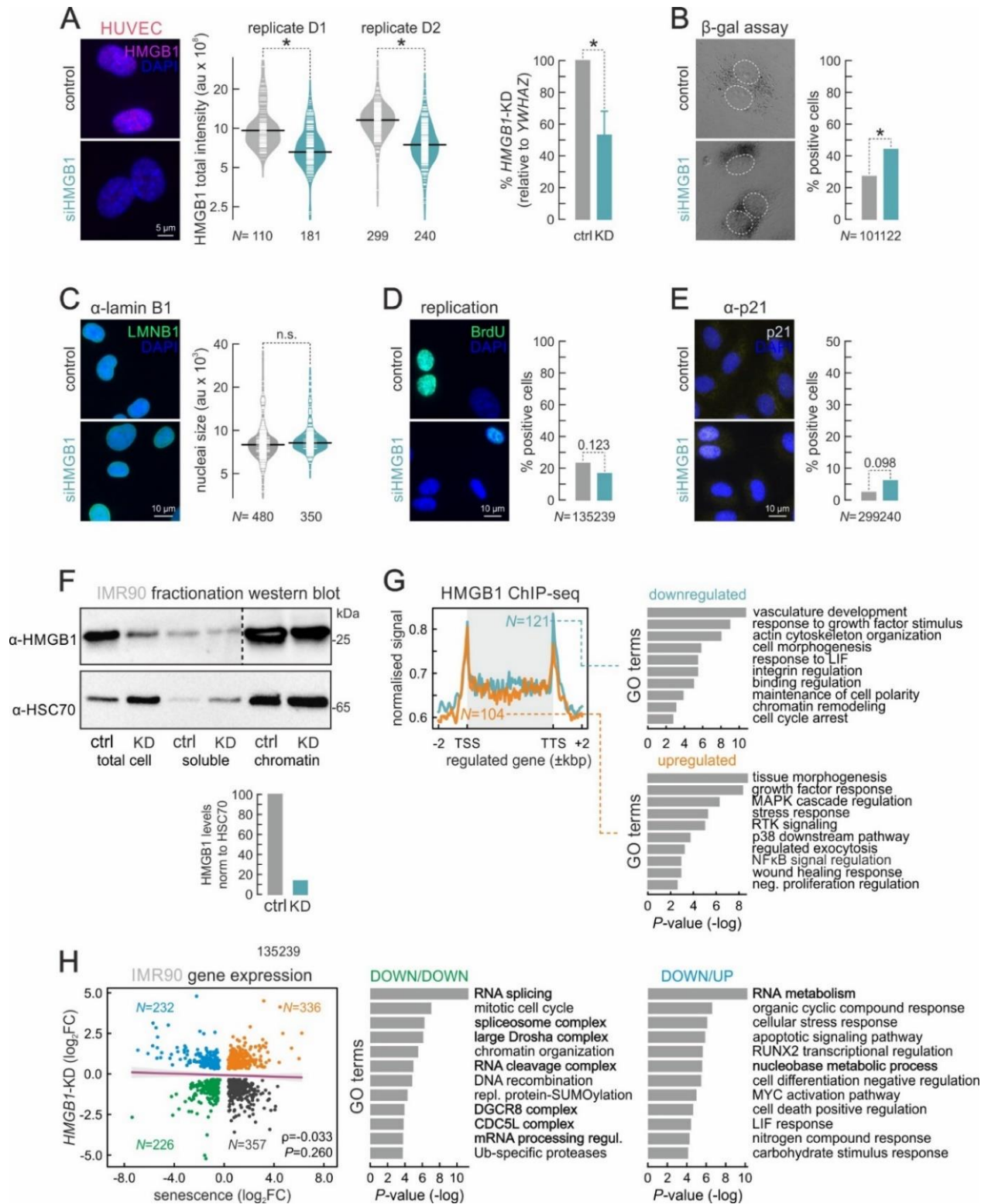


Figure 2.S5. HMGB1-knockdown effects in proliferating HUVECs and IMR90.

(A) Representative immunofluorescence images of HMGB1 (*left*) in siHMGB1-treated and control cells counterstained with DAPI. Bean plots quantify knockdown efficiency (*middle*; N is the number of cells analyzed). Bar graphs (*right*) show reduction in HMGB1 mRNA levels (\pm S.D.; N=2) in knockdown compared to control cells. Bar: 5 μ m. *: P<0.01; Wilcoxon-Mann-Whitney and unpaired two-tailed Student's t-test for bean and bar plots, respectively. (B) Representative brightfield images of siHMGB1-treated HUVECs showing elevated β -galactosidase activity compared to control cells (*left*); bar graphs quantify this increase (*right*; N indicates the number of cells analyzed). *: P<0.01; Fisher's exact test. (C) As in panel A, but for LMNB1 levels. Bar: 10 μ m. No statistically significant difference (n.s.); Wilcoxon-Mann-Whitney test. (D) As in panel A, but for BrdU incorporation. Bar: 10 μ m. P=0.123; Fisher's exact test. (E) As in panel A, but for p21 levels. Bar: 10 μ m. P=0.098; Fisher's exact test. (F) Fractionation western blot (*top*) in control and HMGB1-knockdown IMR90; HSC70 provides a loading control. Bar graphs (*below*) quantify the reduction in chromatin bound HMGB1 after normalization to the loading control. The chromatin fraction of the HMGB1 blot (*dashed line*) is shown at a shorter exposure (1.0 sec) than the rest of the blot (3.3 sec; see also Source Data). (G) Line plots (*left*) showing normalized HMGB1 ChIP-seq signal along genes up- (*orange*) or downregulated upon HMGB1-knockdown in IMR90 (*green*; N indicates the number of HMGB1-bound genes). Bar plots (*right*) showing GO terms associated up/downregulated genes and their enrichment P-values. (H) Scatter plots (*left*) showing correlation between differentially expressed genes (\log_2 FC) upon senescence entry and HMGB1-knockdown in IMR90. Pearson's correlation (ρ) and associated P-value are shown alongside the number of genes (N) in each subgroup. Bar plots (*right*) showing GO terms and enrichment P-values associated with genes commonly downregulated ("DOWN/DOWN") or downregulated in senescence but upregulated upon HMGB1-knockdown ("DOWN/UP"). Terms relevant to RNA processing are highlighted (*bold*).

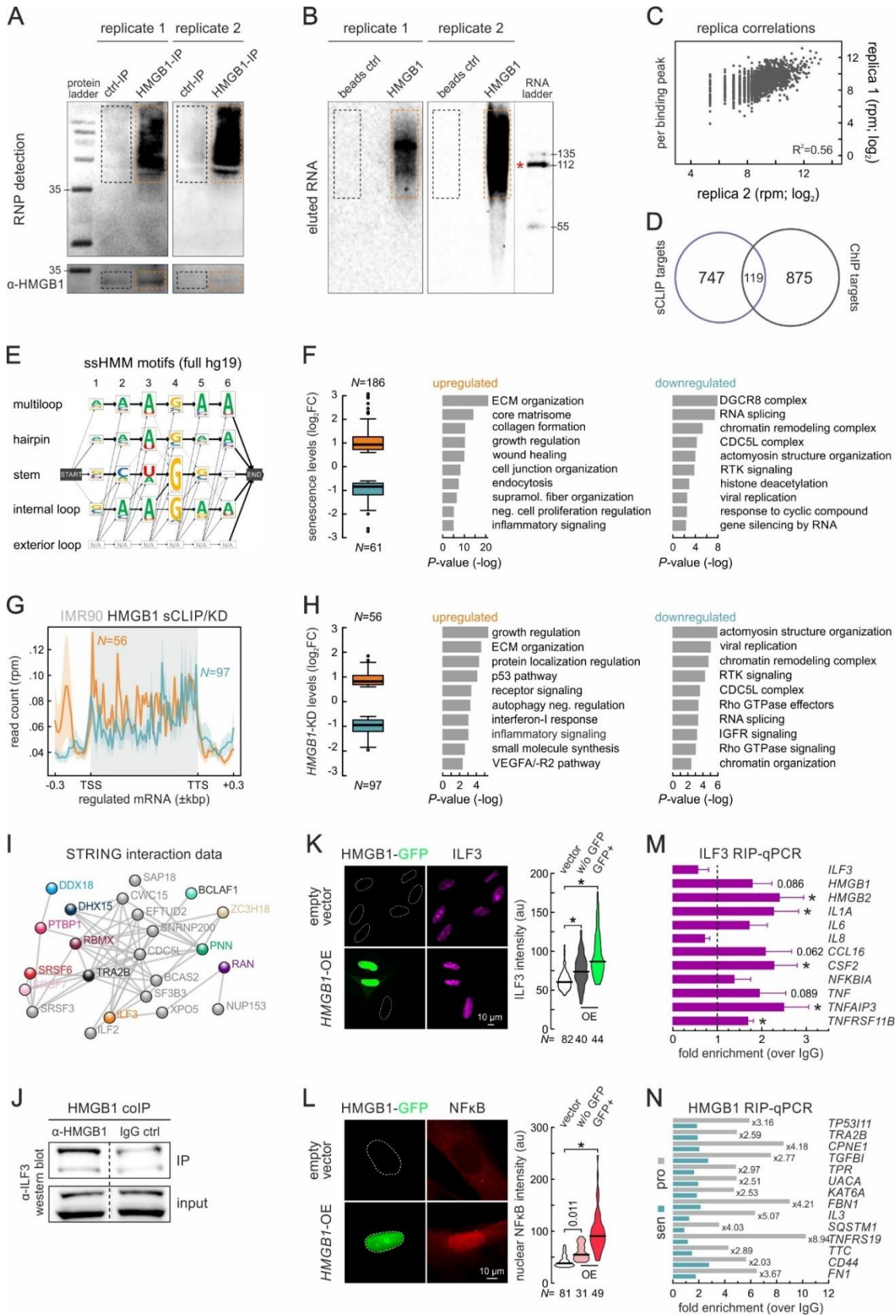


Figure 2.S6. HMGB1 sCLIP controls and analysis of HMGB1-interacting RBPs.

(A) Electrophoretic profiles of control (beads only; *black dotted square*) and HMGB1 IP (*orange dotted rectangle*) probed for RNA (*top*) or HMGB1 (*bottom*) in both sCLIP replicates. The 35-kDa position of the molecular mass ladder is indicated. (B) As in panel A, but for RNA eluted from control (beads only; *black dotted square*) or HMGB1 IP (*orange dotted rectangle*) in both sCLIP replicates. The 112-nt band in the molecular mass ladder (*red star*) corresponds to ~144 pg of RNA. (C) Scatter plot showing correlation of sCLIP data from two independent biological replicates compared per binding peak normalized read count. Spearman correlation values (R2) are shown. (D) Venn diagram showing overlap of HMGB1 sCLIP and CHIP-seq targets in proliferating IMR90. (E) ssHMM motif analysis of sCLIP data showing sequence probabilities in HMGB1-bound motifs predicted to form different structures using the reference human genome (hg19). (F) Box plots (*left*) showing change (log2) in mRNAs differentially regulated upon senescence and bound by HMGB1 in IMR90 sCLIP. The number of mRNAs (N) analyzed is indicated. Bar graphs (*right*) showing GO terms associated with each subgroup and their enrichment P-values. (G) Line plots showing mean HMGB1 sCLIP signal along mRNAs up- (*orange*) or downregulated (*green*) upon HMGB1-knockdown in IMR90. The number of HMGB1-bound mRNAs (N) is indicated. (H) As in panel F, but for the up-/downregulated mRNAs from panel F. (I) Experimentally validated protein-protein interaction network for HMGB1-coimmunoprecipitating RBPs (*colored spheres*) that are also downregulated in senescence. Their secondary interactors (*grey spheres*) from the STRING database are also shown. (J) Western blots showing ILF3 co-immunoprecipitating with HMGB1 (*left*) or IgG (*right*) in IMR90; blots on 10% of input material provide a loading control. (K) Representative images of IMR90 overexpressing HMGB1-GFP and immunostained for ILF3. Outlines of nuclei are indicated (dashed lines based on DAPI counterstaining). IMR90 transfected with an empty vector provides a control. Bar: 10 μ m. Signal intensities are quantified in the violin plots (*right*). *: significantly different to control; P<0.001, Wilcoxon-Mann-Whitney test (L) As in panel K, but immunostained for phospho-NF κ B. Bar: 10 μ m. (M) Bar graphs showing mean fold-enrichments (over IgG controls \pm S.D., N=2) of selected mRNAs from ILF3 RIP experiments in proliferating IMR90. *: significantly different to control; P<0.05, unpaired two-tailed Student's t-test. (N) As in panel M, but showing mean HMGB1 enrichment on selected mRNA targets in proliferating (*grey*) and senescent IMR90 (*green*). Fold differences between the two states are indicated.

Chapter 3 - circRNAs modulate cell proliferation and the immediate-early response by reorganizing gene expression and mRNA processing

My contribution is reflected in the following experimental/bioinformatic analysis as outlined below:

- FISH experiments and knock-down experiment in HUVECs
- RAP-MS together with K.K. Ebbesen
- All the experiments concerning overexpressing cell lines and TNF α treatment
- Computational analysis of all generated RNA-seq

The following figure panels were prepared by myself, with the data input from experiments prepared and/or analyzed by me or indicated contributor:

- Figure 3.1B. Heatmap with the input data prepared by C. Dietrich
- Figure 3.1C. Coverage plots from RNA-seq data
- Figure 3.1D-F. Plots from data generated by A. Zirkel
- Figure 3.1G. FISH experiment
- Figure 3.1H. Fractionation experiment
- Figure 3.2A-C. Knock-down experiment and RNA-seq analysis
- Figure 3.2D-E. GSEA enrichment analysis
- Figure 3.2F-H. and Figure 3.2I Correlation between RNA-seq datasets, GO term enrichment in defined categories and metaplots of RNA-seq signal across defined pathways
- Figure 3.2J. Heatmap of SASP-relevant genes in RNA-seq data
- Figure 3.3A-D., 3.3I., and 3.3D. Analysis of isoform switching in circRAB3IP knock-down dataset
- Figure 3.3E. and 3.4.A. RAP-MS experiment performed together with K.K. Ebbesen
- Figure 3.3F. and 3.4.B. RNA immunoprecipitation experiment
- Figure 3.3G-H. comparison of circRAB3IP knock-down and SF3B1 public eCLIP-seq
- Figure 3.4C-D. comparison of circCAMSAP1 knock-down and RPS3 public eCLIP-seq
- Figure 3.4E-F. and 3.4A-D TNF α treatment of circCAMSAP1 and circRAB3IP overexpressing cells
- Figure 3.4G. MTT assay in overexpressing cells

- Figure 3.4I-J Exosome purification and transfer experiment outlined in Figure 3.4H.
- Figure 3.4K. qPCR test of lupus patient samples
- Figure 3.S2. Overrepresentation of transcription factor binding on differentially expressed genes in circRNA knock-down
- Figure 3.S2B. Correlation of knock-down datasets with public senescence RNA-seq
- Figure S2C-F. RNA-seq analysis of circRNA knock-down in HEK293 cells and correlation with the same in HUVEC
- Figure S4. Exosome purification and analysis of circRNA enrichment

circRNAs modulate cell proliferation and the immediate-early response by reorganizing gene expression and mRNA processing

Natasa Josipovic¹, Karoline K. Ebbesen^{2,3}, Anne Zirkel⁴, Adi Danieli¹, Christoph Dieterich⁵, Leo Kurian^{4,6}, Thomas Birkballe Hansen^{2,3}, Argyris Papantonis^{1,3,*}

¹ Institute of Pathology, University Medical Center Göttingen, 37075 Göttingen

² Department of Molecular Biology and Genetics (MBG), Aarhus University, Aarhus, Denmark

³ Interdisciplinary Nanoscience Centre (iNANO), Aarhus University, Aarhus, Denmark

⁴ Center for Molecular Medicine Cologne, University of Cologne, 50931 Cologne

⁵ Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Str. 9b, 50931 Cologne, Germany; Section of Bioinformatics and Systems Cardiology, University Hospital Heidelberg, 69120 Heidelberg, Germany

⁶ Institute of Neurophysiology, University of Cologne, 50931 Cologne

* Corresponding author: A.P

Abstract

Circular RNAs are endogenous long-lived non-coding RNA species often linked to regulatory roles in various processes relevant for sustaining cellular homeostasis. However, only circRNAs have been experimentally validated and functionally explored to date. Here, we attribute to two circRNAs novel roles as mediators of cellular homeostasis in primary human endothelial cells — circRAB3IP and circCAMSAP1. We show that depletion of either, results in the deregulation of hundreds of genes and senescence-like gene expression changes. Via its interaction with splicing factor SF3B1, circRAB3IP proved a potent mediator of cell cycle progression. Similarly, circCAMSAP1, via its interaction with ribosomal protein RPS3, could be assigned a role as inhibitor of early inflammatory response. Together, our findings link the maintenance of cell homeostasis to the presence of two co-regulated circRNAs.

Introduction

The question of non-coding RNA authenticity and, in many cases, functionality has generally been settled as a result of advances in the technologies cataloguing and studying transcriptomes. However, the circular structure and unconventional biogenesis of circular RNAs (circRNAs) (Salzman *et al*, 2013) led to them often being overlooked in transcriptome studies. circRNAs can be robustly detected across the evolutionary tree (Jeck & Sharpless, 2014; Wu *et al*, 2020), but their expression patterns are cell type- (Salzman *et al*, 2013), tissue- (Maass *et al*, 2017; Xia *et al*, 2017) and context-specific (Farooqi *et al*, 2021; Knupp & Miura, 2018; Venø *et al*, 2015; Lee *et al*, 2019b). It then follows that circRNA expression is highly regulated, advocating for their *in vivo* functionality.

Nevertheless, only recently did circRNAs emerge as important players in numerous cellular processes, such as normal development (Lee *et al*, 2019a), differentiation (Di Agostino *et al*, 2020), senescence (Du *et al*, 2016), immunity (Zhou *et al*, 2019c), cancer (Panda, 2018), Alzheimer's disease (Zhao *et al*, 2016), diabetes or senescence (Braicu *et al*, 2019). Mechanistically, most of the functions ascribed to circRNAs to date involve their interaction with microRNAs or RNA binding proteins (RBPs) (Du *et al*, 2016, 2017a; Abdelmohsen *et al*, 2017; Xia *et al*, 2018; Schneider *et al*, 2016). Thus, on one hand circRNAs have been shown to act as “sponges” sequestering various microRNAs and thereby indirectly regulating the pool of targets of these microRNAs (Hansen *et al*, 2013; Zhao *et al*, 2016; Li *et al*, 2015a). On the other hand, they may act as “decoys” or molecular scaffolds for RBPs, thereby affecting cognate gene processing (Zhang *et al*, 2013; Li *et al*, 2015d; Chen *et al*, 2018b) and translation (Ashwal-Fluss *et al*, 2014; Abdelmohsen *et al*, 2017). Despite compelling data on the mode-of-action of a handful of circRNAs, how this large class of endogenous non-coding RNAs contribute to cellular homeostasis remains rather poorly understood.

In this study, we investigated circRNAs enriched in the nascent RNA pool of primary human endothelial cells following TNF α stimulation. We focused on the exon-derived circCAMSAP1 and circRAB3IP to show that the former, already implicated in type-2 diabetes (Haque *et al*, 2020), cancer cell physiology (Zhou *et al*, 2020; Luo *et al*, 2021b; Chen *et al*, 2021) and chronic inflammation (Liu *et al*, 2019), also acts as a modulator of senescence-relevant inflammatory responses. We also show that the latter, circRAB3IP, can regulate the proliferation potency of cells and, thus, their path to replicative senescence. Together, we demonstrate how depleting a single circRNA from a human cell can profoundly remodel its transcriptome.

Results

Identification and characterization of circRNAs enriched in HUVEC nascent RNA

We first hypothesized that some circRNAs may be directly implicated in acute transcriptional responses. To catalogue these, we used primary human umbilical vein endothelial cells (HUVECs; pooled from three different donors) and short-term TNF α treatments. We then reasoned that circRNAs with a role in transcriptional regulation would be enriched in nascent RNA pools. To this end, we exploited our factory-seq approach (Melnik *et al*, 2016; **Fig. 3.1A**) to catalogue HUVEC nascent transcriptomes before and after 30 min of TNF α treatment (Cheng *et al*, 2016). Following data analysis, we were surprised to find that only two circRNAs, circCAMSAP1 (hsa_circ_0001900) and circRBA3IP (hsa_circ_0000419), satisfied a robust circRNA/mRNA detection ratio (DCC > 0.5), were present in both control and TNF α -treated samples (**Fig. 3.1B**), and could also be identified in total RNA-seq data (**Fig. 3.1C**). In RNA-seq data, circRNA genomic origins are often marked by signal accumulation at the exons and introns that contribute to their structure (Yang *et al*, 2011; Chen & Yang, 2015). CircCAMSAP1 and circRAB3IP arise from the sequences of exons 3 and 4, and 7 and 8, respectively (**Fig. 3.1C, Appendix Fig. 3.S1A**). They also show characteristic intronic RNA enrichment in factory-seq profiles, which is not seen in total transcriptome data (**Fig. 3.1C**) despite their far greater sequencing depth (**Appendix Fig. 3.S1B**).

Due to their covalently-closed circular structure, circRNAs are insensitive to degradation by RNase R and typically display half-lives longer than those of their mRNA counterparts (Jeck & Sharpless, 2014). Moreover, circRNAs are rarely translated (Jeck *et al*, 2013; Guo *et al*, 2014). Despite both circCAMSAP1 and circRAB3IP being previously detected in HUVEC RNA-seq data (Salzman *et al*, 2013; <http://www.circbase.org/>), we confirmed their resistance to RNase R (**Fig. 3.1D**), their lack of polysome association (**Fig. 3.1E**), as well as their increased stability *in vivo* (**Fig. 3.1F**). Thus, both our circRNA candidates are *bona fide* non-coding circRNAs.

circCAMSAP1 and circRAB3IP were enriched in nascent RNA pools, which could imply association with the active sites of transcription in the nucleus (Caudron-Herger *et al*, 2015). However, circRNAs are known to often localize in the cytoplasm (Salzman *et al*, 2013), which is also the case for circCAMSAP1 in colorectal cancer cells (Zhou *et al*, 2020). Therefore, we set out to investigate the subcellular distribution of these two circRNAs. To this end, we designed fluorescently-labelled 50-mers spanning the circRNA back-splicing junction to use as probes selectively detecting circCAMSAP1 or circRAB3IP in single-molecule RNA fluorescence *in situ* hybridization (RNA FISH) experiments (Zirkel & Papantonis, 2018). This approach revealed both nuclear and cytoplasmic circRNA localization in resting HUVECs (**Fig. 3.1G**). Interestingly, in ~5% of cells, circRNA fluorescence accumulated focally at the outer border of the nucleus (**Appendix Fig. 3.S1C**). We complemented this analysis with RT-qPCR performed on RNA from fractionated nuclear and

cytoplasmic compartments. This revealed a slight (in the case of circCAMSAP1) or high (in the case of circRAB3IP) circRNA enrichment in nuclei (Fig. 3.1H).

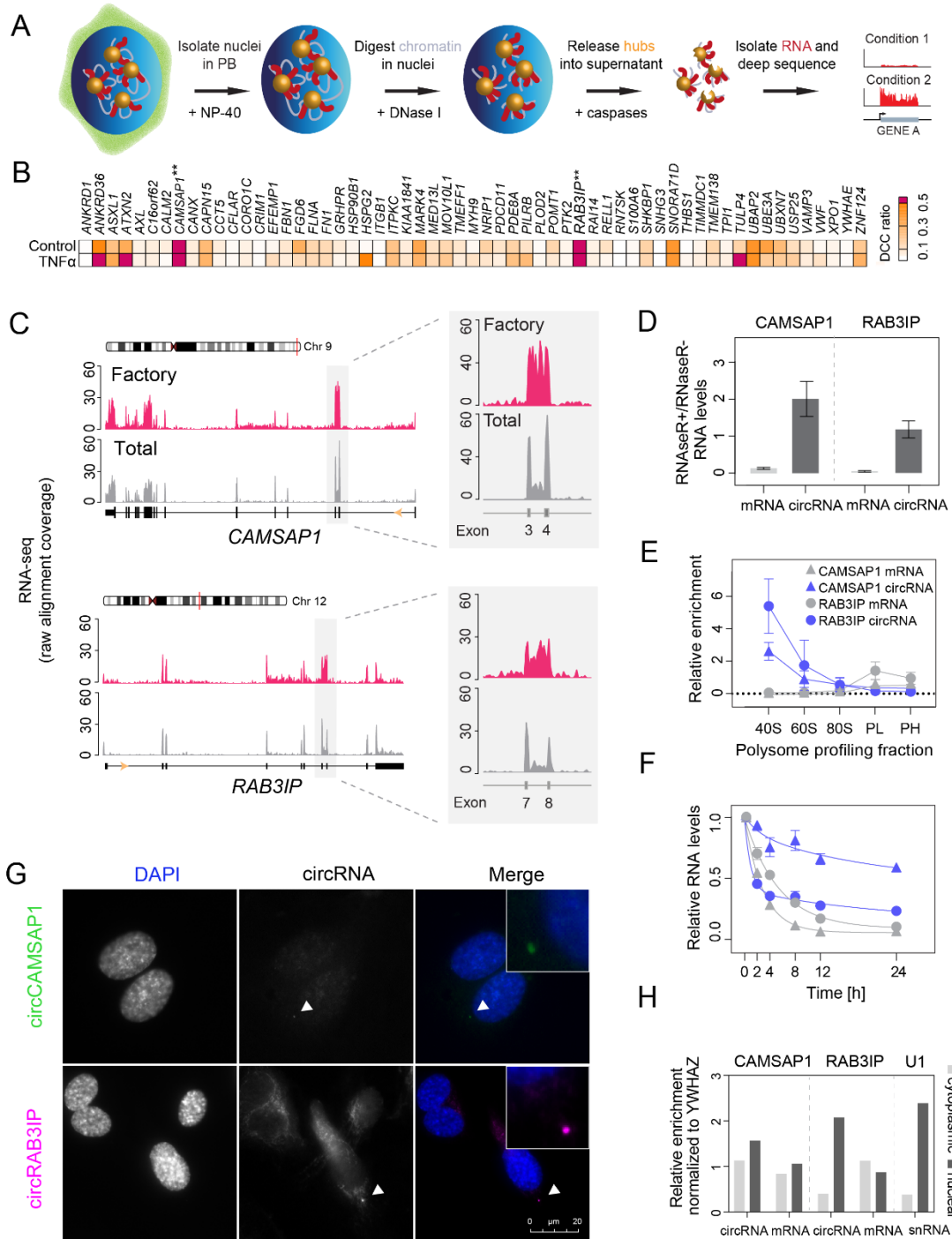


Figure 3.1. circRNAs enriched in nascent RNA during the acute inflammatory response.

(A) Factory-seq strategy overview. HUVECs are stimulated with TNF α for 0 or 30 min, intact nuclei collected in a physiological buffer, and treated with DNase I before Group-III caspases are used to detach transcription factories (*yellow*) from the nuclear substructure. Nascent RNA (*red*) is then purified and sequenced. (B) Heatmap showing relative enrichment of circRNAs over their linear counterparts in Factory-seq data represented as a DCC ratio. circRNA detection in both 0 and 30 min TNF α -treated samples, and a ratio > 0.5 were the criteria for selecting circRNAs (*pink*). (C) Exemplary Factory-seq coverage plots in the *CAMSAP1* and *RAB3IP* loci. Enrichment of intronic signal between two exons indicates circularization (*insets*). (D) Bar graph showing mean fold change of RNase R-treated compared to non-treated samples (\pm S.D., from three technical replicates) determined by RT-qPCR. (E) Relative enrichment of circRNA and their linear counterparts in polysome fractions normalized to whole cell extract levels (\pm S.D., from three technical replicates) determined by RT-qPCR. (F) Decay profiles of the RNAs from panel E, determined by RT-qPCR and fitted with least-squares fit. RNA levels (\pm S.D., from three technical replicates) are normalized to *YWHAZ* and plotted relative to the 0-h time point. Half-life for *CAMSAP1* mRNA = 2.0 h, *CAMSAP1* circRNA = 6.2 h, *RAB3IP* mRNA = 3.7 h, *RAB3IP* circRNA = 1.0 h. (G) Representative RNA FISH signal for circ*CAMSAP1* and circ*RBA3IP* in HUVECs counterstained with DAPI. Bar: 5 μ m. (H) Bar plot showing relative enrichment of circ*CAMSAP1* and circ*RBA3IP* or their linear counterparts in different subcellular fractions determined by RT-qPCR and normalized to *YWHAZ*. The nuclear U1 snRNA provides a control.

circRNA depletion leads to transcriptional reorganization reminiscent of senescence

Given their particular subcellular localization, we asked whether siRNA-mediated depletion of circ*CAMSAP1* or circ*RAB3IP* affected gene expression in HUVECs. Following circRNA-knockdown (KD) and total RNA-seq, principal component analysis indicated separation of the two biological replicates of each condition (**Fig. 3.2A**). This was reflected in the 3134 significant differentially expressed genes (DEGs) in circ*CAMSAP1*-KD (**Fig. 3.2B**) and 920 in circ*RAB3IP*-KD (**Fig. 3.2C**) given cut offs of absolute fold change (\log_2) \geq 0.6 and $P_{\text{adj}} \leq$ 0.05. We next performed gene set enrichment analysis, which revealed that the majority of genes upregulated upon circ*CAMSAP1*-KD contribute to immune response regulation, while downregulated ones to cell cycle progression and ribosome metabolism (**Fig. 3.2D**). Similarly, genes downregulated upon circ*RAB3IP*-KD are predominantly related to cell cycle regulation, as well as to viral responses and RNA metabolism (**Fig. 3.2E**). Unlike circ*CAMSAP1*-KD, the circ*RAB3IP*-KD did not lead to prominent upregulation of inflammatory genes (**Fig. 3.2D,E**). In addition, we analyzed enrichment of transcription factors known to regulate the DEGs in each circRNA-KD. The family of E2F factors, known for exerting cell cycle control (Dimova & Dyson, 2005), dominated analyses of downregulated genes in both knockdown datasets, while pro-inflammatory TFs like NF- κ B, STATs and IRFs (Platanitis & Decker, 2018) were overrepresented in relation to upregulated genes in circ*CAMSAP1*-KD data (**Appendix Fig. 3.S2A**).

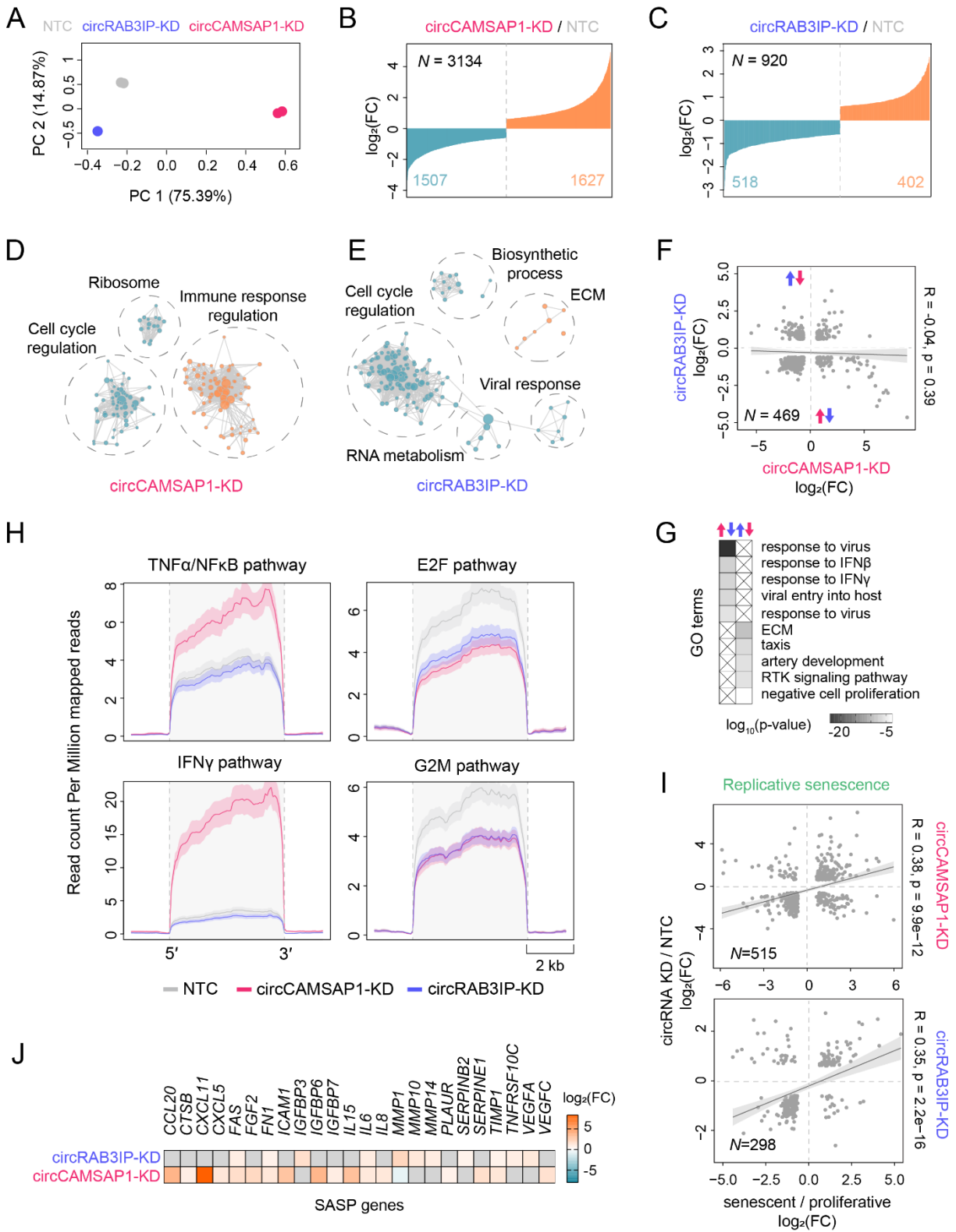


Figure 3.2. Knockdown of circCAMSAP1 or circRAB3IP leads to major transcriptional changes in HUVEC.

(A) PCA clustering of circCAMSAP1- (*magenta*) and circRAB3IP-knockdown (*blue*) and non-targeting control (*grey*) RNA-seq data. (B) Bar plot showing fold change values (\log_2) for up- (*orange*) and downregulated genes (*blue*) upon circCAMSAP1 knockdown at $p_{\text{adj}} \leq 0.05$. (C) As in panel B, but for circRAB3IP knockdown data. (D) Gene set enrichment analysis of the differentially-expressed genes (DEGs) from panel B. Related terms are shown clustered (*circles*) and labelled by the group-dominant gene ontology term. (E) As in panel D, but for circRAB3IP-knockdown DEGs from panel C. (F) Scatter plot correlating circCAMSAP1- and circRAB3IP-knockdown DEG. The Pearson correlation coefficient (R) and associated p-value are calculated for the overlapping DEGs indicated (N). (G) GO term enrichment analysis for those genes from panel F showing divergent regulation in each knockdown experiment. (H) Metaplots of circCAMSAP1 (*pink*) and circRAB3IP (*blue*) RPM-normalized coverage across relevant inflammatory and cell cycle-related genes (from <https://www.gsea-msigdb.org/gsea/msigdb/>). (I) As in panel F, but correlating circCAMSAP1- and circRAB3IP-knockdown DEGs with those from HUVEC replicative senescence. Pearson correlation coefficients (R) and p-values (p) are calculated for the overlapping DEGs indicated (N). (J) Heatmap of SASP genes (from <https://reactome.org/>) that are also DEGs in the circCAMSAP1- and circRAB3IP-knockdown data. The color code reflects fold change values (\log_2) at $p_{\text{adj}} \leq 0.05$.

The gene expression signatures in the two knockdown datasets implied partially overlapping, yet distinct functions for each circRNA. To assess the nature of these functions, we first attempted to correlate the ~500 overlapping DEGs in the two datasets. However, no significant correlation was observed ($R=0.04$; **Fig. 3.2F**), and many genes were divergently regulated in the two knockdown experiments. By focusing on these divergently regulated genes, we found that interferon signaling is stimulated upon circCAMSAP1-, but not in circRAB3IP-KD, while the converse applies to extracellular matrix (ECM) remodeling and receptor tyrosine kinase (RTK) signaling (**Fig. 3.2G**). We validated this by plotting normalized RNA-seq coverage across genes participating in knockdown-specific pathways. Proinflammatory genes in the NF- κ B/TNF α and IFN γ pathways were strongly stimulated upon circCAMSAP1-KD, whereas both knockdown experiments suppressed E2F-controlled pathways and the G2/M cell cycle transition compared to control cells (**Fig. 3.2H**).

We persisted in analyzing the induction of cell cycle arrest that follows the knockdown of both circCAMSAP1 and circRAB3IP as, complemented by inflammation and ECM remodeling, it was very much reminiscent of the changes accompanying replicative senescence entry by HUVECs (Zirkel *et al*, 2018). To this end, we compared each knockdown dataset with DEGs from senescent HUVECs (Zirkel *et al*, 2018) to discover significant correlation ($R > 0.35$; **Fig. 3.2I**). In fact, almost 30% of circRAB3IP-KD and 16% of circCAMSAP1 DEGs are also regulated the same way upon senescence entry. Moreover, even stronger correlation was also observed when comparing a consensus gene expression signature of senescence (Hernandez-Segura *et al*, 2018) with our knockdown data ($R > 0.55$; **Appendix Fig. 3.S2B**), and any genes

showing divergent regulation did not enrich for some particular pathway of GO term. Senescent cells also display cell type-specific secretory phenotypes (i.e., SASP) comprised of molecules acting in a paracrine fashion to promote inflammation and senescence (Coppé *et al*, 2010). Such SASP genes were indeed also induced in our circRNA-KD experiments (**Fig. 3.2J**) and represented ~1/3 of all relevant genes found in the Reactome database (<https://reactome.org/>). In line with our GSEA analysis, circCAMSAP1-KD induced the largest number of SASP genes. Taken together, we could show that depletion of a single circRNA suffices for extensive remodeling of the HUVEC homeostatic gene expression program.

circRAB3IP affects the processing of cell cycle-related transcripts

Recent studies in the context of cancer have suggested that circRNAs may regulate cellular proliferation via microRNA “sponging” (Zeng *et al*, 2018; Li *et al*, 2018; Zheng *et al*, 2016). However, this requires multiple cognate binding sites for the sequestered microRNA in cognate circRNA sponge, which does not seem to be the case for circRAB3IP (targetscan.org). Other recent research has implicated circRNAs in the regulation of splicing (Li *et al*, 2015d; Ashwal-Fluss *et al*, 2014). By reanalyzing our knockdown data, we discovered 231 genes displaying significant (> 30%) change in isoform usage upon circRAB3IP depletion. By contrast, circCAMSAP1-KD only induced such changes in 40 genes. circRAB3IP-related isoform switching is not reflected in differential mRNA expression as < 6% of DEGs in knockdown data also switch isoforms (**Fig. 3.3A**). GO term analysis revealed that most genes displaying differential isoform usage are associated with cell cycle progression and cytoskeleton and centriole regulation, likely enhancing the cell cycle arrest phenotype that follows circRAB3IP-KD (**Fig 3.3B**). As regards the functional consequences of such isoform usage changes, we found that > 90 out of 231 genes now expressed isoforms containing fewer recognizable domains and/or shorter open reading frames (**Fig. 3.3C**). Changes in isoform choice are typically the result of alternative splicing. We therefore complemented our isoform usage data with alternative splicing analyses, which revealed significant and widespread splicing perturbations mostly involving skipping of canonical 5’ and 3’ splicing sites and intron retention that can lead to shorter or non-coding transcripts (**Fig. 3.3D**).

The discovery of widespread splicing changes in circRAB3IP-KD data, prompted us to ask how this circular RNA might exert its action. Given that we could detect no significant miRNA binding potential in circRAB3IP, we set out to explore circRAB3IP association with proteins. To this end, we performed RNA antisense pull-down coupled with mass spectrometry (RAP-MS, McHugh & Guttman, 2018; **Appendix Fig. 3.3A**) in HEK293 cells. This choice of cell type was dictated by the need to obtain large counts (approx. 400 x 10⁶ per replicate) of cells overexpressing the circRNA of interest (**Appendix Fig. 3.3B,C**). Prior to RAP-

MS, we performed knockdown studies in HEK293 largely confirming the transcriptional effects observed in HUVECs, except for cell cycle arrest due to the immortalized nature of HEK cells (**Appendix Fig. 3.S2C-F**). Following circRAB3IP pulldown and mass spectrometry, we could robustly identify 16 proteins as its potential protein partners. The majority of these proteins are functionally linked to cell differentiation and cell death, mRNA splicing and stabilization or chemokine production and immune responses (**Fig. 3.3E**).

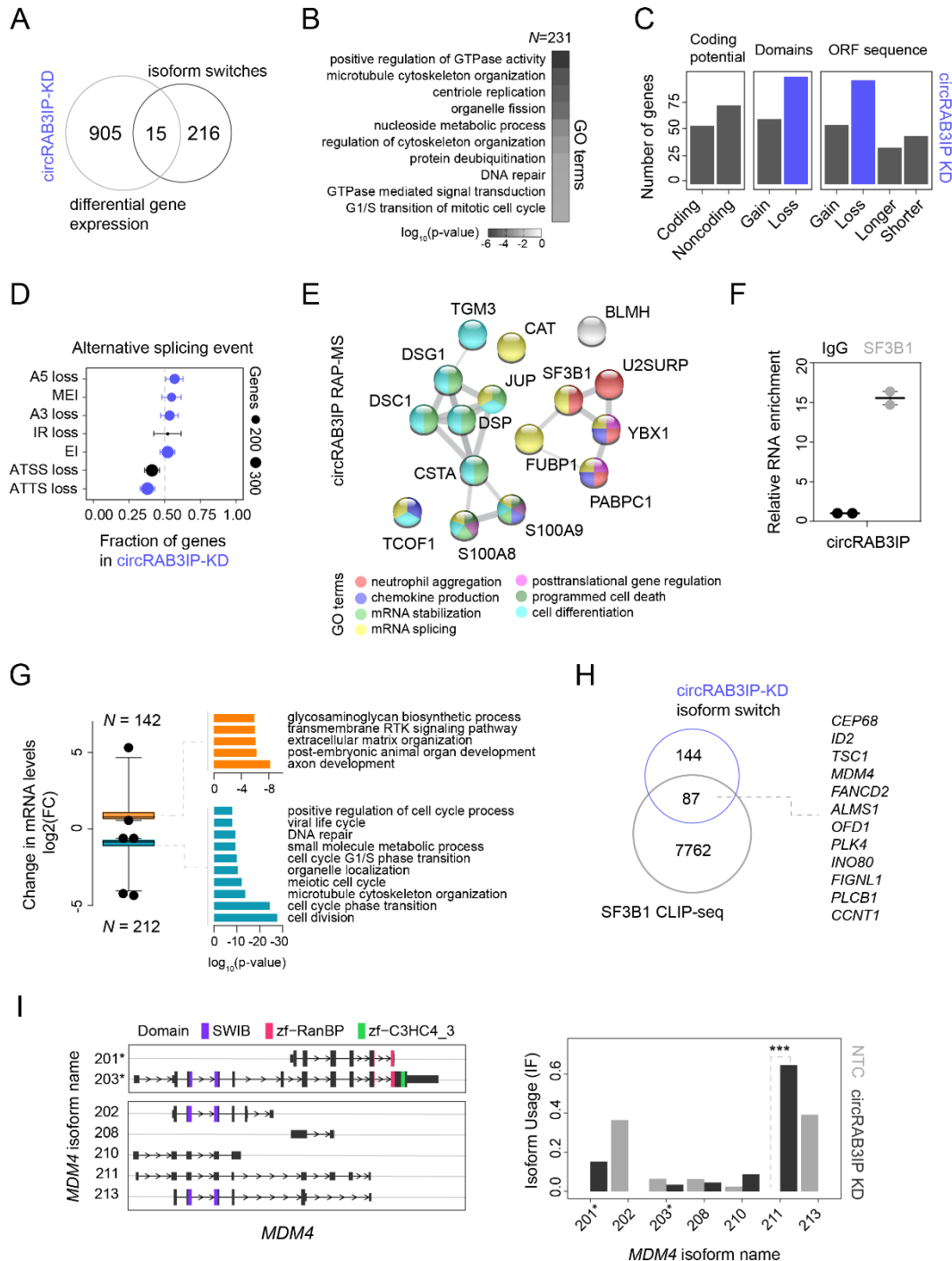


Figure 3.3. circRAB3IP binding to SF3B1 affects RNA splicing.

(A) Venn diagram showing overlap between DEGs and genes displaying isoform usage changes in circRAB3IP-knockdown data. ns: $p = 0.6711$, Fisher's exact test. (B) Heatmap showing overrepresented GO terms for the number of genes (N) displaying isoform usage switch of $\geq 30\%$ upon circRAB3IP knockdown. Color code reflects GO enrichment significance (log). (C) Bar plot showing the type of isoform usage change (of $\geq 30\%$) in circRAB3IP-knockdown data. Statistically-significant changes (blue) are supported by an FDR < 0.05 . (D) Genome-wide changes of alternative splicing events (with $\geq 10\%$ change) in circRAB3IP-knockdown data. Dot size indicates the number of genes and significant changes (blue) are supported by an FDR < 0.05 . A5, alternative 5' donor site; A3, alternative 3' donor site; MEI, multiple exon inclusion; ES, exon inclusion; IR, intron retention; ATSS and ATTS, alternative transcription start and termination sites, respectively. (E) Protein network showing circRAB3IP-bound proteins and their interactions inferred from STRING database (from two biological replicates; <https://string-db.org/>). The color code reflects overrepresented GO terms (below) associated with each protein. (F) Box plot showing enrichment of circRAB3IP in SF3B1 pulldown experiments is normalized to input (from two technical replicates). IgG pulldowns provide a control. (G) Box plot (left) showing fold change values (\log_2) of up- (orange) or downregulated SF3B1-bound mRNAs (blue) upon circRAB3IP knockdown. The number of mRNAs (N) and GO terms significantly associated with each group (right) are shown. (H) Venn diagram showing genes shared between SF3B1 CLIP-seq (gray) and differentially-spliced isoforms upon circRAB3IP-knockdown (blue). *: $p = 0.0001$, Fisher's exact test. Cell cycle relevant genes from the overlap are listed on the right (<https://gsea-msigdb.org>). (I) Illustration (left) showing isoform switching upon circRAB3IP-knockdown for the *MDM4* gene. Functional domains in each detected isoform (colour code, top) and coding isoforms (*) are indicated. Bar plot (right) showing usage changes for each detected isoform between knockdown (black) and control data (grey). ***: FDR ≤ 0.001 .

We chose to focus on SF3B1, a spliceosomal RNA binding protein implicated in mRNA splicing (Sun, 2020), cellular senescence (Yin *et al*, 2019) and aging (Holly *et al*, 2014). First, we orthogonally confirmed its association with circRAB3IP via RNA immunoprecipitation (Fig. 3.3F). Next, we analyzed publicly available SF3B1 enhanced crosslinking and immunoprecipitation (eCLIP) data (ENCSR133QEA) against DEGs from our circRAB3IP-KD. This revealed that $\sim 40\%$ of these DEGs are also bound by SF3B1 and linked to GO terms of cell cycle regulation and ECM remodeling (Fig. 3.3G), both central to the senescent-like phenotype observed in knockdown cells (Fig. 3.2D,E,I). We also analyzed the eCLIP data against mRNAs that switch isoforms upon circRAB3IP-KD. Again, $\sim 40\%$ of these mRNAs were also bound and potentially regulated by SF3B1, with some of them involved in cell cycle progression (Fig. 3.3H). In fact, this list contained 12 key cell cycle regulators like *MDM4*, *ID2*, *ALMS1*, *PLK4*, *FANCD2*, *CCNT1* and *PLCB1* (<https://www.gsea-msigdb.org/gsea/msigdb/genesets.jsp>, GO:0007049). For all these, strong changes in coding versus non-coding/domain-loss isoform usage help explain cell cycle effects in cells lacking circRAB3IP (Fig. 3.3I, Appendix Fig. 3.S2C). Notably, we observed no significant overlap between DEGs from circRAB3IP- and SF3B1-KD experiments, suggesting that the circRAB3IP-SF3B1 interaction concerns only the part of the SF3B1 regulatory repertoire leading to proliferative arrest.

circCAMSAP1 dampens the response to acute inflammatory signaling

Much like in the case of circRAB3IP, circCAMSAP1 has recently been shown to act as a microRNA “sponge” in various cancer cells (Zhou *et al*, 2020; Chen *et al*, 2021; Luo *et al*, 2021b), but its sequence does not appear to contain repeated microRNA-binding sequences (<http://www.targetscan.org/>). Thus, we approached its functional role by RAP-MS (**Appendix Fig. 3.3A**) in HEK293 cells that show pro-inflammatory signatures upon circCAMSAP1-KD and reasonable correlation to HUVEC knockdown data (**Appendix Fig. 3.32C-F**). Overexpressed circCAMSAP1 co-purified with 27 proteins, the majority of which could be linked to inflammation, regulation of gene expression or RNA metabolism (**Fig. 3.4A**). Compared to the circRAB3IP RAP-MS catalogue (**Fig. 3.3E**), only U2SURP, a known component of the general splicing machinery (Will, 2002), was shared. However, none of our two circRNA-KD datasets correlated with previously published U2SURP-KD (not shown, $R = -0.035$ and -0.25 for circRAB3IP-KD and circCAMSAP1-KD, respectively; De Maio *et al*, 2018).

Given the characteristic pro-inflammatory signature in the circCAMSAP1-KD data, we focused on RPS3, a protein implicated in cell cycle progression (Jang *et al*, 2012), E2F1 co-activation (Lee *et al*, 2010), and inflammation via NF- κ B (Wan *et al*, 2007) and the NF- κ B inhibitor, I κ B α (Stanborough *et al*, 2014). First, we verified interaction detected by RAP-MS using RNA immunoprecipitations (**Fig. 3.4B**). Next, by intersecting the publicly available RPS3 eCLIP data with DEGs from circCAMSAP1-KD experiments (**Fig. 3.4C**), we found that > 20% of RPS3-bound mRNAs are also differentially-regulated upon circCAMSAP1 depletion from HUVECs. Intriguingly, approximately half of these genes were upregulated and linked to the aging process, inflammation, and EMC remodeling – all senescence hallmarks (**Fig. 3.4D**). The rest were downregulated and linked to cell cycle regulation (**Fig. 3.4D**), thus explaining the key effects of the circCAMSAP1-KD (**Fig. 3.2D**).

We then speculated that the circCAMSAP1-RPS3 interaction might be prototypically explaining how this circRNA modulates the inflammatory response. Thus, to dissect this, we exploited our circCAMSAP1-pcDNA3 HEK293 overexpressing line (**Appendix Fig. 3.33B**) to assess whether elevated titers of this circRNA can interfere with pro-inflammatory stimulation. We treated overexpressing cell with the TNF α cytokine for 0, 1, and 4 h and saw that circCAMSAP1 levels do not change significantly (**Appendix Fig. 3.34A**), in line with equal circCAMSAP1 detection in our RNA-seq data from HUVECs before and after 30-min TNF α treatment. However, this constant presence of circCAMSAP1 led to a strong inhibition of *TNFAIP3* and *CCL20* expression compared to empty-vector control (**Fig. 3.4E**). circRAB3IP-overexpressing HEK293 cells

do not display circRAB3IP change as well (Appendix Fig. 3.S4B), and do not exhibit such effects on TNF α responsive genes (Appendix Fig. 3.S4C).

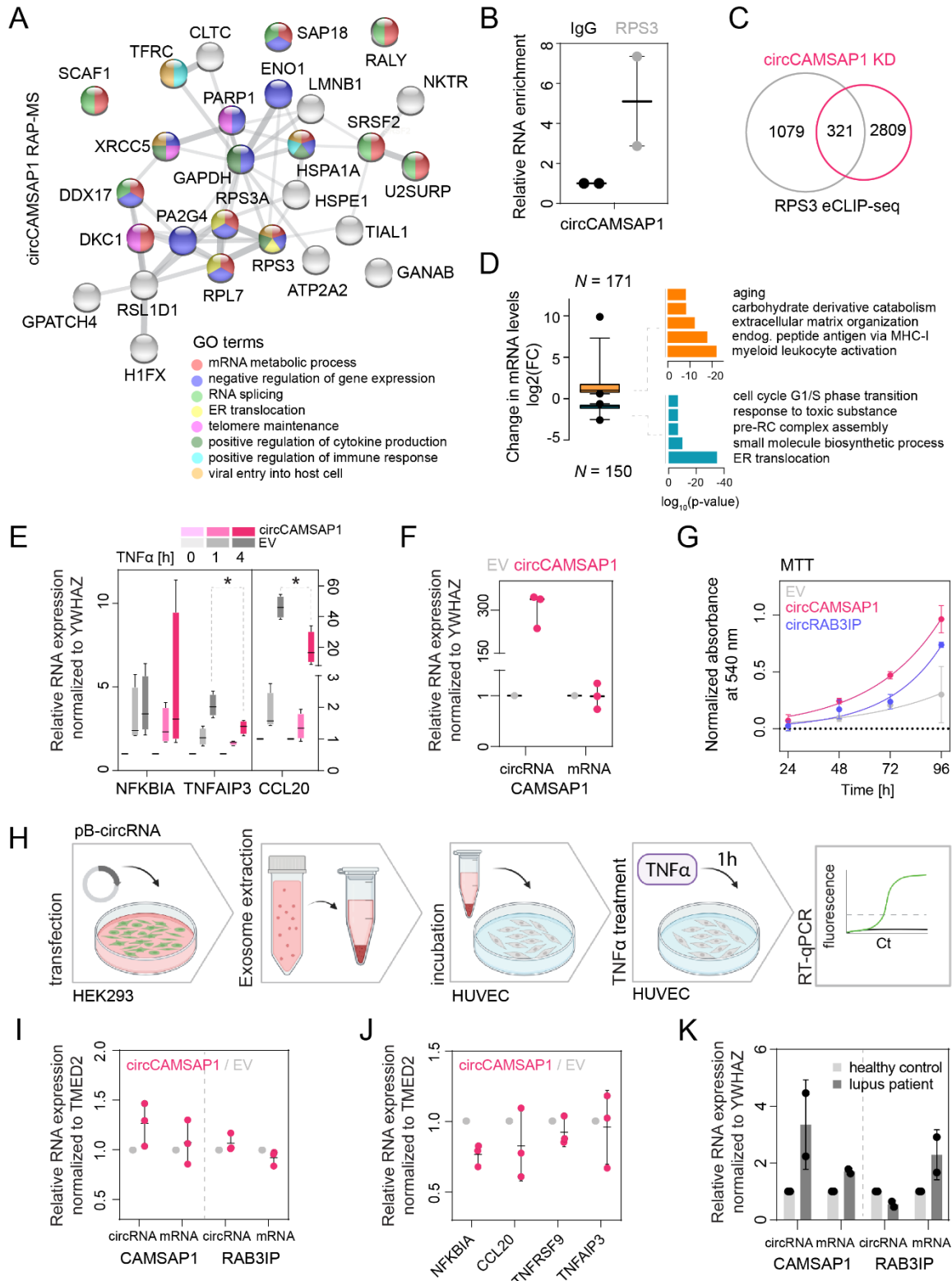


Figure 3.4. circCAMSAP1 levels modulate inflammatory responses.

(A) Protein network showing circCAMSAP1-bound proteins and their interactions inferred from STRING database (from two biological replicates; <https://string-db.org/>). The color code reflects overrepresented GO terms (below) associated with each protein. (B) Box plot showing enrichment of circCAMSAP1 in RPS3 pulldown experiments is normalized to input (from two technical replicates). IgG pulldowns provide a control. (C) Venn diagram showing overlap between DEGs and genes displaying isoform usage changes in circCAMSAP1-knockdown data. *: $p = 0.018$; Fisher's exact test. (D) Box plot (*left*) showing fold change values (\log_2) of up- (*orange*) or downregulated RPS3-bound mRNAs (*blue*) upon circCAMSAP1 knockdown. The number of mRNAs (*N*) and GO terms significantly associated with each group (*right*) are shown. (E) Box plot showing the response of the *NFKBIA*, *TNFAIP3*, and *CCL20* genes to TNF α for the times indicated in HEK293 stably overexpressing circCAMSAP1 or empty vector cells (EV). Relative RNA levels at each time point are normalized to *YWHAZ* and represented as fold change over non-treated samples. *: $p \leq 0.05$, unpaired two-tailed Student's t-test. (F) Box plot showing pB-mediated circCAMSAP1 overexpression in HEK293 cells (three technical replicates) normalized to *YWHAZ* levels and represented as fold change over empty-vector (EV) levels. (G) MTT proliferation assays following induced circCAMSAP1 (*magenta*) or circRAB3IP (*blue*) overexpression in HEK293 cells. Relative absorbance is plotted relative to the 24-h time point. Data was fit with least squares fit and exponential growth calculated as doubling time for circCAMSAP1 = 17.37 h, circRAB3IP = 22.72 h, empty vector control (EV) = 28.10 h. (H) Overview of exosome-transfer strategy. circRNA-pB HEK293 lines are induced and grown in exosome-depleted medium. Exosomes are purified and added to growing HUVECs. 24 h later, 1-h TNF α treatment is performed, and RNA is extracted for RT-qPCR analysis. (I) Box plot showing relative RNA expression for circRNAs and their linear counterparts normalized to *TMED2* and represented as fold change over empty vector (EV) controls. (J) As in panel I, but for the TNF α -responsive genes indicated. (K) Bar plot showing relative RNA expression (\pm S.D.) of circRNAs and their linear counterparts normalized to *YWHAZ* levels in Lupus patients and healthy donors.

During the aforementioned experiments, we observed that maintaining these overexpressing lines in long-term culture resulted in gradual decline in the produced circRNAs. We reasoned that this reflects some sort of negative selection in response to abnormal circRNA levels, and decided to generate inducible lines. We achieved this by stably integrating the circRNA-producing constructs using a piggyBAC (pB) vector in HEK293 (see **Methods** for details), and verified consistent inducible overexpression of circCAMSAP1 (**Fig. 3.4F**) or circRAB3IP over time (**Appendix Fig. 3.S4D**). Interestingly, circCAMSAP1 (and to a lesser extent circRAB3IP) was also enriched in extracellular vesicles collected from cell cultures (**Appendix Fig. 3.S4E**), in line with reports on possible paracrine roles of circRNAs (Li *et al*, 2015c; Dou *et al*, 2016). Thus, we exploited our pB-circRNA-overexpressing HEK293 lines to produce and collect exosomes and test the possibility of paracrine signaling via circCAMSAP1. Exosomes from circCAMSAP1-overexpressing or empty vector cells were then transferred onto HUVEC cultures treated with TNF α (**Fig. 3.4H**). Although the uptake of circCAMSAP1 by HUVECs was only mild (**Fig. 3.4I**), it did dampen the stimulation of the TNF α -responsive *NFKBIA* gene, but not of *TNFRSF9*, *CCL20* or *TNFAIP3* (**Fig. 3.4J**), all of which were upregulated upon

circCAMSAP1-KD. This suggests that elevated circCAMSAP1 levels can interfere with the pro-inflammatory response, and we would expect to see them perturbed also in the context of chronic inflammatory disease. Indeed, performing RT-qPCR on RNA isolated from kidney sections of Lupus nephritis patients, we recorded significantly elevated levels of circCAMSAP1, but not of circRAB3IP, compared to healthy specimens (**Fig. 3.4K**).

Finally, given that cell cycle arrest was a prominent feature of both our circRNA knockdowns, we reasoned that their overexpression could lead to increased proliferation, in line with what has been implied for circCAMSAP1 in cancer cells (Zhou *et al*, 2020; Chen *et al*, 2021; Luo *et al*, 2021b; Liu *et al*, 2019). Indeed, we performed MTT assays showing increased proliferation rates for both circCAMSAP1 and circRAB3IP overexpressing HEK293 cells (**Fig. 3.4G**; circCAMSAP1, circRAB3IP and empty vector control doubling time = 17.37, 22.72, and 28.10 hours, respectively).

Discussion

Despite a growing body of evidence on the roles of circRNAs in various cancer lines, how circRNAs function in normal primary human cells remains largely unexplored. Here, we address this question using human primary endothelial cells to provide the first functional characterization for circRAB3IP and describe a novel mode of action to circCAMSAP1. While the motivation for studying these two circRNAs originated in their converging pro-inflammatory effects, our results showed that circCAMSAP1 and circRAB3IP exhibit both functional specificity and distinct interactomes to undertake their roles.

As regards circRAB3IP, we can assign to it a role in cell homeostasis maintenance, as a prominent senescence-like phenotype follows its depletion. Through correlative transcriptome association with own senescence data from HUVECs and hallmark gene expression signature from different senescence models (Hernandez-Segura *et al*, 2018) we found that circRAB3IP-induced changes to the transcriptome counteract physiological cell cycle progression. Moreover, circRAB3IP depletion also induces genes relevant to ECM remodeling, another senescence hallmark (Levi *et al*, 2020). In contrast, we observed that overexpressing circRAB3IP suffices for accelerated proliferation. We approached the molecular mechanism underlying these effects by cataloguing proteins bound to circRAB3IP and focusing on the general splicing factor SF3B1, loss of which has been implicated in cell cycle arrest (Dolatshad *et al*, 2015). We speculate that the SF3B1-circRAB3IP interaction modulates the splicing factor's efficacy and, in the absence of the circRNA, changes to the transcriptome arise that not only help explain cell cycle arrest (e.g., via the emergence of non-coding

isoforms of the *MDM4* cell cycle regulator), but also resemble those seen in cellular senescence and aging (Deschênes & Chabot, 2017).

Compared to the uncharacterized circRAB3IP, circCAMSAP1 has been studied in the context of cell cycle progression in cancer (Zhou *et al*, 2020; Chen *et al*, 2021; Luo *et al*, 2021b) and chronic inflammation (Liu *et al*, 2019). The former involves circCAMSAP1 acting as a “sponge” for various microRNAs, whereas the latter implicates its RNA folding for interaction with PKR. Here, we describe a different and multifaceted role for circCAMSAP1 in HUVEC homeostasis, also justified by cell type-specific needs. circCAMSAP1 depletion has overlapping but not fully convergent effects to circRAB3IP depletion, affirming the prevailing view of circRNAs as highly specialized molecules carrying out disparate, cell type-specific functions. In addition, while we saw senescence-like cell cycle effects in circCAMSAP1-depleted cells, a prominent pro-inflammatory signature also emerged, suggesting that circCAMSAP1 could play a role in regulating a subset of inflammation-responsive genes. These included a considerable fraction of known SASP-related genes. To this end, the interaction of circCAMSAP1 with the ribosomal protein RPS3, which was recently implicated in NF- κ B signaling via its translocation to the nucleus to direct NF- κ B dimers to a subset of their target genes (Wan *et al*, 2007; Stanborough *et al*, 2014). Like in the case of circRAB3IP, many of the mRNAs differentially regulated upon circCAMSAP1-KD are also bound by RPS3, and circCAMSAP1 overexpression or paracrine action dampen the ability of immediate-early pro-inflammatory genes to be activated. We speculate that this occurs as a result of RPS3 (and possibly other cofactors) sequestration by circCAMSAP1. Then, it follows that the increased circCAMSAP1 levels in patients with autoimmune Lupus nephritis are a response mechanism put in place to restrict chronic pro-inflammatory signaling. Finally, it is worth noting that circCAMSAP1 was also shown to influence E2F factor activity in the context of neurons (Lee *et al*, 2010), which might help explain the cell cycle effects arising from its knockdown in HUVECs.

Together, our results support the notion that modulating the levels of a single circRNA may drive profound changes to gene expression and cellular homeostasis, although the molecular details that allow these long-lived but moderately abundant RNA species to exert their function require further study.

Materials and methods

Cell culture and circRNA verification

HUVECs from pooled, apparently healthy, donors (passage 4-9; Lonza) were grown and passaged in complete Endopan-3 medium kit (PAN-Biotech) under 5% CO₂. HEK293 cell line (kindly gifted by Dr. Anne Zirkel) was passaged in DMEM (Sigma) supplemented with 10% FBS (Gibco) and 1% Pen/Strep (Gibco) under

5% CO₂. For RNase R treatment experiments, proliferating HUVEC were grown in complete media, RNA extracted with Direct-Zol RNA purification kit (Zymo) and treated with 1.7 U of RNase R per µg of RNA (Epicenter) for 13 minutes at 37 °C. Fractionation experiments on HUVECs were performed by nuclei extraction in Physiological buffer (PB; 100 mM CH₃COOK, 30 mM KCl, 10 mM Na₂HPO₄, 1 mM MgCl₂, 1 mM Na₂ATP, 1 mM DTT, 10 mM β-glycerophosphate, 10 mM NaF, 0.2 mM Na₃VO₄, 25 U/ml RiboLock RNase Inhibitor (Thermo-Fisher), and 1:1000 dilution of cOmplete Protease Inhibitor Cocktail (Roche); buffer is adjusted to pH 7.4 with KH₂PO₄) for 30 minutes on ice, followed by 5 minute centrifugation at 1000 x g. Supernatant was collected as cytoplasmic fraction in TRIzol, while nuclear pellet was washed once more in PB before collection in TRIzol and purification with Direct-Zol RNA purification kit (Zymo). Polysome profiling was performed as previously described (Bartsch *et al*, 2018).

Factory-seq protocol and detection of circRNAs

Preparation of nascent RNA from ~10 x 10⁶ HUVEC was performed as previously published (Melnik *et al*, 2016). Briefly, cells were grown and passaged in complete Endopan-2 medium (Lonza) until 90% confluency and scraped in PBS. Nuclei isolation was performed for 30 minutes in ice cold Physiological buffer (PB; 100 mM CH₃COOK, 30 mM KCl, 10 mM Na₂HPO₄, 1 mM MgCl₂, 1 mM Na₂ATP, 1 mM DTT, 10 mM β-glycerophosphate, 10 mM NaF, 0.2 mM Na₃VO₄, 25 U/ml RiboLock RNase Inhibitor (Thermo-Fisher), and 1:1000 dilution of cOmplete Protease Inhibitor Cocktail (Roche); buffer is adjusted to pH 7.4 with KH₂PO₄) and cell lysis efficiency was assessed under the microscope. Nuclei were collected by centrifugation and DNase I treated for 30 min at 37°C (Worthington, 30U per 5 x 10⁶ cells). After the chromatin digestion, nuclei were pelleted and lysed in native lysis buffer (NLB; 40 mM Tris-acetate pH 7.4, 2 M 6-aminocaproic acid and 0.2 M supplemented with 50 U/ml RiboLock RNase Inhibitor and 1:1000 dilution of cOmplete Protease Inhibitor Cocktail) for 20 minutes on ice. Lysates were further treated with 2U per sample of Caspase mix (Active human caspases, group III (6, 8, 9, and 10) mix; BioCat) for 30 minutes with vigorous shaking. Supernatants are collected in TRIzol (Invitrogen) and RNA was precipitated with Direct-Zol RNA purification kit (Zymo). RNA integrity was assessed on Bioanalyzer and sequenced to approximately 25 x 10⁶ paired end reads. Sequencing quality was assessed with FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and reads were aligned to human reference genome with STAR (hg19). For detection of circRNAs enriched in nascent RNA fraction, DCC algorithm was applied (Cheng *et al*, 2016). Raw read coverage plots were generated with Gviz package (Hahne & Ivanek, 2016).

Fluorescence in situ hybridization experiment

FISH was performed with back-splicing junction-targeting fluorescent probes for circCAMSAP1 and circRAB3IP as previously described (Zirkel & Papantonis, 2018). Briefly, HUVEC cells were grown on “etched” coverslips treated with 0.1% hydrofluoric acid until 70% confluent and fixed in fixation buffer for 15 minutes (10 ml of buffer contains 1 ml of 9% NaCl, 2.5 ml of 16 % paraformaldehyde (Electron Microscopy Sciences), 500 μ l of glacial acetic acid and 6 ml of RNase-free water). Coverslips are then washed in PBS and permeabilized in 0.5% Triton-X / 0.5% saponin for 5 minutes. After washing in PBS, coverslips are post-fixed in 3.7% formaldehyde for 5 minutes, washed in PBS and dehydrated in ethanol (70%, 90% and absolute) for 3 min. Hybridization of probes to fixed and permeabilized cells was carried out in hybridization buffer (25% formamide, 2 \times SSC, 200 ng/ μ l sheared salmon sperm DNA (10 mg/mL; Invitrogen), 5 \times Denhardt’s (0.1% Ficoll 400, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin), 50 mM phosphate buffer (20 mM KH_2PO_4 , 30 mM $\text{KHPO}_4 \cdot 2\text{H}_2\text{O}$, pH 7.0), and 1 mM EDTA) mixed with fluorescent probe (25 ng/ μ l) in 9 : 1 ratio. Hybridization mix was denatured at 90°C for 10 minutes, quenched on ice and added to coverslips. Hybridization was performed overnight on 37°C in dark. The next day, coverslips are washed 3x in 2x SCC, one time in RNase-free water and mounted on glass slides with ProLong Gold Antifade reagent with DAPI (Invitrogen). Imaging was performed on Leica DMI8 platform under 100x magnification objective (oil), and images were exported with LAS X software (Leica).

siRNA-mediated circRNA knockdown and analysis

Approximately 35×10^4 HUVEC cells were seeded per 60 mm plate day before the transfection. Each siRNA (180 pmol) was mixed with Opti-MEM (Gibco) and Lipofectamine RNAiMAX reagents according to manufacturer’s instructions (Thermo Fisher). Cells were washed once in PBS and incubated in 2.5 ml of Opti-MEM and 500 μ l of transfection mix for 4 hours, after which the medium was replaced with complete Endopan-3 medium. 48 hours post-transfection, cells were collected in TRIzol (Invitrogen) and RNA extraction was performed with Direct-Zol RNA purification kit (Zymo). RNA integrity was verified with Bioanalyzer, and samples were sequenced to at least 25×10^6 paired end reads. Sequencing quality check was performed with FastQC and reads were aligned to human reference genome (hg19) with default STAR aligner settings (Dobin *et al*, 2013). Aligned reads quantification was performed with *featureCounts* (Liao *et al*, 2014), selecting for uniquely aligned and properly matched read pairs (selected options --primary -p -B -C). Between-sample normalization of raw quantified reads was performed with RUVs method of RUVseq package (Risso *et al*, 2014), that estimates factors of unwanted variation using those genes in the replicates for which the covariates of interest remain constant. Differential gene expression (DEG) was

estimated with DESeq2 package (Love *et al*, 2014), wherein genes with FDR <0.01 and absolute (\log_2) fold-change were considered significantly differentially expressed. DEG list was used as an a priori defined gene list for gene set enrichment analysis (GSEA; Mootha *et al*, 2003; Subramanian *et al*, 2005) to assess the overrepresentation of GO terms in the RNA-seq datasets and produce network plots with Cytoscape (Franz *et al*, 2016). Read counts of mapped RNA-seq were normalized per million mapped reads (RPM) and plotted with ngs.plot (Shen *et al*, 2014) in order to visualize signal enrichment at the genes of chosen pathways from GSEA Molecular Signatures Database (MSigDB IDs: TNF α via NF κ B signalling, M5890; IFN- γ response, M5913; E2F targets, M5925; G2M checkpoint, M5901).

circRNA overexpression constructs

Expression vectors for circCAMSAP1 and circRAB3IP were generated by amplifying genomic DNA containing the two circRNA-generating exons, the intron in between them and ~ 500 and ~ 100 bp of the upstream and downstream flanking introns, respectively. This PCR amplified fragment was inserted into pcDNA3 and piggyBAC (pB) through restriction digest and ligation. To facilitate efficient circRNA production, an artificial inverted repeat was generating around the circRNA exons by PCR amplifying ~ 360 bp of the upstream flanking intron and inserting it in an inverted orientation downstream of the circRNA exons through restriction digest (CutSmart, NEB) and ligation (T4 DNA ligase, NEB) according to manufacturer's instructions. For agarose northern blot and RNase R treatment of the pcDNA3 overexpression vectors, five μ g RNA was digested with one U RNase R (Epicentre) per μ g in a total reaction volume of 10 μ l for ten minutes at 37 $^{\circ}$ C. After this, the sample was prepared for agarose northern blotting by addition of 20 μ l loading buffer. 20 μ l of the loading buffer was added to 5 μ g RNase R treated or control untreated purified RNA. The samples were then loaded onto a 1 % agarose gel containing 3 % formaldehyde and 1 x MOPS and run at 75 V in 1 x MOPS for approximately 3 hours after which RNA was transferred to a Hybond N+ membrane (GE Healthcare) overnight. The RNA was then UV cross-linked to the membrane and pre-hybridized in Church buffer (0.158 M NaH₂PO₄, 0.342 M Na₂HPO₄, 7 % SDS, 1 mM EDTA, 0.5 % BSA, pH 7.5) for one hour. The membrane was probed with a 5' radioactively labelled DNA oligonucleotide at 55 $^{\circ}$ C (60-mer probe) overnight and washed twice in 2 x SSC, 0.1 % SDS for 10 minutes at 45 $^{\circ}$ C before exposure on a phosphoimager screen for data collection. Cell proliferation of pB transfected HEK293 cells was monitored by MTT assay. In brief, $\sim 2,000$ cells are seeded in 96-well format plates in triplicates. On the next day, the medium is replaced with 100 μ l fresh medium plus 10 μ l of a 12 mM MTT stock solution (Invitrogen), and cells are incubated at 37 $^{\circ}$ C for 3 h. Subsequently, all but 25 μ l of the medium is removed from the wells, and formazan dissolved in 50 μ l DMSO, mixed thoroughly and incubated at 37 $^{\circ}$ C for 10 min.

Samples are then mixed again and absorbance read at 530 nm. Measurements are taken at 24, 48, 72 and 96 h post-seeding and background was subtracted.

RNA antisense purification coupled with mass spectrometry (RAP-MS)

Approximately 400×10^6 HEK293 cells were transiently transfected with either circCAMSAP-pcDNA3 or circRAB3IP-pcDNA3 vectors through CaCl₂ transfection protocol. Briefly, cells were transfected 48h prior to experiment with transfection mix containing plasmid, 2.5M CaCl₂ and HEBS buffer (50 mM HEPES, 280 mM NaCl, 1.5 mM Na₂HPO₄, pH 7.05 adjusted with NaOH). Additional sample of HEK293 cells was transfected with control GFP-pcDNA3 to monitor the transfection efficiency. RAP-MS was performed as previously described (McHugh & Guttman, 2018) with 55-mer sense and antisense biotinylated probes designed for circCAMSAP1 and circRAB3IP. The cells were UV crosslinked at 254 nm (UV stratalinker 1800; 0.8 J/cm²), scraped in ice cold PBS, collected by centrifugation, and counted to ensure split into 200×10^6 cells per capture. Cell lysis was performed for 10 minutes on ice in Total cell lysis buffer (10 mM Tris-HCl pH 7.5, 500 mM LiCl, 0.5% dodecyl maltoside (DDM, Sigma), 0.2% sodium dodecyl sulfate (SDS, Sigma), 0.1% sodium deoxycholate (Sigma), 1x cOmplete Protease Inhibitor Cocktail (Roche) and 1000 U of RiboLock RNase Inhibitor (Thermo Fisher)). During cell lysis, samples were passed through 26-G needle 5 times and subsequently sonicated (Bioruptor microtip sonicator, 30% amplitude) for 15 cycles with 5 sec on / 10 sec off settings. Following the sonication, lysates were Dnase I treated (Worthington) with 30 U per vial for 20 minutes at 37°C and pre-cleared for 30 min at 52°C with 1.2 ml per capture of streptavidin Roti MagBeads beads (Carl Roth), which were previously washed in 10 mM Tris-HCl pH 7.5 and Hybridization buffer (10 mM Tris-HCl pH 7.5, 5 mM EDTA, 500 mM LiCl, 0.5% DDM, 0.2% SDS, 0.1% sodium deoxycholate, 4 M urea, 2.5 mM TCEP). After the pre-clearing step, 2 nmol of each probe was denatured at 85°C, quenched on ice, added to the lysates, and incubated at 52°C for 4 hours with intermittent mixing (12 sec on / 9 sec off, 900 rpm). For the probe capture, new set of streptavidin beads was prepared in the same way and added to the samples for another 30 min at 52°C. Beads were then washed 6 times with Hybridization buffer, with each wash being incubated for 5 min at 52°C with intermittent mixing. After the final wash, beads were resuspended in Benzonase elution buffer (20 mM Tris-HCl pH 8.0, 0.05% NLS, 2 mM MgCl₂ and 0.5 mM TCEP) and hybridized probes are eluted from the beads with 120 U of Benzonase (Sigma) for 2 hours at 37°C with intermittent mixing. Supernatant was collected and treated with 5 mM DTT for 30 min at 55°C and subsequently with 40 mM CAA for 30 min at room temperature, centrifuged for 10 min at 20,000 g, transferred to a new tube and stored at -20 °C until the final preparation for mass spectrometry analysis. The final sample preparation and mass spectrometry analysis was carried out at the Proteomics Core

Facility in Cologne, Germany. To assess the efficiency of the RNA pull-down, input and eluted RNA samples were collected during the experiment (as described in McHugh & Guttman, 2018), resuspended in NLS elution buffer (20 mM Tris-HCl pH 8.0, 10 mM EDTA, 2% NLS, 2.5 mM TCEP) and treated with 1 mg/μl of Proteinase K for 1 hour at 55°C. RNA was isolated by phenol:chloroform extraction and precipitated with ethanol. Reverse transcription was carried out using the SuperScript™ II Reverse Transcriptase according to manufacturer's instructions (Invitrogen) and qPCR reactions using the qPCRBIO SyGreen Mix Separate-ROX (NIPPON).

RNA immunoprecipitation coupled with qPCR (RIP-qPCR) experiments

Normal HEK293 cells, grown to confluence in complete media were harvested by gentle scraping in PBS (Sigma). Cells were counted to ensure $\sim 20 \times 10^6$ cells per IP, pelleted by centrifugation and lysed in ice-cold Polysome lysis buffer (100 mM KCl, 5 mM MgCl₂, 10 mM HEPES-NaOH pH 7.0, 1 mM DTT, 200 U/ml RNaseIn, 1x PIC, 0.5% NP-40) for 30 min. Cell lysis was facilitated by passing the lysates through 26-gauge needle 10 times, during the incubation on ice, and 2 x 3 cycles of sonication (30 sec on/30 sec off, low input) on a Bioruptor Pico sonicator (Diagenode). Next, lysates were treated with 40U of DNase I per IP sample (Worthington, LS006343) for 30 min and pelleted by centrifugation to remove the incompletely lysed cells. Input (5%) samples were taken from the supernatants and stored in TRIzol (Invitrogen) until RNA extraction. The rest of the supernatants were subjected to overnight immunoprecipitation with 10 μg of IgG (Milipore, 12-371B; 1 μg/μl), RPS3 (Santa Cruz, sc-376098; 0.2 μg/μl) or SF3B1 (Santa Cruz, sc-514655; 0.2 μg/μl) antibody. Next day, 30 μl of protein-G Dynabeads (Invitrogen) were pre-washed with 1 ml of NT-2 buffer (1x buffer contains 250 mM Tris-HCl pH 7.4, 750 mM NaCl, 5 mM MgCl₂, 0.25% NP-40) and incubated 1 hour with 10 μg/IP of Mouse bridging antibody (Active Motif; 1 μg/μl) at 4°C with end-to-end rotation. Next, beads were washed 3 times in NT-2 buffer, added to the lysates and incubated for 4h at 4°C with end-to-end rotation. Lysates were then washed 6 times in 1 ml of NT-2 buffer for 3 min/wash on end-to-end rotor to ensure removal of unbound material. After the last wash, beads were resuspended in TRIzol (Invitrogen) and RNA was extracted with Direct-zol RNA Microprep kit (Zymo). Reverse transcription was carried out using the SuperScript™ II Reverse Transcriptase according to manufacturer's instructions (Invitrogen) and qPCR reactions using the qPCRBIO SyGreen Mix Separate-ROX (NIPPON).

Extracellular vesicle purification and transfer experiments

Normal HUVEC and circRNA-piggyBAC transfected HEK293 cells were grown in exosome depleted media (normal FBS (Gibco) in complete media replaced with Exosome-Depleted FBS, Life Technologies). HEK293

cells were additionally induced with 3 $\mu\text{g/ml}$ of doxycycline (DOX, Sigma) and 400 $\mu\text{g/ml}$ of G418 (Sigma). After 48 h, cells were collected in TRIzol for qPCR analysis of overexpression levels, while the medium was collected on ice and EVs were purified through series of centrifugation steps. Namely, first round of centrifugation at 750 $\times g$ for 10 min was performed to remove leftover cells from the medium. Supernatants were then collected and spun at 1500 $\times g$ for 10 min, followed by 14000 $\times g$ for 35 min and 100000 $\times g$ for 2 h. Pellets were washed in PBS (Sigma) and subjected to another centrifugation at 100000 $\times g$ for 1h. Pellets representing exosome fraction were resuspended in 200 μl of PBS and kept in -80°C until use. For assessment of circRNA enrichment in HUVEC cells, exosomes were used for RNA extraction (Total Exosome RNA & Protein Isolation Kit, Invitrogen), according to manufacturer's instructions and qPCR analysis (qPCRBIO SyGreen Mix Separate-ROX, NIPPON). For HEK293 exosome transfer onto HUVEC cells, HUVECs were grown in complete media and supplemented with purified exosomes for 24h. HUVECs were then TNF α treated with 10 ng/ml for 1h, washed in PBS 2 times and collected in TRIzol. Extracted RNA (Direct-zol RNA Microprep kit, Zymo) was analyzed by RT-qPCR (SuperScript™ II Reverse Transcriptase, Invitrogen, and qPCRBIO SyGreen Mix Separate-ROX, NIPPON).

Statistical tests

Assessment of correlation between RNA-seq datasets was performed with Pearson correlation core function in R (<https://www.r-project.org/>). T-test and non-linear regression fits analysis of qPCR data was performed with Prism GraphPad (<https://www.graphpad.com/scientific-software/prism/>).

Author contributions

NJ, KKE, AZ, AD performed experiments; NJ performed computational analyses; AP and AZ conceived the study; NJ and AP wrote the manuscript.

Appendix – Supplementary Figures 1-4

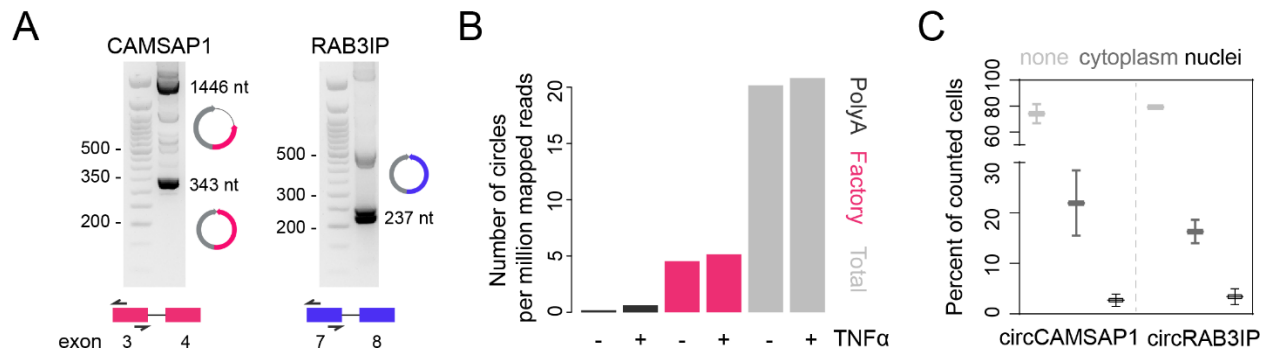


Figure 3.S1. circRNA verification

(A) Agarose gels showing circCAMSAP1 (*left*) and circRAB3IP (*right*) RNA amplification by PCR using divergent primer set (*bottom*). Relevant isoforms are indicated with a cartoon. (B) Bar plot comparing number of circRNAs detected in PolyA (*black*), Factory (*pink*), and Total (*gray*) RNA fractions, ± TNFα. (C) Box plot with whiskers showing percent of counted cells without focal accumulation of either circRNA (*light gray*), one cytoplasmic (*dark gray*) or nuclear focus (*black*).

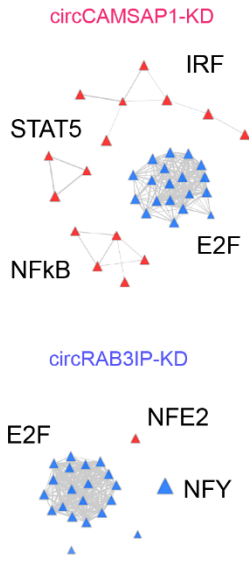
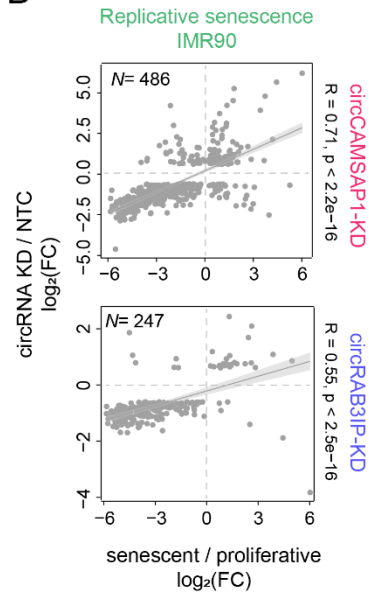
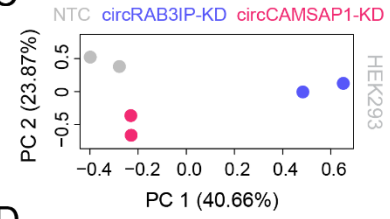
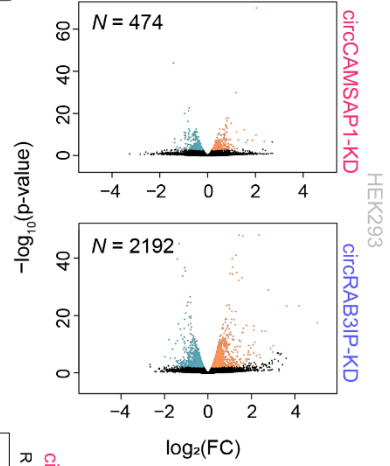
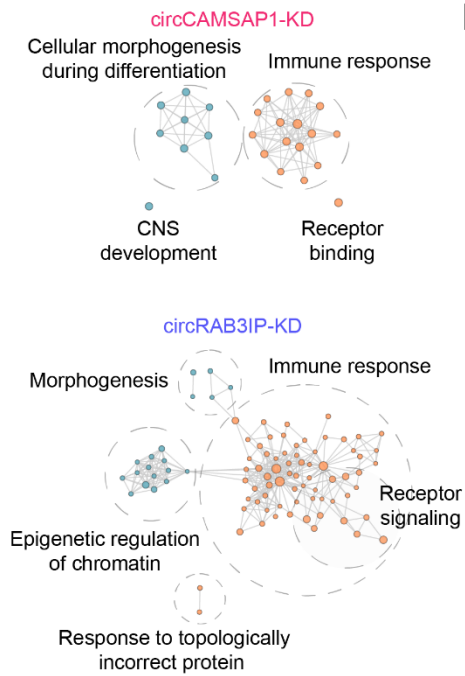
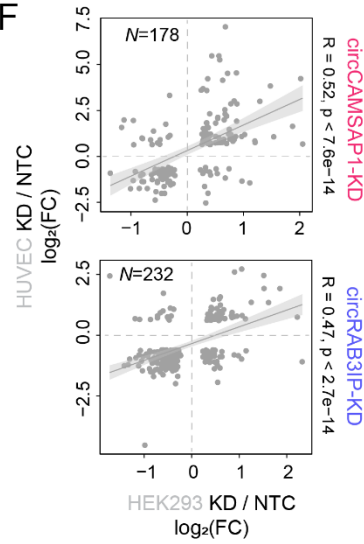
A**B****C****D****E****F**

Figure 3.S2. Knockdown of circCAMSAP1 or circRAB3IP leads to transcriptional changes in HEK293 cells.

(A) Network plot for transcription factor regulation enrichment in circCAMSAP1 (top) or circRAB3IP knock-down (bottom), inferred from overrepresentation analysis (<https://metascape.org/>). (B) Scatter plot correlating circCAMSAP1- (top) and circRAB3IP-knockdown (bottom) DEG with signature DEG in IMR90 replicative senescence. The Pearson correlation coefficient (R) and associated p-value are calculated for the overlapping DEGs indicated (N). (C) PCA scatter plot of gene expression in circCAMSAP1 and circRAB3IP knock-down in HEK293 cells. (D) Top: Volcano plot showing $\log_2(\text{fold change})$ distribution of upregulated (orange) and downregulated (blue) differentially expressed genes upon circCAMSAP1 knock-down. Statistical significance cut-off set at $-\log_{10}(p\text{-value}) \geq 1.3$. Bottom: same as top but for circRAB3IP knock-down. (E) Gene set enrichment analysis of significant differentially expressed genes in circCAMSAP1 (top) and circRAB3IP (bottom) knock-down in HEK293 (absolute $\log_2(\text{fold change}) \geq 0.6$ and $p_{\text{adj}} \leq 0.05$). Clusters of related terms are joined and labelled by the group dominant gene ontology terms. (F) Scatter plot showing correlation of circCAMSAP1 (top) and circRAB3IP (bottom) knock-down datasets in HEK293 and HUVEC cells. Pearson correlation coefficient (R) and p-value (p) are calculated for the N number of overlapping differentially expressed genes.

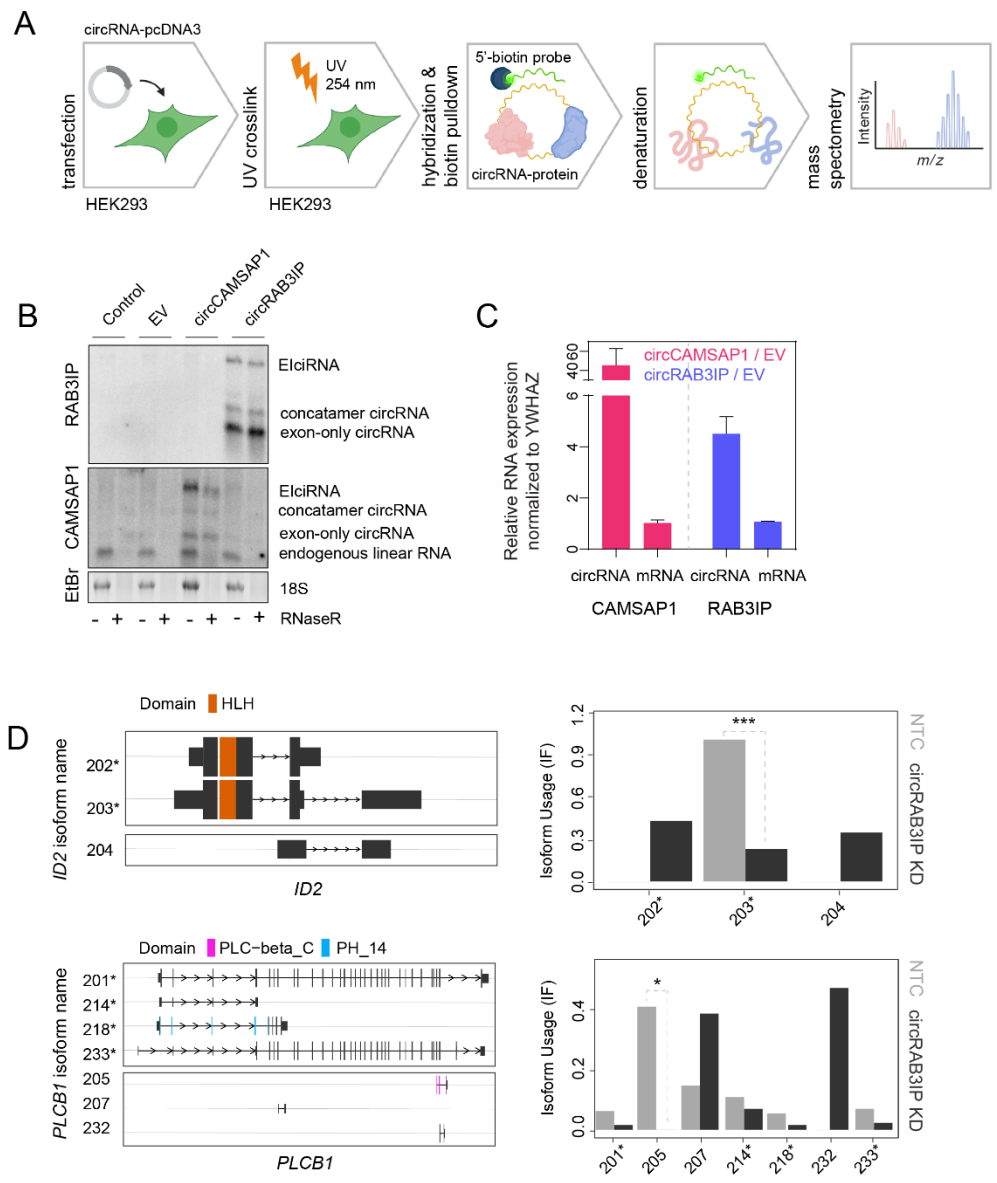


Figure 3.53. RAP-MS and circRAB3IP-knockdown effects on isoform usage.

(A) RNA antisense pull down coupled with mass spectrometry (RAP-MS) strategy overview. Cells are transiently transfected with circRNA-pcDNA3 vectors, prior to the experiment. Transfected cells are UV crosslinked, lysed, and hybridized with biotin labelled probes targeting circRNA backsplicing junction. Pulldown of proteins associated with target circRNAs is followed by the protein denaturation and mass spectrometry. (B) Northern blot for untransfected, empty vector (EV) and pcDNA3-circRNA transfected HEK293 cells, each with or without RNaseR treatment. Ethidium bromide staining of 18S ribosomal RNA was used as a control. (C) Bar plot showing circRNA levels in respective overexpressing HEK293 cells. RNA expression levels are normalized to YWHAZ and presented as fold-change over empty-vector control cells. (D) Left: isoform structure and domain content of selected genes in circRAB3IP knock-down dataset. Isoform names correspond to Ensembl database entries for indicated genes. Right: Bar plots showing changes in isoform usage for the same dataset. Statistical significance of isoform switching is represented by FDR-value ≤ 0.05 (*) and ≤ 0.001 (**).

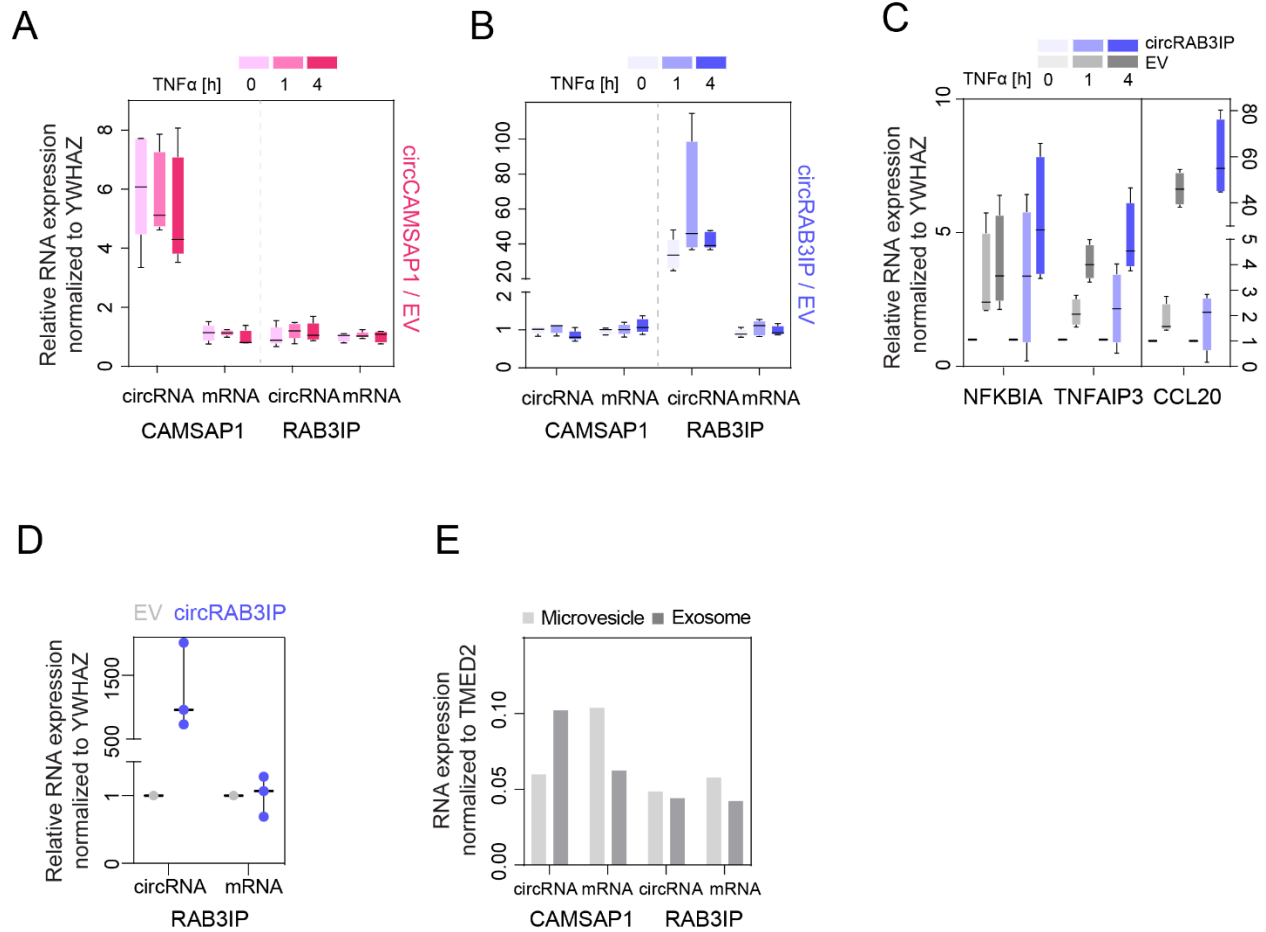


Figure 3.S4. circRNA overexpression and its effects on the inflammatory response.

(A) Box plot showing RNA expression in HEK293 stably overexpressing circCAMSAP1, \pm TNF α treatment. Relative RNA levels of for each TNF α treatment time point are calculated as fold change over non-TNF α treated sample and normalized to *YWHAZ*. (B) Same as panel A, but for circRAB3IP. (C) Box plot showing the response of the *NFKBIA*, *TNFAIP3*, and *CCL20* genes to TNF α for the times indicated in HEK293 stably overexpressing circRAB3IP or empty vector cells (EV). Relative RNA levels at each time point are normalized to *YWHAZ* and represented as fold change over non-treated samples. (D) Box plot with whiskers showing relative overexpression of circRAB3IP in pB-circRNA HEK293 cells (three technical replicates). RNA levels are normalized to *YWHAZ* and represented as a fold change over empty-vector (EV) control samples. (E) Bar plot showing qPCR-based detection of circRNAs in purified extracellular vesicles from HUVEC cells. RNA expression levels were normalized to *TMED2*.

Discussion

Throughout their lifespan, cells undergo continuous cell-cycling accompanied by dramatic alterations in their structure and function. Such consecutive transitions from being metabolically active in interphase to inactive in mitosis, only exemplify a fraction of the cells' activities, as cellular homeostasis inherently also includes the continuous accommodation of intra- and extracellular cues. In normal primary cells, cell cycling eventually culminates in replicative exhaustion, cell cycle arrest, and induction of senescence. However, senescence does not typically result in immediate cell clearance, with senescent cells remaining metabolically active, but physiologically and morphologically altered to reflect their changed cellular activity. Cell cycle progression, responses to environmental cues and senescence entry, therefore, represent distinct and dynamic cellular states, which ultimately diverge in their transcriptional output. The cellular transcriptomes in any of these states represent accurate readouts of the collective activity of thousands of molecules, acting in convergent and opposing manners to regulate gene expression. How these interact to ensure timely execution of gene expression seems to depend on cellular context, but for most regulatory cascades, details of these interactions remain elusive. Similarly, how transcription itself affects other genomic features merits further investigation.

RNA Pol II is essential for the reestablishment of loop extrusion after mitosis

We addressed the question about functional implications of active transcription on genomic features by removing one of its key components, RNA Pol II. As our understanding of principles underlying chromatin folding improved tremendously due the development of high-throughput technologies like Hi-C, emerging evidence pointed towards a dynamic interplay of different processes required for the establishment and maintenance of genome organization (Rada-Iglesias *et al*, 2018). Among those, compartmentalization (Rao *et al*, 2014; Bonev *et al*, 2017; Li *et al*, 2015b), cohesin/CTCF loop extrusion (Nora *et al*, 2019; Rao *et al*, 2017) and active transcription (Papantonis & Cook, 2011; Rowley *et al*, 2017) have been implicated in chromatin folding, but mechanistic insight into their reciprocal dependencies is still lacking.

In our study, we utilized an cell model allowing inducible and reversible RNA Pol II degradation (Yesbolatova *et al*, 2019; Nagashima *et al*, 2019). Corroborating previous reports of RNA Pol II degradation in mouse ES cells (Jiang *et al*, 2020), we observed no prominent impact of RNA Pol II loss on genome folding upon short-term (2 h) depletion. However, longer, more persistent depletion (14 h) affected chromatin

accessibility at transcription start sites (TSS), with overall decrease in accessibility occurring concomitantly with deposition of the H3K27me3 and depletion of the H3K27ac histone marks. Interestingly, the decrease in transcriptional output that naturally follows RNA Pol II degradation and the changes in accessibility and histone marks mentioned above, primarily affected chromatin assembly-related genes. This hints towards the existence of a feedback loop controlling the expression of genes organizing chromatin in the absence of robust transcription, which itself requires chromatin organization control.

As regards higher order chromatin structure, the most prominent effect of RNA Pol II depletion in interphase cells was on chromatin looping. We observed an increase in the number and length of loops that more prominent than what was observed previously (Barutcu *et al*, 2019; Hsieh *et al*, 2015; Jiang *et al*, 2020), but in line with the generally presumed role of active transcription in this process (Vermunt *et al*, 2019). Nevertheless, the relative contribution of initiating versus elongating polymerases, of the cells' RNA content or of the cell cycle phase they are in is a challenging task. In fact, asynchronous cell populations consist of cells with different genomic configurations and transcriptional activities (Naumova *et al*, 2013; Vagnarelli, 2013). In addition, we found that degrading chromatin-bound RNA Pol II marked by Ser5 hyperphosphorylation in its C-terminal domain is less efficient than degradation of the diffuse (unbound) or elongating polymerase pools. And it was recently shown that RNA Pol II-Ser5 constrains chromatin to contribute to its compaction (Nagashima *et al*, 2019). We thus reasoned that our study would benefit from degradation of RNA Pol II in synchronized cells, especially those transitioning through mitosis, when no polymerase is bound to chromatin (Zhang *et al*, 2019). In addition, exit from mitosis and entry into G1 are marked by a prominent transcriptional burst (Hsiung *et al*, 2016), which implies a possible contribution of RNA Pol II in setting up chromatin for the ensuing interphase.

To this end, we synchronized cells at the G2/M checkpoint, initiated RNA Pol II degradation and released cells through mitosis and into G1. Using this experimental set up, the changes we observed across all levels of organization (compartmentalization, TADs and loops) were prominent. Focusing on the loss of A and B-compartment boundary strength, it was accompanied by increased interchromosomal interactions at the expense of intrachromosomal ones. As A/B-compartments are defined by histone modifications and reflect transcriptional activity, the observed decrease in H3K27ac accompanied by increased H3K27me3 levels could explain such changes. Such weakened compartmentalization is a typical feature of late mitotic chromatin (Zhang *et al*, 2019), which is remodeled early after mitosis by carryover histone modifications, "bookmarking" TFs (Teves *et al*, 2016; Pelham-Webb *et al*, 2021) and homotypic chromatin interactions (Hsiung *et al*, 2015; Zhang *et al*, 2019). As early post-mitotic events coincide with a transcriptional "spike"

(Hsiung *et al*, 2016), it follows that complete reestablishment of compartments and accurate deposition of histone marks must at least in part be affected by RNA Pol II. Notably, we could rule out topoisomerase contribution to the observed changes in chromatin topology upon RNA Pol II depletion. We used a TOP2B-knockout cell line allowing conditional depletion of TOP2A. Assessment of the nascent transcriptome in cells lacking both TOP2 factors revealed that a minor fraction of genes is affected, and only changes in chromatin organization were detected in Hi-C data, hence the results obtained in RNA Pol II-depleted cells could not be recapitulated by supercoiling-based changes. This is not likely to be due to redundancy of TOP1 and TOP2 activities, as these are disparate in mammalian cells (Wang, 2002).

Quantitative and qualitative changes in chromatin folding, especially following exit from mitosis, suggest that active RNA Pol II acts as both a converging and opposing force to all others that have been implicated to spatial genome organization to date (Rada-Iglesias *et al*, 2018). However, depletion of RNA Pol II from cells reentering the G1 phase of the cell cycle leads to another intriguing effect. Counter to what intuition would suggest, its absence from chromatin results in increased rather than decreased chromatin accessibility. This phenomenon is widespread across chromosomes and is also accompanied by strikingly increased levels of chromatin-embedded TBP, a general transcription factor and known “bookmarking” protein (Teves *et al*, 2016), as well as of aberrant levels of known chromatin remodelers like BRG1. These data suggest that RNA Pol II and transcription are need for resetting a functional chromatin landscape, already at the level of linear organization, despite successful mitotic bookmarking by other factors. An aberrantly set up landscape inevitably affects gene expression and genomic regulation in interphase, and this is already reflected in our documentation of inefficient loading of the key CTCF and cohesin complexes on chromatin lacking RNA Pol II. It then follows that an intimate relationship between the 1D and 3D structures of chromosomes emerges, and proves critical in the M-to-G1 transition especially.

HMGB1 participates in the regulation of senescence-relevant gene expression and mRNA stability

In sharp contrast to proliferating cells like those studied above, senescent cells undergo extensive morphological and functional remodeling in a cell type- and trigger-dependent manner. This points to senescence being rather dynamic and diverse, instead of a single definitive cellular state (Herranz & Gil, 2018). Despite its intrinsic heterogeneity, common features of senescence include permanent cell cycle arrest primarily regulated at the level of transcription (Narita *et al*, 2003; Hernandez-Segura *et al*, 2018), inflammatory signaling through stimulation of related genes and secretion of pro-inflammatory molecules

(Coppé *et al*, 2010; Davalos *et al*, 2013), as well as extensive chromatin remodeling (Narita, 2007; Criscione *et al*, 2016). Although details of how these events are mutually connected are not yet fully understood, it is apparent that they are intimately related to changes in transcriptional output. In our research, we assigned a regulatory role for all the three common features of senescence to the architectural protein HMGB1, which is the most abundant non-histone protein in proliferating cells, but is rapidly evicted from cell nuclei upon senescence entry.

The senescence-triggered release of HMGB1 as part of the SASP is well documented (Davalos *et al*, 2013), but in the events preceding its secretion, HMGB1 is a regulator in the cell nucleus. Similarly, the mechanism of HMGB1 associating with, bending and unwinding of chromatin is well studied (Stros, 2010), but due to its incompatibility with conventional crosslinking approaches, the positioning of HMGB1 on chromatin genome-wide remained elusive. To overcome this, we applied to HMGB1 the tailored approach for crosslinking that previously allowed mapping of HMGB2 binding sites via ChIP-seq (Zirkel *et al*, 2018). This way, we discovered that HMGB1 in proliferating cells binds predominantly to active gene promoters and gene bodies across cell types, but in a cell type-specific manner. Despite this cell type-specific binding profiles, HMGB1-bound genes belong to the same senescence-relevant pathways like cell cycle transition, ECM remodeling and cell adhesion, as well as inflammatory signaling and the SASP. Curiously, most gene promoters occupied by HMGB1 in proliferating cells are up- rather than downregulated upon its loss in senescence. This identifies in HMGB1 a role as a transcriptional rheostat rather than a classical activator.

Given its architectural nature, we also looked how the role of HMGB1 in gene regulation correlates with chromatin reorganization upon senescence entry. Earlier reports on chromatin organization changes in senescent cells described increased chromatin accessibility, reorganization of heterochromatin, and TAD remodeling (Criscione *et al*, 2016; De Cecco *et al*, 2013; Narita, 2007; Chandra *et al*, 2015; Zirkel *et al*, 2018). TAD boundaries are often marked by CTCF binding and/or active gene promoters (Dixon *et al*, 2012; Nora *et al*, 2017) and we recently showed that the sister protein of HMGB1, HMGB2, occupies TAD boundaries. These boundaries are remodeled upon senescence entry and nuclear loss of HMGB2 (Zirkel *et al*, 2018), Therefore we have tested if this also holds true for HMGB1. Perhaps counterintuitively, HMGB1 is enriched at just a small fraction of TAD boundaries, but largely distributed within TADs. It is noteworthy that HMGB1 anti-correlated with CTCF-bound positions at TAD boundaries, pointing away from HMGB1 playing a role in TAD insulation. However, we did find HMGB1 decorating almost 20% of proliferating cohesin/CTCF loop anchors, although >1/4 of those remained unaltered in senescence. These are considered architectural features of chromosomes that span an average size of 285 kbp (Rao *et al*, 2014, 2017). Interestingly, they

also harbor a considerable fraction of senescence-related genes, mainly associated with inflammatory activation and cell cycle arrest. As this was true for both cell types, it asserted the notion of HMGB1 topological contribution to gene regulation upon senescence entry.

Transient siRNA-mediated knockdown of HMGB1 returned ~2000 differentially-expressed genes in HMGB1-depleted cells, linking the loss of HMGB1 to the upregulation of pro-inflammatory signaling and the downregulation of hallmark senescence genes related to chromatin organization, transcriptional silencing, and the p53 pathway (Hernandez-Segura *et al*, 2018). However, correlation of senescence-derived and HMGB1-knockdown transcriptomes was rather moderate. For example, a significant number of genes upregulated upon HMGB1-knockdown, but downregulated in senescence concerned the pathways of RNA metabolism and splicing. This was significant because it indicated that HMGB1 depletion alone cannot reproduce the whole spectrum of senescent-relevant changes (e.g., changes in heterochromatin identity and emergence of senescence-induced CTCF clusters is only seen upon HMGB2 knockdown; Zirkel *et al*, 2018). It also revealed that RNA processing might also rely on HMGB1, and thus be altered in senescence as a result of its nuclear loss.

This last finding prompted us to catalogue HMGB1 protein partners through co-IP coupled with mass spectrometry. This revealed that a significant portion of HMGB1 protein partners are, in fact, RNA binding proteins (RBPs). Through sCLIP-seq we have detected more than 800 mRNAs being directly bound by HMGB1, some of which are also regulated in senescence, again with a strong tendency towards being up- rather than downregulated. Moreover, > 25% of HMGB1-bound mRNAs were affected as regards splicing in senescence and/or HMGB1 knockdown. Intriguingly, some of the RBP partners of HMGB1 are also HMGB1 targets in sCLIP data, and had further been linked to senescence and SASP regulation (Tominaga-Yamanaka *et al*, 2012; Wu *et al*, 2015; Shao *et al*, 2016; Chen *et al*, 2017). This way, we exemplify a novel regulatory role of HMGB1 on the path to senescence, where its nuclear eviction also affects the cells' transcriptional program via altered RNA processing.

In summary, unveiling the binding preferences of HMGB1 on chromatin, its relative contribution to the senescence gene expression program, and demonstrating its RNA-binding and regulatory potential, added important pieces of the puzzle as regards the multifaceted role of this protein in homeostasis and cellular ageing. Thus, we bridge two different aspects of HMGB1 activity in senescence: one where release of nuclear HMGB1 triggers a cascade of events within senescent nuclei, and the other where HMGB1 acts as a part of the pro-inflammatory SASP in extracellular space. This represents a unique regulatory example

of functional crosstalk between genome organization and paracrine cell-to-cell communication allowing for the switching of states in a cell population.

circRNAs contribute to cellular homeostasis through RNA-protein interactions

As somatic cells display finite proliferative capacity, replicative senescence represents a physiological mechanism normally triggered close to the end of their lifespan. However, senescence can also be induced under conditions of elevated stress, when homeostasis is disrupted, as a mechanism facilitating either cell survival or death. Regulatory roles in senescent events have been attributed to various proteins, as discussed above for the architectural protein HMGB1, but less often to non-coding RNAs. On the other hand, regulatory roles for non-coding RNAs have generally been investigated in cancer and other pathological conditions, but rarely in normal primary cells. Additionally, while more than ten thousand circRNAs are registered in human cells, functional validations for their vast majority are still missing. Here, we dissect the functionality of circRAB3IP, while also describing a novel role for the well-studied circCAMSAP1. We assign these circRNAs with regulatory roles in homeostasis maintenance of primary cells, whereby their loss invokes a senescence-like response.

The initial key finding for this study was the profound transcriptional changes induced in primary human endothelial cells upon siRNA-mediated knockdown of either circRNA. For comparison, the combined elimination of TOP2A/B led to significant expression changes of < 400 genes. For either circRNA, this number was significantly higher, despite the prevalence and involvement of TOP2A/B in transcription. To add to this, the genes affected by the circRNA depletion were functionally enriched for a very defined set of cellular pathways. A shared feature was the induction of proliferation arrest, with a gene expression signature correlating well with that of senescence (Hernandez-Segura *et al*, 2018). By contrast, we earlier observed little such correlation upon knockdown of HMGB1, a protein with a central role in senescence. Additionally, overexpression of either circRNA markedly increased the proliferative potential of stably-transfected cells. However, the rest of the pathways affected were disparate, arguing for some functional specificity of each circRNA.

As circRNAs have been recently implicated in splicing regulation (Ashwal-Fluss *et al*, 2014; Li *et al*, 2015d), we analyzed genes displaying a shift in isoform usage in our KD data, only to find circRAB3IP- but not circCAMSAP1-knockdown was followed by an extensive change in isoform usage for a large set of genes. Reassuringly, many of the genes displaying such changes were associated with cell cycle progression and

displayed a tendency for expressing shorter, often non-coding or devoid of functional domains isoforms. As circRAB3IP is not predicted to act as a miRNA “sponge”, we catalogued its protein partners using RAP-MS. Of a handful for robustly-detected proteins, we focused on SF3B1 due to its established roles in mRNA splicing (Sun, 2020), senescence (Yin *et al*, 2019) and ageing (Holly *et al*, 2014). We showed that more than 40% of differentially-expressed genes and 40% of isoform-switching genes in circRAB3IP-knockdown data were also regulated by SF3B1, and a subset of these regulates cell cycle progression. However, the correlation between genes differentially expressed upon SF3B1- and circRAB3IP-knockdown was poor. This suggests that circRAB3IP interaction with SF3B1 could specifically only facilitate splicing of cell cycle-related genes, and its disruption upon circRAB3IP depletion invokes proliferation arrest. Thus, it is tempting to speculate that circRAB3IP serves to guide SF3B1 to those targets relevant for cell cycle progression.

We were not able to implicate circCAMSAP1 in alternative splicing based on knockdown data, and this was by and large also reflected on its putative protein partners as catalogued by RAP-MS. However, its larger interactome (in comparison to that of circRAB3IP) included a number of factors participating in inflammatory and RNA metabolism pathways. Among those, RPS3 has been implicated in cell cycle progression (Jang *et al*, 2012), E2F1 co-activation (Lee *et al*, 2010), and inflammation via NF- κ B (Wan *et al*, 2007) and the NF- κ B inhibitor, I κ B α (Stanborough *et al*, 2014). Moreover, > 20% of mRNAs bound by RPS3 are also regulated upon circCAMSAP1-knockdown and included in the senescence hallmark signature. We thus speculated that the circCAMSAP1-RPS3 interaction mediates inflammation in primary cells. Depletion of circCAMSAP1 leads to strong inflammatory gene expression, thus its overexpression (and possibly over-association with RPS3 to sequester) might act to constrain it. Indeed, circCAMSAP1 overexpression suffices for the dampening the response of two prototypic immediate-early genes, *TNFAIP3* and *CCL20*, following acute TNF α stimulation. In addition, and following reports on circRNAs acting in a paracrine manner via exosome transport (Li *et al*, 2015c; Dou *et al*, 2016), we treated endothelial cells with exosomes collected from circCAMSAP1-overexpressing HEK293 cells to find that, again, stimulation of the immediate-early *NFKBIA* gene was suppressed in response to acute TNF α stimulation. Nevertheless, a higher throughput approach should follow such experiments in order to obtain a clearer picture of the selectivity of inflammatory modulation by circCAMSAP1. In summary, both circRAB3IP and circCAMSAP1 seem to be essential for normal cellular proliferation and depletion of either results in rapid cell-cycle arrest, and in case of CAMSAP1 also in a pro-inflammatory phenotype.

Outlook

In our effort to tackle seemingly unrelated biological questions, we have stumbled upon a prominent connection between them. The contribution of RNA molecules seems to represent a common identifier between individual roles of examined proteins and respective cellular processes. This finding agrees with emerging view of “pan-genomic” regulatory networks (Brackley *et al*, 2020), reinforcing the idea of multivalent small-scale interactions governing the large-scale cellular processes.

Investigation of RNA Pol II contribution to chromatin refolding via RNA Pol II depletion furthered our understanding of chromatin dynamics following exit from mitosis. However, in this experimental setup delineating the indirect effects of transcriptional attenuation remains an issue. This is mostly so because we utilize a 14-hour-long RNA Pol II depletion and RNA species carrying potential regulatory roles have relatively short half-lives. Thus, we cannot rule out an additive effect of RNA Pol II depletion and RNA degradation. This is evident in the recently reported RNA-binding capacity of CTCF, which seems to be required for its robust binding to a subset of chromatin sites and loss of such RNA-protein interactions promptly invokes the collapse of chromatin loops (Saldaña-Meyer *et al*, 2019; Hansen *et al*, 2019; Kuang & Wang, 2020).

Regarding HMGB1, we corroborated previous research showing that depletion of HMGB1 suffices for senescence induction, however we did not characterize the long-term effect of such depletion and whether senescence markers persist in the continuous absence of HMGB1. It is also worth noting that while HMGB1 depletion in senescence is gradual and tightly connected to paracrine HMGB1, knock-down experiments invoke rapid depletion of both intra- and extracellular HMGB1 activities that might trigger stress mechanisms which would normally not be detected in senescence. The gradual dynamic of HMGB1 release from chromatin would allow interrogation of HMGB1 binding profiles at multiple timepoints on the path to senescence, and thus possible reshuffling of remaining HMGB1 on chromatin and its regulatory effects.

Until now, circRNAs are reported to generally act as protein and/or microRNA scaffolds, yet they are often short in length (Memczak *et al*, 2013), take up various secondary structures (Liu *et al*, 2019) and often don't carry multiple microRNA binding motifs (Guo *et al*, 2014). These circRNA features are hard to reconcile with experimental evidence of protein/miRNA interactions with functional implications. However, investigation of non-coding RNAs as potent factors driving the phase-separation in cells recently gained traction (Luo *et al*, 2021a). lncRNAs, for example, are reported to regulate various cellular processes by

means of phase-separation, such as X-chromosome silencing (Cerase *et al*, 2019), gene expression (Daneshvar *et al*, 2020) and cell proliferation (Elguindy & Mendell, 2021). Here, we identified in RNA binding proteins SF3B1 and RPS3 two factors putatively sequestered by circRAB3IP and circCAMSAP1, respectively. Although mechanistic details of these interactions remain to be illuminated, a similar “phase separation” mode of action for these circRNAs might explain how they exert their regulatory roles and how removal of a single circRNA can have such profound effects on cell homeostasis.

Literature

- Abdelmohsen K, Panda AC, Munk R, Grammatikakis I, Dudekula DB, De S, Kim J, Noh JH, Kim KM, Martindale JL, *et al* (2017) Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol* 14: 361–369
- Abramo K, Valton AL, Venev SV, Ozadam H, Fox AN & Dekker J (2019) A chromosome folding intermediate at the condensin-to-cohesin transition during telophase. *Nat Cell Biol* 21: 1393–1402
- Achar YJ, Adhil M, Choudhary R, Gilbert N & Foiani M (2020) Negative supercoil at gene boundaries modulates gene topology. *Nature* 577: 701–705
- Acosta JC, Banito A, Wuestefeld T, Georgilis A, Janich P, Morton JP, Athineos D, Kang T-W, Lasitschka F, Andrulis M, *et al* (2013) A complex secretory program orchestrated by the inflammasome controls paracrine senescence. *Nat Cell Biol* 15: 978–990
- Aktaş T, Avşar İlik İ, Maticzka D, Bhardwaj V, Pessoa Rodrigues C, Mittler G, Manke T, Backofen R & Akhtar A (2017) DHX9 suppresses RNA processing defects originating from the Alu invasion of the human genome. *Nature* 544: 115–119
- Alexandrova EA & Beltchev BG (1988) Acetylated HMG1 protein interacts specifically with homologous DNA polymerase alpha in vitro. *Biochem Biophys Res Commun* 154: 918–927
- Amat R, Böttcher R, Dily FL, Vidal E, Quilez J, Cuartero Y, Beato M, Nadal E & Posas F (2019) Rapid reversible changes in compartments and local chromatin organization revealed by hyperosmotic shock. *Genome Res* 29: 18–28
- Arrigoni L, Richter AS, Betancourt E, Bruder K, Diehl S, Manke T & Bonisch U (2016) Standardizing chromatin research: a simple and universal method for ChIP-seq. *Nucleic Acids Res* 44: 67
- Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N & Kadener S (2014) circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* 56: 55–66
- Azam S, Hou S, Zhu B, Wang W, Hao T, Bu X, Khan M & Lei H (2019) Nuclear retention element recruits U1 snRNP components to restrain spliced lncRNAs in the nucleus. *RNA Biol* 16: 1001–1009
- Baker DJ, Childs BG, Durik M, Wijers ME, Sieben CJ, Zhong J, Saltness R, Jeganathan KB, Versoza GC, Pezeshki A-M, *et al* (2016) Naturally occurring p16Ink4a-positive cells shorten healthy lifespan. *Nature* 530: 184–189
- Banani SF, Lee HO, Hyman AA & Rosen MK (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* 18: 285–298
- Banigan EJ, van den Berg AA, Brandão HB, Marko JF & Mirny LA (2020) Chromosome organization by one-sided and two-sided loop extrusion. *eLife* 9: e53558
- Baranello L, Wojtowicz D, Cui K, Devaiah BN, Chung HJ, Chan-Salis KY, Guha R, Wilson K, Zhang X, Zhang H, *et al* (2016) RNA polymerase II regulates topoisomerase 1 activity to favor efficient transcription. *Cell* 165: 357–371

- Bartsch D, Zirkel A & Kurian L (2018) Characterization of Circular RNAs (circRNA) Associated with the Translation Machinery. In *Circular RNAs*, Dieterich C & Papantonis A (eds) pp 159–166. New York, NY: Springer New York
- Barutcu AR, Blencowe BJ & Rinn JL (2019) Differential contribution of steady-state RNA and active transcription in chromatin organization. *EMBO reports* 20: e48068
- Bauer EM, Shapiro R, Billiar TR & Bauer PM (2013) High mobility group Box 1 inhibits human pulmonary artery endothelial cell migration via a Toll-like receptor 4- and interferon response factor 3-dependent mechanism(s). *J Biol Chem* 288: 1365–1373
- Beagan JA, Duong MT, Titus KR, Zhou L, Cao Z, Ma J, Lachanski CV, Gillis DR & Phillips-Cremens JE (2017) YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Res* 27: 1139–1152
- Beagan JA & Phillips-Cremens JE (2020) On the existence and functionality of topologically associating domains. *Nat Genet* 52: 8–16
- Bediaga NG, Coughlan HD, Johanson TM, Garnham AL, Naselli G, Schröder J, Fearnley LG, Bandala-Sanchez E, Allan RS, Smyth GK, *et al* (2021) Multi-level remodelling of chromatin underlying activation of human T cells. *Scientific Reports* 11: 528
- Bianchi ME, Crippa MP, Manfredi AA, Mezzapelle R, Rovere Querini P & Venereau E (2017) High-mobility group box 1 protein orchestrates responses to tissue damage via inflammation, innate and adaptive immunity, and tissue repair. *Immunol Rev* 280: 74–82
- Bianco S, Lupiáñez DG, Chiariello AM, Annunziatella C, Kraft K, Schöpflin R, Wittler L, Andrey G, Vingron M, Pombo A, *et al* (2018) Polymer physics predicts the effects of structural variants on chromatin architecture. *Nature Genetics* 50: 662–667
- Bintu B, Mateo LJ, Su J-H, Sinnott-Armstrong NA, Parker M, Kinrot S, Yamaya K, Boettiger AN & Zhuang X (2018) Super-resolution chromatin tracing reveals domains and cooperative interactions in single cells. *Science* 362
- Birger Y, West KL, Postnikov YV, Lim J-H, Furusawa T, Wagner JP, Laufer CS, Kraemer KH & Bustin M (2003) Chromosomal protein HMG1 enhances the rate of DNA repair in chromatin. *EMBO J* 22: 1665–1675
- Blair RH, Horn AE, Pazhani Y, Grado L, Goodrich JA & Kugel JF (2016) The HMGB1 C-terminal tail regulates DNA bending. *J Mol Biol* 428: 4060–4072
- Boehning M, Dugast-Darzacq C, Rankovic M, Hansen AS, Yu T, Marie-Nelly H, McSwiggen DT, Kocic G, Dailey GM, Cramer P, *et al* (2018) RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat Struct Mol Biol* 25: 833–840
- Boeynaems S, Alberti S, Fawzi NL, Mittag T, Polymenidou M, Rousseau F, Schymkowitz J, Shorter J, Wolozin B, Van Den Bosch L, *et al* (2018) Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol* 28: 420–435

Boija A, Klein IA, Sabari BR, Dall’Agnese A, Coffey EL, Zamudio AV, Li CH, Shrinivas K, Manteiga JC, Hannett NM, *et al* (2018) Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* 175: 1842-1855.e16

Bonaldi T, Längst G, Strohner R, Becker PB & Bianchi ME (2002) The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. *EMBO J* 21: 6865–6873

Bonaldi T, Talamo F, Scaffidi P, Ferrera D, Porto A, Bachi A, Rubartelli A, Agresti A & Bianchi ME (2003) Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. *EMBO J* 22: 5551–5560

Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot J-P, Tanay A, *et al* (2017) Multiscale 3D Genome Rewiring during Mouse Neural Development. *Cell* 171: 557-572.e24

Bonne C, Duguet M & de Recondo AM (1979) Rat liver DNA binding proteins: physiological variations. *FEBS Lett* 106: 292–296

Bonne-Andrea C, Harper F, Puvion E, Delpech M & De Recondo AM (1986) Nuclear accumulation of HMG1 protein is correlated to DNA synthesis. *Biol Cell* 58: 185–194

Bouwman BA & de Laat W (2015) Getting the genome in shape: the formation of loops, domains and compartments. *Genome Biology* 16: 154

Brackley CA, Gilbert N, Michieletto D, Papantonis A, Pereira MCF, Cook PR & Marenduzzo D (2020) Complex small-world regulatory networks emerge from the 3D organisation of the human genome. *bioRxiv*: 2020.05.12.091041

Brackley CA, Johnson J, Kelly S, Cook PR & Marenduzzo D (2016) Simulated binding of transcription factors to active and inactive regions folds human chromosomes into loops, rosettes and topological domains. *Nucleic Acids Res* 44: 3503–3512

Braicu C, Zimta A-A, Gulei D, Olariu A & Berindan-Neagoe I (2019) Comprehensive analysis of circular RNAs in pathological states: biogenesis, cellular regulation, and therapeutic relevance. *Cell Mol Life Sci* 76: 1559–1577

Brandão HB, Paul P, Berg AA, Rudner DZ, Wang X & Mirny LA (2019) RNA polymerases as moving barriers to condensin loop extrusion. *Proc Natl Acad Sci USA* 116: 20489–20499

Brangwynne CP, Mitchison TJ & Hyman AA (2011) Active liquid-like behavior of nucleoli determines their size and shape in *Xenopus laevis* oocytes. *Proc Natl Acad Sci U S A* 108: 4334–4339

Brant L, Georgomanolis T, Nikolic M, Brackley CA, Kolovos P, Ijcken W, Grosveld FG, Marenduzzo D & Papantonis A (2016) Exploiting native forces to capture chromosome conformation in mammalian cell nuclei. *Mol Syst Biol* 12: 891

Buckle A, Brackley CA, Boyle S, Marenduzzo D & Gilbert N (2018) Polymer simulations of heteromorphic chromatin predict the 3D folding of complex genomic loci. *Mol Cell* 72: 786–797

Buenrostro J, Wu B, Chang H & Greenleaf W (2015) ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Curr Protoc Mol Biol* 109: 21.29.1-21.29.9

Burd CE, Sorrentino JA, Clark KS, Darr DB, Krishnamurthy J, Deal AM, Bardeesy N, Castrillon DH, Beach DH & Sharpless NE (2013) Monitoring tumorigenesis and senescence in vivo with a p16(INK4a)-luciferase model. *Cell* 152: 340–351

Buschbeck M & Hake SB (2017) Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat Rev Mol Cell Biol* 18: 299–314

Busslinger GA, Stocsits RR, van der Lelij P, Axelsson E, Tedeschi A, Galjart N & Peters J-M (2017) Cohesin is positioned in mammalian genomes by transcription, CTCF and Wapl. *Nature* 544: 503–507

Bustin M (2001) Revised nomenclature for high mobility group (HMG) chromosomal proteins. *Trends Biochem Sci* 26: 152–153

Cai D, Feliciano D, Dong P, Flores E, Gruebele M, Porat-Shliom N, Sukenik S, Liu Z & Lippincott-Schwartz J (2019) Phase separation of YAP reorganizes genome topology for long-term YAP target gene expression. *Nat Cell Biol* 21: 1578–1589

Calandrelli R, Xu L, Luo Y, Wu W, Fan X, Nguyen T, Chen C-J, Sriram K, Tang X, Burns AB, *et al* (2020) Stress-induced RNA–chromatin interactions promote endothelial dysfunction. *Nature Communications* 11: 5211

Calcinotto A, Kohli J, Zagato E, Pellegrini L, Demaria M & Alimonti A (2019) Cellular Senescence: Aging, Cancer, and Injury. *Physiol Rev* 99: 1047–1078

Calogero S, Grassi F, Aguzzi A, Voigtländer T, Ferrier P, Ferrari S & Bianchi ME (1999) The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. *Nat Genet* 22: 276–280

Canat A, Veillet A, Bonnet A & Therizols P (2020) Genome anchoring to nuclear landmarks drives functional compartmentalization of the nuclear space. *Brief Funct Genomics* 19: 101–110

Carlevaro-Fita J, Rahim A, Guigó R, Vardy LA & Johnson R (2016) Cytoplasmic long noncoding RNAs are frequently bound to and degraded at ribosomes in human cells. *RNA* 22: 867–882

Casa V, Moronta Gines M, Gade Gusmao E, Slotman JA, Zirkel A, Josipovic N, Oole E, van IJcken WFJ, Houtsmuller AB, Papantonis A, *et al* (2020) Redundant and specific roles of cohesin STAG subunits in chromatin looping and transcriptional control. *Genome Res* 30: 515–527

Castello A, Fischer B, Frese CK, Horos R, Alleaume A-M, Foehr S, Curk T, Krijgsveld J & Hentze MW (2016) Comprehensive Identification of RNA-Binding Domains in Human Cells. *Mol Cell* 63: 696–710

Catena R, Escoffier E, Caron C, Khochbin S, Martianov I & Davidson I (2009) HMGB4, a novel member of the HMGB family, is preferentially expressed in the mouse testis and localizes to the basal pole of elongating spermatids. *Biol Reprod* 80: 358–366

Cato L, Stott K, Watson M & Thomas JO (2008) The interaction of HMGB1 and linker histones occurs through their acidic and basic tails. *J Mol Biol* 384: 1262–1272

Caudron-Herger M, Cook PR, Rippe K & Papantonis A (2015) Dissecting the nascent human transcriptome by analysing the RNA content of transcription factories. *Nucleic Acids Res* 43: e95

- Cekan P, Hasegawa K, Pan Y, Tubman E, Odde D, Chen J-Q, Herrmann MA, Kumar S & Kalab P (2016) RCC1-dependent activation of Ran accelerates cell cycle and DNA repair, inhibiting DNA damage-induced cell senescence. *Mol Biol Cell* 27: 1346–1357
- Celona B, Weiner A, Di Felice F, Mancuso FM, Cesarini E, Rossi RL, Gregory L, Baban D, Rossetti G, Grianti P, *et al* (2011) Substantial histone reduction modulates genomewide nucleosomal occupancy and global transcriptional output. *PLoS Biol* 9: e1001086
- Cerase A, Armaos A, Neumayer C, Avner P, Guttman M & Tartaglia GG (2019) Phase separation drives X-chromosome inactivation: a hypothesis. *Nat Struct Mol Biol* 26: 331–334
- Chandra T, Ewels PA, Schoenfelder S, Furlan-Magaril M, Wingett SW, Kirschner K, Thuret J-Y, Andrews S, Fraser P & Reik W (2015) Global reorganization of the nuclear landscape in senescent cells. *Cell Rep* 10: 471–483
- Chandra T, Kirschner K, Thuret J-Y, Pope BD, Ryba T, Newman S, Ahmed K, Samarajiwa SA, Salama R, Carroll T, *et al* (2012) Independence of Repressive Histone Marks and Chromatin Compaction during Senescent Heterochromatic Layer Formation. *Molecular Cell* 47: 203–214
- Chen C-K, Blanco M, Jackson C, Aznauryan E, Ollikainen N, Surka C, Chow A, Cerase A, McDonel P & Guttman M (2016) Xist recruits the X chromosome to the nuclear lamina to enable chromosome-wide silencing. *Science* 354: 468–472
- Chen CY & Sarnow P (1995) Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* 268: 415–417
- Chen J, Crutchley J, Zhang D, Owzar K & Kastan MB (2017) Identification of a DNA Damage-Induced Alternative Splicing Pathway That Regulates p53 and Cellular Senescence Markers. *Cancer Discov* 7: 766–781
- Chen L-L & Yang L (2015) Regulation of circRNA biogenesis. *RNA Biology* 12: 381–388
- Chen M, Lyu G, Han M, Nie H, Shen T, Chen W, Niu Y, Song Y, Li X, Li H, *et al* (2018a) 3' UTR lengthening as a novel mechanism in regulating cellular senescence. *Genome Res* 28: 285–294
- Chen N, Zhao G, Yan X, Lv Z, Yin H, Zhang S, Song W, Li X, Li L, Du Z, *et al* (2018b) A novel FLI1 exonic circular RNA promotes metastasis in breast cancer by coordinately regulating TET1 and DNMT1. *Genome Biol* 19: 218
- Chen Z, Xu W, Zhang D, Chu J, Shen S, Ma Y, Wang Q, Liu G, Yao T, Huang Y, *et al* (2021) circCAMSAP1 promotes osteosarcoma progression and metastasis by sponging miR-145-5p and regulating FLI1 expression. *Molecular Therapy - Nucleic Acids* 23: 1120–1135
- Cheng J, Metge F & Dieterich C (2016) Specific identification and quantification of circular RNAs from sequencing data. *Bioinformatics* 32: 1094–1096
- Chieffi P, Battista S, Barchi M, Di Agostino S, Pierantoni GM, Fedele M, Chiariotti L, Tramontano D & Fusco A (2002) HMGA1 and HMGA2 protein expression in mouse spermatogenesis. *Oncogene* 21: 3644–3650
- Cho W-K, Spille J-H, Hecht M, Lee C, Li C, Grube V & Cisse II (2018) Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* 361: 412–415

Chong S, Dugast-Darzacq C, Liu Z, Dong P, Dailey GM, Cattoglio C, Heckert A, Banala S, Lavis L, Darzacq X, *et al* (2018) Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* 361

Chopra VS, Kong N & Levine M (2012) Transcriptional repression via antilooping in the *Drosophila* embryo. *Proc Natl Acad Sci U S A* 109: 9460–9464

Clark MB & Mattick JS (2011) Long noncoding RNAs in cell biology. *Seminars in Cell & Developmental Biology* 22: 366–376

Clemson CM, McNeil JA, Willard HF & Lawrence JB (1996) XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* 132: 259–275

Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA & Goodall GJ (2015) The RNA binding protein quaking regulates formation of circRNAs. *Cell* 160: 1125–1134

Coppé J-P, Desprez P-Y, Krtolica A & Campisi J (2010) The senescence-associated secretory phenotype: the dark side of tumor suppression. *Annu Rev Pathol* 5: 99–118

Corces MR, Trevino AE, Hamilton EG, Greenside PG, Sinnott-Armstrong NA, Vesuna S, Satpathy AT, Rubin AJ, Montine KS, Wu B, *et al* (2017) An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat Methods* 14: 959–962

Cotmore SF & Tattersall P (1998) High-mobility group 1/2 proteins are essential for initiating rolling-circle-type DNA replication at a parvovirus hairpin origin. *J Virol* 72: 8477–8484

Crane E, Bian Q, McCord RP, Lajoie BR, Wheeler BS, Ralston EJ, Uzawa S, Dekker J & Meyer BJ (2015) Condensin-driven remodelling of X chromosome topology during dosage compensation. *Nature* 523: 240–244

Cremer T & Cremer C (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2: 292–301

Criscione SW, De Cecco M, Siranosian B, Zhang Y, Kreiling JA, Sedivy JM & Neretti N (2016) Reorganization of chromosome architecture in replicative cellular senescence. *Sci Adv* 2: e1500882

Cruikshanks HA, McBryan T, Nelson DM, Vanderkraats ND, Shah PP, van Tuyn J, Singh Rai T, Brock C, Donahue G, Dunican DS, *et al* (2013) Senescent cells harbour features of the cancer epigenome. *Nat Cell Biol* 15: 1495–1506

Daneshvar K, Ardehali MB, Klein IA, Hsieh F-K, Kratkiewicz AJ, Mahpour A, Cancelliere SOL, Zhou C, Cook BM, Li W, *et al* (2020) lncRNA DIGIT and BRD3 protein form phase-separated condensates to regulate endoderm differentiation. *Nat Cell Biol* 22: 1211–1222

Davaapil H, Brockes JP & Yun MH (2017) Conserved and novel functions of programmed cellular senescence during vertebrate development. *Development* 144: 106–114

Davalos AR, Kawahara M, Malhotra GK, Schaum N, Huang J, Ved U, Beausejour CM, Coppe J-P, Rodier F & Campisi J (2013) p53-dependent release of Alarmin HMGB1 is a central mediator of senescent phenotypes. *J Cell Biol* 201: 613–629

Davidson IF, Bauer B, Goetz D, Tang W, Wutz G & Peters J-M (2019) DNA loop extrusion by human cohesin. *Science* 366: 1338–1345

Davidson IF, Goetz D, Zaczek MP, Molodtsov MI, Huis In 't Veld PJ, Weissmann F, Litos G, Cisneros DA, Ocampo-Hafalla M, Ladurner R, *et al* (2016) Rapid movement and transcriptional re-localization of human cohesin on DNA. *EMBO J* 35: 2671–2685

De Cecco M, Criscione SW, Peterson AL, Neretti N, Sedivy JM & Kreiling JA (2013) Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging (Albany NY)* 5: 867–883

De Maio A, Yalamanchili HK, Adamski CJ, Gennarino VA, Liu Z, Qin J, Jung SY, Richman R, Orr H & Zoghbi HY (2018) RBM17 Interacts with U2SURP and CHERP to Regulate Expression and Splicing of RNA-Processing Proteins. *Cell Reports* 25: 726-736.e7

Degryse B, Bonaldi T, Scaffidi P, Müller S, Resnati M, Sanvito F, Arrighoni G & Bianchi ME (2001) The high mobility group (HMG) boxes of the nuclear protein HMG1 induce chemotaxis and cytoskeleton reorganization in rat smooth muscle cells. *J Cell Biol* 152: 1197–1206

Dekker J, Marti-Renom MA & Mirny LA (2013) Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. *Nat Rev Genet* 14: 390–403

Dekker J, Rippe K, Dekker M & Kleckner N (2002) Capturing chromosome conformation. *Science* 295: 1306–1311

Demaria M, Ohtani N, Youssef SA, Rodier F, Toussaint W, Mitchell JR, Laberge R-M, Vijg J, Van Steeg H, Dollé MET, *et al* (2014) An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev Cell* 31: 722–733

Denholtz M, Zhu Y, He Z, Lu H, Isoda T, Döhrmann S, Nizet V & Murre C (2020) Upon microbial challenge, human neutrophils undergo rapid changes in nuclear architecture and chromatin folding to orchestrate an immediate inflammatory gene program. *Genes Dev* 34: 149–165

Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, *et al* (2012) The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. *Genome Research* 22: 1775–1789

Deschênes M & Chabot B (2017) The emerging role of alternative splicing in senescence and aging. *Aging Cell* 16: 918–933

Di Agostino S, Riccioli A, De Cesaris P, Fontemaggi G, Blandino G, Filippini A & Fazi F (2020) Circular RNAs in Embryogenesis and Cell Differentiation With a Focus on Cancer Development. *Front Cell Dev Biol* 8: 389

Dimova DK & Dyson NJ (2005) The E2F transcriptional network: old acquaintances with new faces. *Oncogene* 24: 2810–2826

Dixon JR, Gorkin DU & Ren B (2016) Chromatin Domains: the Unit of Chromosome Organization. *Mol Cell* 62: 668–680

Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, *et al* (2015) Chromatin Architecture Reorganization during Stem Cell Differentiation. *Nature* 518: 331–336

Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS & Ren B (2012) Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions. *Nature* 485: 376–380

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21

Dolatshad H, Pellagatti A, Fernandez-Mercado M, Yip BH, Malcovati L, Attwood M, Przychodzen B, Sahgal N, Kanapin AA, Lockstone H, *et al* (2015) Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells. *Leukemia* 29: 1092–1103

Dörr JR, Yu Y, Milanovic M, Beuster G, Zasada C, Däbritz JHM, Lisec J, Lenze D, Gerhardt A, Schleicher K, *et al* (2013) Synthetic lethal metabolic targeting of cellular senescence in cancer therapy. *Nature* 501: 421–425

Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, *et al* (2006) Chromosome Conformation Capture Carbon Copy (5C): a massively parallel solution for mapping interactions between genomic elements. *Genome Res* 16: 1299–1309

Dou Y, Cha DJ, Franklin JL, Higginbotham JN, Jeppesen DK, Weaver AM, Prasad N, Levy S, Coffey RJ, Patton JG, *et al* (2016) Circular RNAs are down-regulated in KRAS mutant colon cancer cells and can be transferred to exosomes. *Sci Rep* 6: 37982

Du WW, Fang L, Yang W, Wu N, Awan FM, Yang Z & Yang BB (2017a) Induction of tumor apoptosis through a circular RNA enhancing Foxo3 activity. *Cell Death Differ* 24: 357–370

Du WW, Yang W, Chen Y, Wu Z-K, Foster FS, Yang Z, Li X & Yang BB (2017b) Foxo3 circular RNA promotes cardiac senescence by modulating multiple factors associated with stress and senescence responses. *European Heart Journal* 38: 1402–1412

Du WW, Yang W, Liu E, Yang Z, Dhaliwal P & Yang BB (2016) Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Res* 44: 2846–2858

Du Z, Zheng H, Huang B, Ma R, Wu J, Zhang X, He J, Xiang Y, Wang Q, Li Y, *et al* (2017c) Allelic reprogramming of 3D chromatin architecture during early mammalian development. *Nature* 547: 232–235

Duguet M, Soussi T, Rossignol JM, Méchali M & De Recondo AM (1977) Stimulation of rat liver alpha- and beta-type DNA polymerases by an homologous DNA-unwinding protein. *FEBS Lett* 79: 160–164

Durand NC, Shamim MS, Machol I, Rao SSP, Huntley MH, Lander ES & Aiden EL (2016) Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Syst* 3: 95–98

El Gazzar M, Yoza BK, Chen X, Garcia BA, Young NL & McCall CE (2009) Chromatin-specific remodeling by HMGB1 and linker histone H1 silences proinflammatory genes during endotoxin tolerance. *Mol Cell Biol* 29: 1959–1971

El Khattabi L, Zhao H, Kalchschmidt J, Young N, Jung S, Van Blerkom P, Kieffer-Kwon P, Kieffer-Kwon K-R, Park S, Wang X, *et al* (2019) A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell* 178: 1145-1158.e20

- Elguindy MM & Mendell JT (2021) NORAD-induced Pumilio phase separation is required for genome stability. *Nature*: 1–6
- ENCODE Project Consortium (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489: 57–74
- Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, Surka C, Kadri S, Xing J, Goren A, Lander ES, *et al* (2013) The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341: 1237973
- Erdel F, Rademacher A, Vlijm R, Tünnermann J, Frank L, Weinmann R, Schweigert E, Yserentant K, Hummert J, Bauer C, *et al* (2020) Mouse Heterochromatin Adopts Digital Compaction States without Showing Hallmarks of HP1-Driven Liquid-Liquid Phase Separation. *Molecular Cell* 78: 236-249.e7
- Fanucchi S, Fok ET, Dalla E, Shibayama Y, Börner K, Chang EY, Stoychev S, Imakaev M, Grimm D, Wang KC, *et al* (2019) Immune genes are primed for robust transcription by proximal long noncoding RNAs located in nuclear compartments. *Nat Genet* 51: 138–150
- Farooqi AA, Naureen H & Attar R (2021) Regulation of cell signaling pathways by circular RNAs and microRNAs in different cancers: Spotlight on Wnt/ β -catenin, JAK/STAT, TGF/SMAD, SHH/GLI, NOTCH and Hippo pathways. *Semin Cell Dev Biol*
- Fedele M, Fidanza V, Battista S, Pentimalli F, Klein-Szanto AJP, Visone R, De Martino I, Curcio A, Morisco C, Del Vecchio L, *et al* (2006) Haploinsufficiency of the Hmga1 gene causes cardiac hypertrophy and myelolymphoproliferative disorders in mice. *Cancer Res* 66: 2536–2543
- Fišerová J, Efenberková M, Sieger T, Maninová M, Uhlířová J & Hozák P (2017) Chromatin organization at the nuclear periphery as revealed by image analysis of structured illumination microscopy data. *J Cell Sci* 130: 2066–2077
- Flyamer IM, Illingworth RS & Bickmore WA (2020) Coolpup.py: versatile pile-up analysis of Hi-C data. *Bioinformatics* 36: 2980–2985
- Franz M, Lopes CT, Huck G, Dong Y, Sumer O & Bader GD (2016) Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics* 32: 309–311
- Fraser J, Ferrai C, Chiariello AM, Schueler M, Rito T, Laudanno G, Barbieri M, Moore BL, Kraemer DCA, Aitken S, *et al* (2015) Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. *Mol Syst Biol* 11: 852
- Frottin F, Schueder F, Tiwary S, Gupta R, Körner R, Schlichthaerle T, Cox J, Jungmann R, Hartl FU & Hipp MS (2019) The nucleolus functions as a phase-separated protein quality control compartment. *Science* 365: 342–347
- Fudenberg G, Abdennur N, Imakaev M, Goloborodko A & Mirny LA (2017) Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harb Symp Quant Biol* 82: 45–55
- Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N & Mirny LA (2016) Formation of Chromosomal Domains by Loop Extrusion. *Cell Reports* 15: 2038–2049

Furusawa T, Lim J-H, Catez F, Birger Y, Mackem S & Bustin M (2006) Down-regulation of nucleosomal binding protein HMGN1 expression during embryogenesis modulates Sox9 expression in chondrocytes. *Mol Cell Biol* 26: 592–604

Gambetta MC & Furlong EEM (2018) The Insulator Protein CTCF Is Required for Correct Hox Gene Expression, but Not for Embryonic Development in *Drosophila*. *Genetics* 210: 129–136

Ganji M, Shaltiel IA, Bisht S, Kim E, Kalichava A, Haering CH & Dekker C (2018) Real-time imaging of DNA loop extrusion by condensin. *Science* 360: 102–105

García-Prat L, Martínez-Vicente M, Perdiguero E, Ortet L, Rodríguez-Ubreva J, Rebollo E, Ruiz-Bonilla V, Gutarra S, Ballestar E, Serrano AL, *et al* (2016) Autophagy maintains stemness by preventing senescence. *Nature* 529: 37–42

Gassler J, Brandão HB, Imakaev M, Flyamer IM, Ladstätter S, Bickmore WA, Peters J-M, Mirny LA & Tachibana K (2017) A mechanism of cohesin-dependent loop extrusion organizes zygotic genome architecture. *The EMBO Journal* 36: 3600–3618

Gdynia G, Keith M, Kopitz J, Bergmann M, Fassel A, Weber ANR, George J, Kees T, Zentgraf H-W, Wiestler OD, *et al* (2010) Danger signaling protein HMGB1 induces a distinct form of cell death accompanied by formation of giant mitochondria. *Cancer Res* 70: 8558–8568

Genschel J & Modrich P (2009) Functions of MutLalpha, replication protein A (RPA), and HMGB1 in 5'-directed mismatch repair. *J Biol Chem* 284: 21536–21544

Georgilis A, Klotz S, Hanley CJ, Herranz N, Weirich B, Morancho B, Leote AC, D'Artista L, Gallage S, Seehawer M, *et al* (2018) PTBP1-Mediated Alternative Splicing Regulates the Inflammatory Secretome and the Pro-tumorigenic Effects of Senescent Cells. *Cancer Cell* 34: 85-102.e9

Gerlitz G, Hock R, Ueda T & Bustin M (2009) The dynamics of HMG protein-chromatin interactions in living cells. *Biochem Cell Biol* 87: 127–137

Ghamari A, van de Corput MPC, Thongjuea S, van Cappellen WA, van Ijcken W, van Haren J, Soler E, Eick D, Lenhard B & Grosveld FG (2013) In vivo live imaging of RNA polymerase II transcription factories in primary cells. *Genes Dev* 27: 767–777

Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, Huber W & Furlong EEM (2014) Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 512: 96–100

Giavara S, Kosmidou E, Hande MP, Bianchi ME, Morgan A, d'Adda di Fagnana F & Jackson SP (2005) Yeast Nhp6A/B and mammalian Hmgb1 facilitate the maintenance of genome stability. *Curr Biol* 15: 68–72

Gibcus JH, Samejima K, Goloborodko A, Samejima I, Naumova N, Nuebler J, Kanemaki MT, Xie L, Paulson JR, Earnshaw WC, *et al* (2018) A pathway for mitotic chromosome formation. *Science* 359

Giorgetti L, Lajoie BR, Carter AC, Attia M, Zhan Y, Xu J, Chen CJ, Kaplan N, Chang HY, Heard E, *et al* (2016) Structural organization of the inactive X chromosome in the mouse. *Nature* 535: 575–579

Goel VY & Hansen AS (2021) The macro and micro of chromosome conformation capture. *WIREs Developmental Biology* n/a: e395

Golfier S, Quail T, Kimura H & Brugués J (2020) Cohesin and condensin extrude DNA loops in a cell cycle-dependent manner. *eLife* 9: e53885

Gong W, Zheng Y, Chao F, Li Y, Xu Z, Huang G, Gao X, Li S & He F (2010) The anti-inflammatory activity of HMGB1 A box is enhanced when fused with C-terminal acidic tail. *J Biomed Biotechnol* 2010: 915234

Goodwin GH, Sanders C & Johns EW (1973) A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. *Eur J Biochem* 38: 14–19

Groesser T, Chang H, Fontenay G, Chen J, Costes SV, Helen Barcellos-Hoff M, Parvin B & Rydberg B (2011) Persistence of γ -H2AX and 53BP1 foci in proliferating and non-proliferating human mammary epithelial cells after exposure to γ -rays or iron ions. *Int J Radiat Biol* 87: 696–710

Gu B, Comerici CJ, McCarthy DG, Saurabh S, Moerner WE & Wysocka J (2020) Opposing effects of cohesin and transcription on CTCF organization revealed by super-resolution imaging. *Mol Cell* 80: 699–711

Gu S, Ran S, Liu B & Liang J (2016) miR-152 induces human dental pulp stem cell senescence by inhibiting SIRT7 expression. *FEBS Lett* 590: 1123–1131

Guo JU, Agarwal V, Guo H & Bartel DP (2014) Expanded identification and characterization of mammalian circular RNAs. *Genome Biol* 15: 409

Guo Y, Wei X & Peng Y (2020) Structure-Mediated Degradation of CircRNAs. *Trends Cell Biol* 30: 501–503

Guo YE, Manteiga JC, Henninger JE, Sabari BR, Dall’Agnese A, Hannett NM, Spille J-H, Afeyan LK, Zamudio AV, Shrinivas K, *et al* (2019) Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* 572: 543–548

Gusmao EG, Dieterich C, Zenke M & Costa IG (2014) Detection of active transcription factor binding sites with the combination of DNase hypersensitivity and histone modifications. *Bioinformatics* 30: 3143–3151

Haarhuis JHI, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, Krijger PHL, Teunissen H, Medema RH, van Steensel B, *et al* (2017) The Cohesin Release Factor WAPL Restricts Chromatin Loop Extension. *Cell* 169: 693-707.e14

Hahne F & Ivanek R (2016) Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol* 1418: 335–351

Halford SE & Marko JF (2004) How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res* 32: 3040–3052

Handwerger KE, Cordero JA & Gall JG (2005) Cajal Bodies, Nucleoli, and Speckles in the *Xenopus* Oocyte Nucleus Have a Low-Density, Sponge-like Structure. *Mol Biol Cell* 16: 202–211

Hansen AS, Cattoglio C, Darzacq X & Tjian R (2018) Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* 9: 20–32

Hansen AS, Hsieh T-HS, Cattoglio C, Pustova I, Saldaña-Meyer R, Reinberg D, Darzacq X & Tjian R (2019) Distinct Classes of Chromatin Loops Revealed by Deletion of an RNA-Binding Region in CTCF. *Mol Cell* 76: 395-411.e13

Hansen AS, Pustova I, Cattoglio C, Tjian R & Darzacq X (2017) CTCF and cohesin regulate chromatin loop stability with distinct dynamics. *eLife* 6: 25776

Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK & Kjems J (2013) Natural RNA circles function as efficient microRNA sponges. *Nature* 495: 384–388

Hansen TB, Wiklund ED, Bramsen JB, Villadsen SB, Statham AL, Clark SJ & Kjems J (2011) miRNA-dependent gene silencing involving Ago2-mediated cleavage of a circular antisense RNA. *EMBO J* 30: 4414–4422

Hanssen LLP, Kassouf MT, Oudelaar AM, Biggs D, Preece C, Downes DJ, Gosden M, Sharpe JA, Sloane-Stanley JA, Hughes JR, *et al* (2017) Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nat Cell Biol* 19: 952–961

Haque S, Ames RM, Moore K, Lee BP, Jeffery N & Harries LW (2020) Islet-expressed circular RNAs are associated with type 2 diabetes status in human primary islets and in peripheral blood. *BMC Med Genomics* 13: 64

Harr JC, Luperchio TR, Wong X, Cohen E, Wheelan SJ & Reddy KL (2015) Directed targeting of chromatin to the nuclear lamina is mediated by chromatin state and A-type lamins. *J Cell Biol* 208: 33–52

Hayflick L & Moorhead PS (1961) The serial cultivation of human diploid cell strains. *Experimental Cell Research* 25: 585–621

Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, Rialdi A, White KM, Albrecht RA, Pache L, *et al* (2018) Transcription Elongation Can Affect Genome 3D Structure. *Cell* 174: 1522-1536.e22

Heisterkamp D & Johnson J (2013) Pinch Ratio Clustering from a Topologically Intrinsic Lexicographic Ordering. In *Proceedings of the 2013 SIAM International Conference on Data Mining* pp 560–568. Society for Industrial and Applied Mathematics

Heller D, Krestel R, Ohler U, Vingron M & Marsico A (2017) ssHMM: extracting intuitive sequence-structure motifs from high-throughput RNA-binding protein data. *Nucleic Acids Res* 45: 11004–11018

Hernandez-Segura A, Nehme J & Demaria M (2018) Hallmarks of Cellular Senescence. *Trends in Cell Biology* 28: 436–453

Herranz N & Gil J (2018) Mechanisms and functions of cellular senescence. *J Clin Invest* 128: 1238–1246

Hezroni H, Koppstein D, Schwartz MG, Avrutin A, Bartel DP & Ulitsky I (2015) Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep* 11: 1110–1122

Hildebrand EM & Dekker J (2020) Mechanisms and Functions of Chromosome Compartmentalization. *Trends in Biochemical Sciences* 45: 385–396

Hnisz D, Shrinivas K, Young RA, Chakraborty AK & Sharp PA (2017) A Phase Separation Model for Transcriptional Control. *Cell* 169: 13–23

Hock R, Furusawa T, Ueda T & Bustin M (2007) HMG chromosomal proteins in development and disease. *Trends Cell Biol* 17: 72–79

Holly AC, Pilling LC, Hernandez D, Lee BP, Singleton A, Ferrucci L, Melzer D & Harries LW (2014) Splicing factor 3B1 hypomethylation is associated with altered SF3B1 transcript expression in older humans. *Mechanisms of Ageing and Development* 135: 50–56

Hsieh T-HS, Cattoglio C, Slobodyanyuk E, Hansen AS, Rando OJ, Tjian R & Darzacq X (2020) Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Mol Cell* 78: 539-553.e8

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

Hsiung CCS, Bartman CR, Huang P, Ginart P, Stonestrom AJ, Keller CA, Face C, Jahn KS, Evans P, Sankaranarayanan L, *et al* (2016) A hyperactive transcriptional state marks genome reactivation at the mitosis-G1 transition. *Genes Dev* 30: 1423–1439

Hsiung CC-S, Morrissey CS, Udugama M, Frank CL, Keller CA, Baek S, Giardine B, Crawford GE, Sung M-H, Hardison RC, *et al* (2015) Genome accessibility is widely preserved and locally modulated during mitosis. *Genome Res* 25: 213–225

Hsu MT & Coca-Prados M (1979) Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature* 280: 339–340

Hug CB, Grimaldi AG, Kruse K & Vaquerizas JM (2017) Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* 169: 216-228.e19

Huttunen HJ, Fages C, Kuja-Panula J, Ridley AJ & Rauvala H (2002) Receptor for advanced glycation end products-binding COOH-terminal motif of amphoterin inhibits invasive migration and metastasis. *Cancer Res* 62: 4805–4811

Iborra FJ, Pombo A, Jackson DA & Cook PR (1996) Active RNA polymerases are localized within discrete transcription “factories” in human nuclei. *Journal of Cell Science* 109: 1427–1436

Isoda T, Moore AJ, He Z, Chandra V, Aida M, Denholtz M, Piet van Hamburg J, Fisch KM, Chang AN, Fahl SP, *et al* (2017) Non-coding Transcription Instructs Chromatin Folding and Compartmentalization to Dictate Enhancer-Promoter Communication and T Cell Fate. *Cell* 171: 103-119.e18

Ito H, Fujita K, Tagawa K, Chen X, Homma H, Sasabe T, Shimizu J, Shimizu S, Tamura T, Muramatsu S, *et al* (2015) HMGB1 facilitates repair of mitochondrial DNA damage and extends the lifespan of mutant ataxin-1 knock-in mice. *EMBO Mol Med* 7: 78–101

Ito Y, Hoare M & Narita M (2017) Spatial and Temporal Control of Senescence. *Trends Cell Biol* 27: 820–832

Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, Piechotta M, Levanon EY, Landthaler M, Dieterich C, *et al* (2015) Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. *Cell Rep* 10: 170–177

Ivanov IP, Shin B-S, Loughran G, Tzani I, Young-Baird SK, Cao C, Atkins JF & Dever TE (2018) Polyamine Control of Translation Elongation Regulates Start Site Selection on Antizyme Inhibitor mRNA via Ribosome Queuing. *Mol Cell* 70: 254-264.e6

- Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, Barrette TR, Prensner JR, Evans JR, Zhao S, *et al* (2015) The Landscape of Long Noncoding RNAs in the Human Transcriptome. *Nat Genet* 47: 199–208
- Jang C-Y, Kim HD, Zhang X, Chang J-S & Kim J (2012) Ribosomal protein S3 localizes on the mitotic spindle and functions as a microtubule associated protein in mitosis. *Biochemical and Biophysical Research Communications* 429: 57–62
- Jarrard DF, Sarkar S, Shi Y, Yeager TR, Magrane G, Kinoshita H, Nassif N, Meisner L, Newton MA, Waldman FM, *et al* (1999) p16/pRb pathway alterations are required for bypassing senescence in human prostate epithelial cells. *Cancer Res* 59: 2957–2964
- Javaherian K, Liu JF & Wang JC (1978) Nonhistone proteins HMG1 and HMG2 change the DNA helical structure. *Science* 199: 1345–1346
- Jeck WR & Sharpless NE (2014) Detecting and characterizing circular RNAs. *Nat Biotechnol* 32: 453–461
- Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF & Sharpless NE (2013) Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* 19: 141–157
- Jiang Y, Huang J, Lun K, Li B, Zheng H, Li Y, Zhou R, Duan W, Wang C, Feng Y, *et al* (2020) Genome-wide analyses of chromatin interactions after the loss of Pol I, Pol II, and Pol III. *Genome Biology* 21: 158
- Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, Yen C-A, Schmitt AD, Espinoza C & Ren B (2013) A high-resolution map of three-dimensional chromatin interactome in human cells. *Nature* 503: 290–294
- Jing H, Vakoc CR, Ying L, Mandat S, Wang H, Zheng X & Blobel GA (2008) Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol Cell* 29: 232–242
- Johnson J (2012) Topological graph clustering with thin position. *arXiv:12060771 [cs, math, stat]*
- Kaaij LJT, van der Weide RH, Ketting RF & de Wit E (2018) Systemic Loss and Gain of Chromatin Architecture throughout Zebrafish Development. *Cell Rep* 24: 1-10.e4
- Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, van Berkum NL, Ebmeier CC, Goossens J, Rahl PB, Levine SS, *et al* (2010) Mediator and Cohesin Connect Gene Expression and Chromatin Architecture. *Nature* 467: 430–435
- Kang C, Xu Q, Martin TD, Li MZ, Demaria M, Aron L, Lu T, Yankner BA, Campisi J & Elledge SJ (2015) The DNA damage response induces inflammation and senescence by inhibiting autophagy of GATA4. *Science* 349: aaa5612
- Kang H, Shokhirev MN, Xu Z, Chandran S, Dixon JR & Hetzer MW (2020) Dynamic regulation of histone modifications and long-range chromosomal interactions during postmitotic transcriptional reactivation. *Genes Dev* 34: 913–930
- Kang R, Chen R, Zhang Q, Hou W, Wu S, Cao L, Huang J, Yu Y, Fan X, Yan Z, *et al* (2014) HMGB1 in Health and Disease. *Mol Aspects Med* 0: 1–116

Kargapolova Y, Levin M, Lackner K & Danckwardt S (2017) sCLIP—an integrated platform to study RNA–protein interactomes in biomedical research: identification of CSTF2tau in alternative processing of small nuclear RNAs. *Nucleic Acids Res* 45: 6074–6086

Kawase T, Sato K, Ueda T & Yoshida M (2008) Distinct domains in HMGB1 are involved in specific intramolecular and nucleosomal interactions. *Biochemistry* 47: 13991–13996

Ke Y, Xu Y, Chen X, Feng S, Liu Z, Sun Y, Yao X, Li F, Zhu W, Gao L, *et al* (2017) 3D Chromatin Structures of Mature Gametes and Structural Reprogramming during Mammalian Embryogenesis. *Cell* 170: 367–381.e20

Kim S & Shendure J (2019) Mechanisms of interplay between transcription factors and the 3D genome. *Mol Cell* 76: 306–319

Kim WY & Sharpless NE (2006) The Regulation of INK4/ARF in Cancer and Aging. *Cell* 127: 265–275

Kim Y, Shi Z, Zhang H, Finkelstein IJ & Yu H (2019) Human cohesin compacts DNA by loop extrusion. *Science* 366: 1345–1349

Knupp D & Miura P (2018) CircRNA accumulation: A new hallmark of aging? *Mech Ageing Dev* 173: 71–79

Kohlstaedt LA & Cole RD (1994) Specific interaction between H1 histone and high mobility protein HMG1. *Biochemistry* 33: 570–575

Kolovos P, Georgomanolis T, Koeflerle A, Larkin JD, Brant L, Nikolic M, Gusmao EG, Zirkel A, Knoch TA, van Ijcken WF, *et al* (2016) Binding of nuclear factor κ B to noncanonical consensus sites reveals its multimodal role during the early inflammatory response. *Genome Res* 26: 1478–1489

Koopman P, Münsterberg A, Capel B, Vivian N & Lovell-Badge R (1990) Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature* 348: 450–452

Körner U, Bustin M, Scheer U & Hock R (2003) Developmental role of HMGN proteins in *Xenopus laevis*. *Mech Dev* 120: 1177–1192

Krizhanovsky V, Yon M, Dickins RA, Hearn S, Simon J, Miething C, Yee H, Zender L & Lowe SW (2008) Senescence of activated stellate cells limits liver fibrosis. *Cell* 134: 657–667

Kruse K, Hug CB, Hernández-Rodríguez B & Vaquerizas JM (2016) TADtool: visual parameter identification for TAD-calling algorithms. *Bioinformatics* 32: 3190–3192

Kruse K, Hug CB & Vaquerizas JM (2020) FAN-C: a feature-rich framework for the analysis and visualisation of chromosome conformation capture data. *Genome Biology* 21: 303

Kuang S & Wang L (2020) Identification and analysis of consensus RNA motifs binding to the genome regulator CTCF. *NAR Genomics and Bioinformatics* 2

Kuilman T & Peeper DS (2009) Senescence-messaging secretome: SMS-ing cellular stress. *Nat Rev Cancer* 9: 81–94

Kung JTY, Colognori D & Lee JT (2013) Long Noncoding RNAs: Past, Present, and Future. *Genetics* 193: 651–669

- Lai WKM & Pugh BF (2017) Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nature Reviews Molecular Cell Biology* 18: 548–562
- Lange SS & Vasquez KM (2009) HMGB1: the jack-of-all-trades protein is a master DNA repair mechanic. *Mol Carcinog* 48: 571–580
- Langmead B & Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357–359
- Larson AG, Elnatan D, Keenen MM, Trnka MJ, Johnston JB, Burlingame AL, Agard DA, Redding S & Narlikar GJ (2017) Liquid droplet formation by HP1 α suggests a role for phase separation in heterochromatin. *Nature* 547: 236–240
- Lasda E & Parker R (2014) Circular RNAs: diversity of form and function. *RNA* 20: 1829–1842
- Latos PA, Pauler FM, Koerner MV, Şenergin HB, Hudson QJ, Stocsits RR, Allhoff W, Stricker SH, Klement RM, Warczok KE, *et al* (2012) Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science* 338: 1469–1472
- Lawrence M, Daujat S & Schneider R (2016) Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet* 32: 42–56
- Le TBK, Imakaev MV, Mirny LA & Laub MT (2013) High-resolution mapping of the spatial organization of a bacterial chromosome. *Science* 342: 731–734
- Lee ECS, Elhassan SAM, Lim GPL, Kok WH, Tan SW, Leong EN, Tan SH, Chan EWL, Bhattamisra SK, Rajendran R, *et al* (2019a) The roles of circular RNAs in human development and diseases. *Biomedicine & Pharmacotherapy* 111: 198–208
- Lee J-J, Park IH, Kwak MS, Rhee WJ, Kim SH & Shin J-S (2021) HMGB1 orchestrates STING-mediated senescence via TRIM30 α modulation in cancer cells. *Cell Death Discov* 7: 28
- Lee SB, Kwon I-S, Park J, Lee K-H, Ahn Y, Lee C, Kim J, Choi SY, Cho S-W & Ahn J-Y (2010) Ribosomal Protein S3, a New Substrate of Akt, Serves as a Signal Mediator between Neuronal Apoptosis and DNA Repair. *Journal of Biological Chemistry* 285: 29457–29468
- Lee W, Ku S-K, Lee Y-M & Bae J-S (2014) Anti-septic effects of glyceollins in HMGB1-induced inflammatory responses in vitro and in vivo. *Food Chem Toxicol* 63: 1–8
- Lee W-J, Moon J, Jeon D, Shin Y-W, Yoo J-S, Park D-K, Lee S-T, Jung K-H, Park K-I, Jung K-Y, *et al* (2019b) Possible epigenetic regulatory effect of dysregulated circular RNAs in Alzheimer’s disease model. *Scientific Reports* 9: 11956
- Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, *et al* (2017) Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol Cell* 66: 22-37.e9
- Léveillé N, Melo CA, Rooijers K, Díaz-Lagares A, Melo SA, Korkmaz G, Lopes R, Moqadam FA, Maia AR, Wijchers PJ, *et al* (2015) Genome-wide profiling of p53-regulated enhancer RNAs uncovers a subset of enhancers controlled by a lncRNA. *Nature Communications* 6: 6520

Levi N, Papismadov N, Solomonov I, Sagi I & Krizhanovsky V (2020) The ECM path of senescence in aging: components and modifiers. *FEBS J* 287: 2636–2646

Li F, Zhang L, Li W, Deng J, Zheng J, An M, Lu J & Zhou Y (2015a) Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/ β -catenin pathway. *Oncotarget* 6: 6001–6013

Li J, Kokkola R, Tabibzadeh S, Yang R, Ochani M, Qiang X, Harris HE, Czura CJ, Wang H, Ulloa L, *et al* (2003) Structural Basis for the Proinflammatory Cytokine Activity of High Mobility Group Box 1. *Mol Med* 9: 37–45

Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, Ong C-T, Cubeñas-Potts C, Hu M, Lei EP, Bosco G, *et al* (2015b) Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol Cell* 58: 216–231

Li Q, Li J, Wen T, Zeng W, Peng C, Yan S, Tan J, Yang K, Liu S, Guo A, *et al* (2014) Overexpression of HMGB1 in melanoma predicts patient survival and suppression of HMGB1 induces cell cycle arrest and senescence in association with p21 (Waf1/Cip1) up-regulation via a p53-independent, Sp1-dependent pathway. *Oncotarget* 5: 6387–6403

Li R, Ke S, Meng F, Lu J, Zou X, He Z, Wang W & Fang M (2018) CiRS-7 promotes growth and metastasis of esophageal squamous cell carcinoma via regulation of miR-7/HOXB13. *Cell Death Dis* 9: 838

Li Y, Haarhuis JHI, Ángela SC, Roel O, Ruiten MS, Willems L, Teunissen H, Muir KW, Wit E, Rowland BD, *et al* (2020) The structural basis for cohesin-CTCF-anchored loops. *Nature* 578: 472–476

Li Y, Zheng Q, Bao C, Li S, Guo W, Zhao J, Chen D, Gu J, He X & Huang S (2015c) Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. *Cell Res* 25: 981–984

Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, *et al* (2015d) Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol* 22: 256–264

Liang W-C, Wong C-W, Liang P-P, Shi M, Cao Y, Rao S-T, Tsui SK-W, Waye MM-Y, Zhang Q, Fu W-M, *et al* (2019) Translation of the circular RNA circ β -catenin promotes liver cancer cell growth through activation of the Wnt pathway. *Genome Biology* 20: 84

Liao Y, Smyth GK & Shi W (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30: 923–930

Lieberman-Aiden E, van Berkum NL, Williams L, Imaekae M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, *et al* (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326: 289–293

Little AJ, Corbett E, Ortega F & Schatz DG (2013) Cooperative recruitment of HMGB1 during V(D)J recombination through interactions with RAG1 and DNA. *Nucleic Acids Res* 41: 3289–3301

Liu C-X, Li X, Nan F, Jiang S, Gao X, Guo S-K, Xue W, Cui Y, Dong K, Ding H, *et al* (2019) Structure and Degradation of Circular RNAs Regulate PKR Activation in Innate Immunity. *Cell* 177: 865-880.e21

Loayza-Puch F, Drost J, Rooijers K, Lopes R, Elkon R & Agami R (2013) p53 induces transcriptional and translational programs to suppress cell proliferation and growth. *Genome Biol* 14: R32

Lohani N & Rajeswari MR (2016) Dichotomous Life of DNA Binding High Mobility Group Box1 Protein in Human Health and Disease. *Curr Protein Pept Sci* 17: 762–775

Lotze MT & Tracey KJ (2005) High-mobility group box 1 protein (HMGB1): nuclear weapon in the immune arsenal. *Nat Rev Immunol* 5: 331–342

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

Lu Y, Wu T, Gutman O, Lu H, Zhou Q, Henis YI & Luo K (2020) Phase separation of TAZ compartmentalizes the transcription machinery to promote gene expression. *Nat Cell Biol* 22: 453–464

Luger K, Mäder AW, Richmond RK, Sargent DF & Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389: 251–260

Luo J, Qu L, Gao F, Lin J, Liu J & Lin A (2021a) LncRNAs: Architectural Scaffolds or More Potential Roles in Phase Separation. *Front Genet* 12

Luo Z, Lu L, Tang Q, Wei W, Chen P, Chen Y, Pu J & Wang J (2021b) CircCAMSAP1 promotes hepatocellular carcinoma progression through miR-1294/GRAMD1A pathway. *J Cell Mol Med* 25: 3793–3802

Maass PG, Glažar P, Memczak S, Dittmar G, Hollfinger I, Schreyer L, Sauer AV, Toka O, Aiuti A, Luft FC, *et al* (2017) A map of human circular RNAs in clinically relevant tissues. *J Mol Med (Berl)* 95: 1179–1189

Madsen JGS, Schmidt SF, Larsen BD, Loft A, Nielsen R, & S. (2015) Mandrup, iRNA-seq: computational method for genome-wide assessment of acute transcriptional regulation from total RNA-seq data. *Nucleic Acids Res* 43: 40

Maher JF & Nathans D (1996) Multivalent DNA-binding properties of the HMG-1 proteins. *Proc Natl Acad Sci U S A* 93: 6716–6720

Marchal C, Sima J & Gilbert DM (2019) Control of DNA replication timing in the 3D genome. *Nat Rev Mol Cell Biol* 20: 721–737

Marques AC & Ponting CP (2009) Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness. *Genome Biol* 10: R124

Marrugo J, Marsh DG & Ghosh B (1996) The conserved lymphokine element-0 in the IL5 promoter binds to a high mobility group-1 protein. *Mol Immunol* 33: 1119–1125

McConnell BB, Starborg M, Brookes S & Peters G (1998) Inhibitors of cyclin-dependent kinases induce features of replicative senescence in early passage human diploid fibroblasts. *Curr Biol* 8: 351–354

McHugh CA & Guttman M (2018) RAP-MS: A Method to Identify Proteins that Interact Directly with a Specific RNA Molecule in Cells. *Methods Mol Biol* 1649: 473–488

Meers MP, Tenenbaum D & Henikoff S (2019) Peak calling by Sparse Enrichment Analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin* 12: 42

- Melloni E, Sparatore B, Patrone M, Pessino A, Passalacqua M & Pontremoli S (1995) Extracellular release of the 'differentiation enhancing factor', a HMG1 protein type, is an early step in murine erythroleukemia cell differentiation. *FEBS Lett* 368: 466–470
- Melnik S, Caudron-Herger M, Brant L, Carr IM, Rippe K, Cook PR & Papantonis A (2016) Isolation of the protein and RNA content of active sites of transcription from mammalian cells. *Nat Protoc* 11: 553–565
- Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, *et al* (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* 495: 333–338
- Minajigi A, Froberg J, Wei C, Sunwoo H, Kesner B, Colognori D, Lessing D, Payer B, Boukhali M, Haas W, *et al* (2015) A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* 349
- Misteli T (2020) The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* 183: 28–45
- Mitchell JA & Fraser P (2008) Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev* 22: 20–25
- Mitsouras K, Wong B, Arayata C, Johnson RC & Carey M (2002) The DNA architectural protein HMGB1 displays two distinct modes of action that promote enhanceosome assembly. *Mol Cell Biol* 22: 4390–4401
- Mootha VK, Lindgren CM, Eriksson K-F, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstråle M, Laurila E, *et al* (2003) PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics* 34: 267–273
- Mukherjee A & Vasquez KM (2016) HMGB1 interacts with XPA to facilitate the processing of DNA interstrand crosslinks in human cells. *Nucleic Acids Res* 44: 1151–1160
- Müller S, Ronfani L & Bianchi ME (2004) Regulated expression and subcellular localization of HMGB1, a chromatin protein with a cytokine function. *J Intern Med* 255: 332–343
- Nacarelli T, Liu P & Zhang R (2017) Epigenetic Basis of Cellular Senescence and Its Implications in Aging. *Genes (Basel)* 8
- Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, Dean W, Laue ED, Tanay A & Fraser P (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502: 59–64
- Nagano T, Lubling Y, Várnai C, Dudley C, Leung W, Baran Y, Mendelson Cohen N, Wingett S, Fraser P & Tanay A (2017) Cell-cycle dynamics of chromosomal organization at single-cell resolution. *Nature* 547: 61–67
- Nagashima R, Hibino K, Ashwin SS, Babokhov M, Fujishiro S, Imai R, Nozaki T, Tamura S, Tani T, Kimura H, *et al* (2019) Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. *J Cell Biol* 218: 1511–1530
- Narita M (2007) Cellular senescence and chromatin organisation. *Br J Cancer* 96: 686–691

Narita M, Narita M, Krizhanovsky V, Nuñez S, Chicas A, Hearn SA, Myers MP & Lowe SW (2006) A novel role for high-mobility group a proteins in cellular senescence and heterochromatin formation. *Cell* 126: 503–514

Narita M, Nuñez S, Heard E, Narita M, Lin AW, Hearn SA, Spector DL, Hannon GJ & Lowe SW (2003) Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* 113: 703–716

Narita M, Young ARJ, Arakawa S, Samarajiwa SA, Nakashima T, Yoshida S, Hong S, Berry LS, Reichelt S, Ferreira M, *et al* (2011) Spatial coupling of mTOR and autophagy augments secretory phenotypes. *Science* 332: 966–970

Naughton C, Avlonitis N, Corless S, Prendergast JG, Mati IK, Eijk PP, Cockroft SL, Bradley M, Ylstra B & Gilbert N (2013) Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nat Struct Mol Biol* 20: 387–395

Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA & Dekker J (2013) Organization of the mitotic chromosome. *Science* 342: 948–953

Nemeth MJ, Curtis DJ, Kirby MR, Garrett-Beal LJ, Seidel NE, Cline AP & Bodine DM (2003) Hmgb3: an HMG-box family member expressed in primitive hematopoietic cells that inhibits myeloid and B-cell differentiation. *Blood* 102: 1298–1306

Nielsen CF, Zhang T, Barisic M, Kalitsis P & Hudson DF (2020) Topoisomerase II α is essential for maintenance of mitotic chromosome structure. *Proc Natl Acad Sci USA* 117: 12131–12142

Nora EP, Caccianini L, Fudenberg G, So K, Kameswaran V, Nagle A, Uebersohn A, Hajj B, Saux AL, Coulon A, *et al* (2020) Molecular basis of CTCF binding polarity in genome folding. *Nat Commun* 11: 5612

Nora EP, Goloborodko A, Valton A-L, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA & Bruneau BG (2017) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* 169: 930-944.e22

Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA & Bruneau BG (2019) Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* 169: 930–944

Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, *et al* (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485: 381–385

Nuebler J, Fudenberg G, Imakaev M, Abdennur N & Mirny LA (2018) Chromatin organization by an interplay of loop extrusion and compartmental segregation. *Proc Natl Acad Sci U S A* 115: E6697–E6706

Oertlin C, Lorent J, Murie C, Furic L, Topisirovic I & Larsson O (2019) Generally applicable transcriptome-wide analysis of translation using anota2seq. *Nucleic Acids Research* 47: e70–e70

Olan I, Parry AJ, Schoenfelder S, Narita M, Ito Y, Chan ASL, Slater GSC, Bihary D, Bando M, Shirahige K, *et al* (2020) Transcription-dependent cohesin repositioning rewires chromatin loops in cellular senescence. *Nat Commun* 11: 6049

Oomen ME, Hansen AS, Liu Y, Darzacq X & Dekker J (2019) CTCF sites display cell cycle-dependent dynamics in factor binding and nucleosome positioning. *Genome Res* 29: 236–249

Orlova VV, Choi EY, Xie C, Chavakis E, Bierhaus A, Ihanus E, Ballantyne CM, Gahmberg CG, Bianchi ME, Nawroth PP, *et al* (2007) A novel pathway of HMGB1-mediated inflammatory cell recruitment that requires Mac-1-integrin. *EMBO J* 26: 1129–1139

Oudet P, Gross-Bellard M & Chambon P (1975) Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* 4: 281–300

Pal S & Tyler JK (2016) Epigenetics and aging. *Sci Adv* 2: e1600584

Palikyras S & Papantonis A (2019) Modes of phase separation affecting chromatin regulation. *Open Biol* 9: 190167

Pallier C, Scaffidi P, Chopineau-Proust S, Agresti A, Nordmann P, Bianchi ME & Marechal V (2003) Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes. *Mol Biol Cell* 14: 3414–3426

Palumbo R & Bianchi ME (2004) High mobility group box 1 protein, a cue for stem cell recruitment. *Biochem Pharmacol* 68: 1165–1170

Panda AC (2018) Circular RNAs Act as miRNA Sponges. *Adv Exp Med Biol* 1087: 67–79

Papantonis A & Cook PR (2011) Fixing the model for transcription: the DNA moves, not the polymerase. *Transcription* 2: 41–44

Papantonis A & Cook PR (2013) Transcription factories: genome organization and gene regulation. *Chem Rev* 113: 8683–8705

Park OH, Ha H, Lee Y, Boo SH, Kwon DH, Song HK & Kim YK (2019) Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. *Mol Cell* 74: 494-507.e8

Park PJ (2009) ChIP–seq: advantages and challenges of a maturing technology. *Nat Rev Genet* 10: 669–680

Parmar JJ, Woringer M & Zimmer C (2019) How the Genome Folds: The Biophysics of Four-Dimensional Chromatin Organization. *Annu Rev Biophys* 48: 231–253

Parry AJ, Hoare M, Bihary D, Hänsel-Hertsch R, Smith S, Tomimatsu K, Mannion E, Smith A, D’Santos P, Russell IA, *et al* (2018) NOTCH-mediated non-cell autonomous regulation of chromatin structure during senescence. *Nat Commun* 9: 1840

Parry AJ & Narita M (2016) Old cells, new tricks: chromatin structure in senescence. *Mamm Genome* 27: 320–331

Passarge E (1979) Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am J Hum Genet* 31: 106–115

Pelham-Webb B, Polyzos A, Wojenski L, Kloetgen A, Li J, Giammartino DCD, Sakellaropoulos T, Tsirigos A, Core L & Apostolou E (2021) H3K27ac bookmarking promotes rapid post-mitotic activation of the pluripotent stem cell program without impacting 3D chromatin reorganization. *Mol Cell* 81: 1732–1748

Penagos-Puig A & Furlan-Magaril M (2020) Heterochromatin as an Important Driver of Genome Organization. *Front Cell Dev Biol* 8

Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SWM, Solovei I, Brugman W, Gräf S, Flicek P, Kerkhoven RM, van Lohuizen M, *et al* (2010) Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol Cell* 38: 603–613

Pil PM, Chow CS & Lippard SJ (1993) High-mobility-group 1 protein mediates DNA bending as determined by ring closures. *PNAS* 90: 9465–9469

Platanitis E & Decker T (2018) Regulatory Networks Involving STATs, IRFs, and NFκB in Inflammation. *Front Immunol* 9: 2542

Plimpton S (1995) Fast parallel algorithms for short-range molecular dynamics. *J Comp Phys* 1: 1–19

Plys AJ, Davis CP, Kim J, Rizki G, Keenen MM, Marr SK & Kingston RE (2019) Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev* 33: 799–813

Potier M-C, Rivals I, Mercier G, Ettwiller L, Moldrich RX, Laffaire J, Personnaz L, Rossier J & Dauphinot L (2006) Transcriptional disruptions in Down syndrome: a case study in the Ts1Cje mouse cerebellum during post-natal development. *J Neurochem* 97 Suppl 1: 104–109

Prasad R, Liu Y, Deterding LJ, Poltoratsky VP, Kedar PS, Horton JK, Kanno S-I, Asagoshi K, Hou EW, Khodyreva SN, *et al* (2007) HMGB1 is a cofactor in mammalian base excision repair. *Mol Cell* 27: 829–841

Prasad S & Thakur MK (1990) Distribution of high mobility group proteins in different tissues of rats during aging. *Biochem Int* 20: 687–695

Quinlan AR & Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841–842

Rada-Iglesias A, Grosveld FG & Papantonis A (2018) Forces driving the three-dimensional folding of eukaryotic genomes. *Mol Syst Biol* 14: 8214

Rai TS, Cole JJ, Nelson DM, Dikovskaya D, Faller WJ, Vizioli MG, Hewitt RN, Anannya O, McBryan T, Manoharan I, *et al* (2014) HIRA orchestrates a dynamic chromatin landscape in senescence and is required for suppression of neoplasia. *Genes Dev* 28: 2712–2725

Ramírez F, Dünder F, Diehl S, Grüning BA & Manke T (2014) DeepTools: A flexible platform for exploring deep-sequencing data. *Nucleic Acids Res* 42: 187–191

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dünder F & Manke T (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* 44: W160–W165

Rao SSP, Huang S-C, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-R, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID, *et al* (2017) Cohesin Loss Eliminates All Loop Domains. *Cell* 171: 305-320.e24

Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, *et al* (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159: 1665–1680

Reeve JN (2003) Archaeal chromatin and transcription. *Molecular Microbiology* 48: 587–598

Reeves R (2001) Molecular biology of HMGA proteins: hubs of nuclear function. *Gene* 277: 63–81

Reeves R & Beckerbauer L (2001) HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim Biophys Acta* 1519: 13–29

Rhodes JDP, Feldmann A, Hernandez-Rodriguez B, Diaz N, Brown JM, Fursova NA, Blackledge NP, Prathapan P, Dobrinic P, Huseyin MK, *et al* (2020) Cohesin disrupts polycomb-dependent chromosome interactions in embryonic stem cells. *Cell Rep* 30: 820–835

Risso D, Ngai J, Speed TP & Dudoit S (2014) Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat Biotechnol* 32: 896–902

Ronfani L, Ferraguti M, Croci L, Ovitt CE, Schöler HR, Consalez GG & Bianchi ME (2001) Reduced fertility and spermatogenesis defects in mice lacking chromosomal protein Hmgb2. *Development* 128: 1265–1273

Rosencrance CD, Ammouri HN, Yu Q, Ge T, Rendleman EJ, Marshall SA & Eagen KP (2020) Chromatin hyperacetylation impacts chromosome folding by forming a nuclear subcompartment. *Mol Cell* 78: 112–126

Roukos V, Pegoraro G, Voss TC & Misteli T (2015) Cell cycle staging of individual cells by fluorescence microscopy. *Nat Protoc* 10: 334–348

Rowell JP, Simpson KL, Stott K, Watson M & Thomas JO (2012) HMGB1-facilitated p53 DNA binding occurs via HMG-Box/p53 transactivation domain interaction, regulated by the acidic tail. *Structure* 20: 2014–2024

Rowley MJ & Corces VG (2018) Organizational principles of 3D genome architecture. *Nat Rev Genet* 19: 789–800

Rowley MJ, Lyu X, Rana V, Ando-Kuri M, Karns R, Bosco G & Corces VG (2019) Condensin II counteracts cohesin and RNA polymerase II in the establishment of 3D chromatin organization. *Cell Rep* 26: 2890–2903

Rowley MJ, Nichols MH, Lyu X, Ando-Kuri M, Rivera ISM, Hermetz K, Wang P, Ruan Y & Corces VG (2017) Evolutionarily Conserved Principles Predict 3D Chromatin Organization. *Molecular Cell* 67: 837-852.e7

Rowley MJ, Poulet A, Nichols MH, Bixler BJ, Sanborn AL, Brouhard EA, Hermetz K, Linsenbaum H, Csankovszki G, Lieberman-Aiden E, *et al* (2020) Analysis of Hi-C data using SIP effectively identifies loops in organisms from *C. elegans* to mammals. *Genome Res* 30: 447–458

Rybak-Wolf A, Stottmeister C, Glažar P, Jens M, Pino N, Giusti S, Hanan M, Behm M, Bartok O, Ashwal-Fluss R, *et al* (2015) Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Mol Cell* 58: 870–885

Sabari BR, Dall’Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, *et al* (2018) Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 361

Sadaie M, Salama R, Carroll T, Tomimatsu K, Chandra T, Young ARJ, Narita M, Pérez-Mancera PA, Bennett DC, Chong H, *et al* (2013) Redistribution of the Lamin B1 genomic binding profile affects rearrangement of heterochromatic domains and SAHF formation during senescence. *Genes Dev* 27: 1800–1808

Sage J, Miller AL, Pérez-Mancera PA, Wysocki JM & Jacks T (2003) Acute mutation of retinoblastoma gene function is sufficient for cell cycle re-entry. *Nature* 424: 223–228

Salama R, Sadaie M, Hoare M & Narita M (2014) Cellular senescence and its effector programs. *Genes Dev* 28: 99–114

Saldaña-Meyer R, Rodriguez-Hernaez J, Escobar T, Nishana M, Jácome-López K, Nora EP, Bruneau BG, Tsirigos A, Furlan-Magaril M, Skok J, *et al* (2019) RNA Interactions Are Essential for CTCF-Mediated Genome Organization. *Mol Cell* 76: 412-422.e5

Salminen A, Kauppinen A & Kaarniranta K (2012) Emerging role of NF- κ B signaling in the induction of senescence-associated secretory phenotype (SASP). *Cell Signal* 24: 835–845

Salzman J, Chen RE, Olsen MN, Wang PL & Brown PO (2013) Cell-type specific features of circular RNA expression. *PLoS Genet* 9: e1003777

Salzman J, Gawad C, Wang PL, Lacayo N & Brown PO (2012) Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* 7: e30733

Sanyal A, Lajoie BR, Jain G & Dekker J (2012) The long-range interaction landscape of gene promoters. *Nature* 489: 109–113

Sati S, Bonev B, Szabo Q, Jost D, Bensadoun P, Serra F, Loubiere V, Papadopoulos GL, Rivera-Mulia JC, Fritsch L, *et al* (2020) 4D genome rewiring during oncogene-induced and replicative senescence. *Mol Cell* 78: 522–538

Scaffidi P, Misteli T & Bianchi ME (2002) Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. *Nature* 418: 191–195

Schlackow M, Nojima T, Gomes T, Dhir A, Carmo-Fonseca M & Proudfoot NJ (2017) Distinctive Patterns of Transcription and RNA Processing for Human lincRNAs. *Mol Cell* 65: 25–38

Schneider, Fg W, D K, Aam F, M H, J T, H Y & W W (2021) Liquid-liquid phase separation of light-inducible transcription factors increases transcription activation in mammalian cells and mice. *Sci Adv* 7

Schneider T, Hung L-H, Schreiner S, Starke S, Eckhof H, Rossbach O, Reich S, Medenbach J & Bindereif A (2016) CircRNA-protein complexes: IMP3 protein component defines subfamily of circRNPs. *Sci Rep* 6: 31313

Schoenfelder S & Fraser P (2019) Long-range enhancer–promoter contacts in gene expression control. *Nat Rev Genet* 20: 437–455

Schwarzer W, Abdennur N, Goloborodko A, Pekowska A, Fudenberg G, Loe-Mie Y, Fonseca NA, Huber W, Haering CH, Mirny L, *et al* (2017) Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 551: 51–56

Seif E, Kang JJ, Sasseville C, Senkovich O, Kaltashov A, Boulier EL, Kapur I, Kim CA & Francis NJ (2020) Phase separation by the polyhomeotic sterile alpha motif compartmentalizes Polycomb Group proteins and enhances their activity. *Nature Communications* 11: 5609

Servant N, Lajoie BR, Nora EP, Giorgetti L, Chen C-J, Heard E, Dekker J & Barillot E (2012) HiTC: exploration of high-throughput ‘C’ experiments. *Bioinformatics* 28: 2843–2844

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

Sgarra R, Rustighi A, Tessari MA, Di Bernardo J, Altamura S, Fusco A, Manfioletti G & Giancotti V (2004) Nuclear phosphoproteins HMGA and their relationship with chromatin structure and cancer. *FEBS Lett* 574: 1–8

Shah PP, Donahue G, Otte GL, Capell BC, Nelson DM, Cao K, Aggarwala V, Cruickshanks HA, Rai TS, McBryan T, *et al* (2013) Lamin B1 depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. *Genes Dev* 27: 1787–1799

Shao A -w, Sun H, Geng Y, Peng Q, Wang P, Chen J, Xiong T, Cao R & Tang J (2016) Bclaf1 is an important NF- κ B signaling transducer and C/EBP β regulator in DNA damage-induced senescence. *Cell Death & Differentiation* 23: 865–875

Shen L, Shao N, Liu X & Nestler E (2014) ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* 15: 284

Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B & de Laat W (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat Genet* 38: 1348–1354

Singh J & Dixon GH (1990) High mobility group proteins 1 and 2 function as general class II transcription factors. *Biochemistry* 29: 6295–6302

Singh S, Szlachta K, Manukyan A, Raimer HM, Dinda M, Bekiranov S & Wang YH (2020) Pausing sites of RNA polymerase II on actively transcribed genes are enriched in DNA double-stranded breaks. *J Biol Chem* 295: 3990–4000

Sobuz SU, Sato Y, Yoshizawa T, Karim F, Ono K, Sawa T, Miyamoto Y, Oka M & Yamagata K (2019) SIRT7 regulates the nuclear export of NF- κ B p65 by deacetylating Ran. *Biochim Biophys Acta Mol Cell Res* 1866: 1355–1367

Sparatore B, Passalacqua M, Patrone M, Melloni E & Pontremoli S (1996) Extracellular high-mobility group 1 protein is essential for murine erythroleukaemia cell differentiation. *Biochem J* 320 (Pt 1): 253–256

Stadhouders R, Filion GJ & Graf T (2019) Transcription factors and 3D genome conformation in cell-fate decisions. *Nature* 569: 345–354

- Stadhouders R, Vidal E, Serra F, Di Stefano B, Le Dily F, Quilez J, Gomez A, Collombet S, Berenguer C, Cuartero Y, *et al* (2018) Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat Genet* 50: 238–249
- Stanborough T, Niederhauser J, Koch B, Bergler H & Pertschy B (2014) Ribosomal protein S3 interacts with the NF- κ B inhibitor I κ B α . *FEBS Letters* 588: 659–664
- Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung L-H & Bindereif A (2015) Exon circularization requires canonical splice signals. *Cell Rep* 10: 103–111
- Starkov D, Parfenyev V & Belan S (2021) Conformational statistics of non-equilibrium polymer loops in Rouse model with active loop extrusion. *J Chem Phys* 154: 164106
- Statello L, Guo C-J, Chen L-L & Huarte M (2021) Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol* 22: 96–118
- van Steensel B & Belmont AS (2017) Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. *Cell* 169: 780–791
- van Steensel B & Furlong EEM (2019) The role of transcription in shaping the spatial organization of the genome. *Nat Rev Mol Cell Biol* 20: 327–337
- Sterne-Weiler T, Weatheritt RJ, Best AJ, Ha KCH & Blencowe BJ (2018) Efficient and Accurate Quantitative Profiling of Alternative Splicing Patterns of Any Complexity on a Laptop. *Mol Cell* 72: 187-200.e6
- Stevens TJ, Lando D, Basu S, Atkinson LP, Cao Y, Lee SF, Leeb M, Wohlfahrt KJ, Boucher W, O’Shaughnessy-Kirwan A, *et al* (2017) 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 544: 59–64
- Strom AR, Emelyanov AV, Mir M, Fyodorov DV, Darzacq X & Karpen GH (2017) Phase separation drives heterochromatin domain formation. *Nature* 547: 241–245
- Stros M (1998) DNA bending by the chromosomal protein HMG1 and its high mobility group box domains. Effect of flanking sequences. *J Biol Chem* 273: 10355–10361
- Stros M (2010) HMGB proteins: interactions with DNA and chromatin. *Biochim Biophys Acta* 1799: 101–113
- Stros M, Nishikawa S & Dixon GH (1994) cDNA sequence and structure of a gene encoding trout testis high-mobility-group-1 protein. *Eur J Biochem* 225: 581–591
- Stros M, Ozaki T, Bacikova A, Kageyama H & Nakagawara A (2002) HMGB1 and HMGB2 cell-specifically down-regulate the p53- and p73-dependent sequence-specific transactivation from the human Bax gene promoter. *J Biol Chem* 277: 7157–7164
- Stros M, Polanská E, Struncová S & Pospíšilová S (2009) HMGB1 and HMGB2 proteins up-regulate cellular expression of human topoisomerase II α . *Nucleic Acids Res* 37: 2070–2086
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, *et al* (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA* 102: 15545

Sun C (2020) The SF3b complex: splicing and beyond. *Cell Mol Life Sci* 77: 3583–3595

Sundberg E, Fasth AER, Palmblad K, Harris HE & Andersson U (2009) High mobility group box chromosomal protein 1 acts as a proliferation signal for activated T lymphocytes. *Immunobiology* 214: 303–309

Sutherland H & Bickmore WA (2009) Transcription factories: gene expression in unions? *Nat Rev Genet* 10: 457–466

Swanson EC, Manning B, Zhang H & Lawrence JB (2013) Higher-order unfolding of satellite heterochromatin is a consistent and early event in cell senescence. *J Cell Biol* 203: 929–942

Symmons O, Uslu VV, Tsujimura T, Ruf S, Nassari S, Schwarzer W, Ettwiller L & Spitz F (2014) Functional and topological characteristics of mammalian regulatory domains. *Genome Res* 24: 390–400

Szabo Q, Bantignies F & Cavalli G (2019) Principles of genome folding into topologically associating domains. *Sci Adv* 5: 1668

Szabo Q, Jost D, Chang J-M, Cattoni DI, Papadopoulos GL, Bonev B, Sexton T, Gurgo J, Jacquier C, Nollmann M, *et al* (2018) TADs are 3D structural units of higher-order chromosome organization in Drosophila. *Science Advances* 4: eaar8082

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, *et al* (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47: D607–D613

Tadie J-M, Bae H-B, Deshane JS, Bell CP, Lazarowski ER, Chaplin DD, Thannickal VJ, Abraham E & Zmijewski JW (2012) Toll-like receptor 4 engagement inhibits adenosine 5'-monophosphate-activated protein kinase activation through a high mobility group box 1 protein-dependent mechanism. *Mol Med* 18: 659–668

Tang D, Kang R, Livesey KM, Cheh C-W, Farkas A, Loughran P, Hoppe G, Bianchi ME, Tracey KJ, Zeh HJ, *et al* (2010) Endogenous HMGB1 regulates autophagy. *J Cell Biol* 190: 881–892

Tatavosian R, Kent S, Brown K, Yao T, Duc HN, Huynh TN, Zhen CY, Ma B, Wang H & Ren X (2019) Nuclear condensates of the Polycomb protein chromobox 2 (CBX2) assemble through phase separation. *J Biol Chem* 294: 1451–1463

Teves SS, An L, Bhargava-Shah A, Xie L, Darzacq X & Tjian R (2018) A stable mode of bookmarking by TBP recruits RNA polymerase II to mitotic chromosomes

Teves SS, An L, Hansen AS, Xie L, Darzacq X & Tjian R (2016) A dynamic mode of mitotic bookmarking by transcription factors. *Elife* 5

Thakurela S, Garding A, Jung J, Schübeler D, Burger L & Tiwari VK (2013) Gene regulation and priming by topoisomerase II α in embryonic stem cells. *Nat Commun* 4: 2478

Therizols P, Illingworth RS, Courilleau C, Boyle S, Wood AJ & Bickmore WA (2014) Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. *Science* 346: 1238–1242

Thiecke MJ, Wutz G, Muhar M, Tang W, Bevan S, Malysheva V, Stocsits R, Neumann T, Zuber J, Fraser P, *et al* (2020) Cohesin-dependent and -independent mechanisms mediate chromosomal contacts between promoters and enhancers. *Cell Rep* 32: 107929

- Thomas JO & Stott K (2012) H1 and HMGB1: modulators of chromatin structure. *Biochem Soc Trans* 40: 341–346
- Thomas JO & Travers AA (2001) HMG1 and 2, and related ‘architectural’ DNA-binding proteins. *Trends Biochem Sci* 26: 167–174
- Thuault S, Valcourt U, Petersen M, Manfioletti G, Heldin C-H & Moustakas A (2006) Transforming growth factor-beta employs HMGA2 to elicit epithelial-mesenchymal transition. *J Cell Biol* 174: 175–183
- Tjalsma SJ & Laat W (2020) Novel orthogonal methods to uncover the complexity and diversity of nuclear architecture. *Curr Opin Genet Dev* 67: 10–17
- Tolhuis B, Palstra RJ, Splinter E, Grosveld F & de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10: 1453–1465
- Toma ID, Rossetti G, Zambrano S, Bianchi ME & Agresti A (2014) Nucleosome loss facilitates the chemotactic response of macrophages. *Journal of Internal Medicine* 276: 454–469
- Tominaga-Yamanaka K, Abdelmohsen K, Martindale JL, Yang X, Taub DD & Gorospe M (2012) NF90 coordinately represses the senescence-associated secretory phenotype. *Aging* 4: 695–708
- Trendel J, Schwarzl T, Horos R, Prakash A, Bateman A, Hentze MW & Krijgsveld J (2019) The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell* 176: 391-403.e19
- Tsai A, Alves MR & Crocker J (2019) Multi-enhancer transcriptional hubs confer phenotypic robustness. *eLife* 8: e45325
- Übelmesser N & Papantonis A (2019) Technologies to study spatial genome organization: beyond 3C. *Briefings in Functional Genomics* 18: 395–401
- Ueda T, Chou H, Kawase T, Shirakawa H & Yoshida M (2004) Acidic C-tail of HMGB1 is required for its target binding to nucleosome linker DNA and transcription stimulation. *Biochemistry* 43: 9901–9908
- Ueda T & Yoshida M (2010) HMGB proteins and transcriptional regulation. *Biochim Biophys Acta* 1799: 114–118
- Ugrinova I, Pasheva EA, Armengaud J & Pashev IG (2001) In vivo acetylation of HMG1 protein enhances its binding affinity to distorted DNA structures. *Biochemistry* 40: 14655–14660
- Ulianov SV, Khrameeva EE, Gavrilov AA, Flyamer IM, Kos P, Mikhaleva EA, Penin AA, Logacheva MD, Imakaev MV, Chertovich A, *et al* (2016) Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res* 26: 70–84
- Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigó R & Johnson R (2018) Towards a complete map of the human long non-coding RNA transcriptome. *Nat Rev Genet* 19: 535–548
- Uusküla-Reimand L, Hou H, Samavarchi-Tehrani P, Rudan MV, Liang M, Medina-Rivera A, Mohammed H, Schmidt D, Schwalie P, Young EJ, *et al* (2016) Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biol* 17: 182
- Vagnarelli P (2013) Chromatin reorganization through mitosis. *Adv Protein Chem Struct Biol* 90: 179–224

Vaňková Hausnerová V & Lanctôt C (2017) Chromatin decondensation is accompanied by a transient increase in transcriptional output. *Biol Cell* 109: 65–79

Venø MT, Hansen TB, Venø ST, Clausen BH, Grebing M, Finsen B, Holm IE & Kjems J (2015) Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development. *Genome Biol* 16: 245

Vermunt MW, Zhang D & Blobel GA (2019) The interdependence of gene-regulatory elements and the 3D genome. *J Cell Biol* 218: 12–26

Vian L, Pekowska A, Rao SSP, Kieffer-Kwon KR, Jung S, Baranello L, Huang SC, Khattabi LE, Dose M, Pruett N, *et al* (2018) The energetics and physiological impact of cohesin extrusion. *Cell* 173: 1165–1178

Wan F, Anderson DE, Barnitz RA, Snow A, Bidere N, Zheng L, Hegde V, Lam LT, Staudt LM, Levens D, *et al* (2007) Ribosomal Protein S3: A KH Domain Subunit in NF- κ B Complexes that Mediates Selective Gene Regulation. *Cell* 131: 927–939

Wang JC (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat Rev Mol Cell Biol* 3: 430–440

Wang L, Gao Y, Zheng X, Liu C, Dong S, Li R, Zhang G, Wei Y, Qu H, Li Y, *et al* (2019) Histone Modifications Regulate Chromatin Compartmentalization by Contributing to a Phase Separation Mechanism. *Mol Cell* 76: 646-659.e6

Wang PL, Bao Y, Yee M-C, Barrett SP, Hogan GJ, Olsen MN, Dinneny JR, Brown PO & Salzman J (2014) Circular RNA is expressed across the eukaryotic tree of life. *PLoS One* 9: e90859

Wang Y, Lu JJ, He L & Yu Q (2011) Triptolide (TPL) inhibits global transcription by inducing proteasome-dependent degradation of RNA polymerase II (Pol II). *PLoS One* 6: 23993

Watrin E, Schleiffer A, Tanaka K, Eisenhaber F, Nasmyth K & Peters JM (2006) Human Scc4 is required for cohesin binding to chromatin, sister-chromatid cohesion, and mitotic progression. *Curr Biol* 16: 863–874

Weintraub AS, Li CH, Zamudio AV, Sigova AA, Hannett NM, Day DS, Abraham BJ, Cohen MA, Nabet B, Buckley DL, *et al* (2017) YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* 171: 1573-1588.e28

Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, Celniker SE, Graveley BR & Lai EC (2014) Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. *Cell Rep* 9: 1966–1980

Wheeler JR, Matheny T, Jain S, Abrisch R & Parker R (2016) Distinct stages in stress granule assembly and disassembly. *eLife* 5: e18413

Wijchers PJ, Geeven G, Eyres M, Bergsma AJ, Janssen M, Verstegen M, Zhu Y, Schell Y, Vermeulen C, de Wit E, *et al* (2015) Characterization and dynamics of pericentromere-associated domains in mice. *Genome Res* 25: 958–969

Wijchers PJ, Krijger PHL, Geeven G, Zhu Y, Denker A, Verstegen MJAM, Valdes-Quezada C, Vermeulen C, Janssen M, Teunissen H, *et al* (2016) Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments. *Molecular Cell* 61: 461–473

- Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J & Bähler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453: 1239–1243
- Will CL (2002) Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *The EMBO Journal* 21: 4978–4988
- de Wit E & de Laat W (2012) A decade of 3C technologies: insights into nuclear organization. *Genes Dev* 26: 11–24
- de Wit E, Vos ESM, Holwerda SJB, Valdes-Quezada C, Verstegen MJAM, Teunissen H, Splinter E, Wijchers PJ, Krijger PHL & de Laat W (2015) CTCF Binding Polarity Determines Chromatin Looping. *Mol Cell* 60: 676–684
- Wu C-L, Wang Y, Jin B, Chen H, Xie B-S & Mao Z-B (2015) Senescence-associated Long Non-coding RNA (SALNR) Delays Oncogene-induced Senescence through NF90 Regulation. *J Biol Chem* 290: 30175–30192
- Wu Q, Zhang W, Pwee K-H & Kumar PP (2003) Cloning and characterization of rice HMGB1 gene. *Gene* 312: 103–109
- Wu W, Ji P & Zhao F (2020) CircAtlas: an integrated resource of one million highly accurate circular RNAs from 1070 vertebrate transcriptomes. *Genome Biology* 21: 101
- Wutz G, Ladurner R, Hilaire BGS, Stocsits RR, Nagasaka K, Pignard B, Sanborn A, Tang W, Varnai C, Ivanov MP, *et al* (2020) ESCO1 and CTCF enable formation of long chromatin loops by protecting cohesinSTAG1 from WAPL. *eLife* 9: 52091
- Xia P, Wang S, Ye B, Du Y, Li C, Xiong Z, Qu Y & Fan Z (2018) A Circular RNA Protects Dormant Hematopoietic Stem Cells from DNA Sensor cGAS-Mediated Exhaustion. *Immunity* 48: 688-701.e7
- Xia S, Feng J, Lei L, Hu J, Xia L, Wang J, Xiang Y, Liu L, Zhong S, Han L, *et al* (2017) Comprehensive characterization of tissue-specific circular RNAs in the human and mouse genomes. *Brief Bioinform* 18: 984–992
- Xiang J-F, Yin Q-F, Chen T, Zhang Y, Zhang X-O, Wu Z, Zhang S, Wang H-B, Ge J, Lu X, *et al* (2014) Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res* 24: 513–531
- Xie W, Kagiampakis I, Pan L, Zhang YW, Murphy L, Tao Y, Kong X, Kang B, Xia L, Carvalho FLF, *et al* (2018) DNA Methylation Patterns Separate Senescence from Transformation Potential and Indicate Cancer Risk. *Cancer Cell* 33: 309-321.e5
- Xue W, Zender L, Miething C, Dickins RA, Hernando E, Krizhanovsky V, Cordon-Cardo C & Lowe SW (2007) Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* 445: 656–660
- Yamazaki T, Souquere S, Chujo T, Kobelke S, Chong YS, Fox AH, Bond CS, Nakagawa S, Pierron G & Hirose T (2018) Functional Domains of NEAT1 Architectural lncRNA Induce Paraspeckle Assembly through Phase Separation. *Mol Cell* 70: 1038-1053.e7

- Yang D, Postnikov YV, Li Y, Tewary P, de la Rosa G, Wei F, Klinman D, Gioannini T, Weiss JP, Furusawa T, *et al* (2012) High-mobility group nucleosome-binding protein 1 acts as an alarmin and is critical for lipopolysaccharide-induced immune responses. *J Exp Med* 209: 157–171
- Yang L, Duff MO, Graveley BR, Carmichael GG & Chen L-L (2011) Genomewide characterization of non-polyadenylated RNAs. *Genome Biology* 12: R16
- Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen L-L, Wang Y, *et al* (2017) Extensive translation of circular RNAs driven by N6-methyladenosine. *Cell Res* 27: 626–641
- Yang Y, Gao X, Zhang M, Yan S, Sun C, Xiao F, Huang N, Yang X, Zhao K, Zhou H, *et al* (2018) Novel Role of FBXW7 Circular RNA in Repressing Glioma Tumorigenesis. *JNCI: Journal of the National Cancer Institute* 110: 304–315
- Yesbolatova A, Natsume T, Hayashi KI & Kanemaki MT (2019) Generation of conditional auxin-inducible degron (AID) cells and tight control of degron-fused proteins using the degradation inhibitor auxinole. *Methods* 164–165: 73–80
- Yin S, Gambe RG, Sun J, Martinez AZ, Cartun ZJ, Regis FFD, Wan Y, Fan J, Brooks AN, Herman SEM, *et al* (2019) A Murine Model of Chronic Lymphocytic Leukemia Based on B Cell-Restricted Expression of Sf3b1 Mutation and Atm Deletion. *Cancer Cell* 35: 283-296.e5
- Yin Y, Lu JY, Zhang X, Shao W, Xu Y, Li P, Hong Y, Cui L, Shan G, Tian B, *et al* (2020) U1 snRNP regulates chromatin retention of noncoding RNAs. *Nature* 580: 147–150
- Yoshida M, Makiguchi K, Chida Y & Shimura K (1984) Unwinding of DNA by nonhistone protein HMG1 and HMG2. *Nucleic Acids Symp Ser*: 181–184
- Yu G, Wang L-G & He Q-Y (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* 31: 2382–2383
- Yumoto M, Nishida O, Moriyama K, Shimomura Y, Nakamura T, Kuriyama N, Hara Y & Yamada S (2011) In vitro evaluation of high mobility group box 1 protein removal with various membranes for continuous hemofiltration. *Ther Apher Dial* 15: 385–393
- Yusufzai TM, Tagami H, Nakatani Y & Felsenfeld G (2004) CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* 13: 291–298
- Zagelbaum J, Shimazaki N, Esguerra ZA, Watanabe G, Lieber MR & Rothenberg E (2016) Real-time analysis of RAG complex activity in V(D)J recombination. *Proc Natl Acad Sci U S A* 113: 11853–11858
- Zeng K, Chen X, Xu M, Liu X, Hu X, Xu T, Sun H, Pan Y, He B & Wang S (2018) CircHIPK3 promotes colorectal cancer growth and metastasis by sponging miR-7. *Cell Death Dis* 9: 417
- Zeng Y, Du WW, Wu Y, Yang Z, Awan FM, Li X, Yang W, Zhang C, Yang Q, Yee A, *et al* (2017) A Circular RNA Binds To and Activates AKT Phosphorylation and Nuclear Localization Reducing Apoptosis and Enhancing Cardiac Repair. *Theranostics* 7: 3842–3855
- Zetterström CK, Strand M-L & Söder O (2006) The high-mobility group box chromosomal protein 1 is expressed in the human and rat testis where it may function as an antibacterial factor. *Hum Reprod* 21: 2801–2809

Zhang H, Emerson DJ, Gilgenast TG, Titus KR, Lan Y, Huang P, Zhang D, Wang H, Keller CA, Giardine B, *et al* (2019) Chromatin structure dynamics during the mitosis-to-G1 phase transition. *Nature* 576: 158–162

Zhang M, Zhao K, Xu X, Yang Y, Yan S, Wei P, Liu H, Xu J, Xiao F, Zhou H, *et al* (2018) A peptide encoded by circular form of LINC-PINT suppresses oncogenic transcriptional elongation in glioblastoma. *Nat Commun* 9: 4475

Zhang X-O, Wang H-B, Zhang Y, Lu X, Chen L-L & Yang L (2014) Complementary Sequence-Mediated Exon Circularization. *Cell* 159: 134–147

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, *et al* (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biology* 9: R137

Zhang Y, Zhang X-O, Chen T, Xiang J-F, Yin Q-F, Xing Y-H, Zhu S, Yang L & Chen L-L (2013) Circular Intronic Long Noncoding RNAs. *Molecular Cell* 51: 792–806

Zhao J, Sun BK, Erwin JA, Song J-J & Lee JT (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322: 750–756

Zhao Y, Alexandrov PN, Jaber V & Lukiw WJ (2016) Deficiency in the Ubiquitin Conjugating Enzyme UBE2A in Alzheimer's Disease (AD) is Linked to Deficits in a Natural Circular miRNA-7 Sponge (circRNA; ciRS-7). *Genes (Basel)* 7

Zheng Q, Bao C, Guo W, Li S, Chen J, Chen B, Luo Y, Lyu D, Li Y, Shi G, *et al* (2016) Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat Commun* 7: 11215

Zheng X, Chen L, Zhou Y, Wang Q, Zheng Z, Xu B, Wu C, Zhou Q, Hu W, Wu C, *et al* (2019) A novel protein encoded by a circular RNA circPPP1R12A promotes tumor pathogenesis and metastasis of colon cancer via Hippo-YAP signaling. *Molecular Cancer* 18: 47

Zhou C, Liu H, Wang F, Hu T, Liang Z, Lan N, He X, Zheng X, Wu X, Xie D, *et al* (2020) circCAMSAP1 Promotes Tumor Growth in Colorectal Cancer via the miR-328-5p/E2F1 Axis. *Molecular Therapy* 28: 914–928

Zhou KI, Shi H, Lyu R, Wylder AC, Matuszek Ż, Pan JN, He C, Parisien M & Pan T (2019a) Regulation of co-transcriptional pre-mRNA splicing by m6A through the low-complexity protein hnRNPG. *Mol Cell* 76: 70-81.e9

Zhou X, Benson KF, Ashar HR & Chada K (1995) Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature* 376: 771–774

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C & Chanda SK (2019b) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications* 10: 1523

Zhou Z, Sun B, Huang S & Zhao L (2019c) Roles of circular RNAs in immune regulation and autoimmune diseases. *Cell Death Dis* 10: 503

Zhu Y, Denholtz M, Lu H & Murre C (2021) Calcium signaling instructs NIPBL recruitment at active enhancers and promoters via distinct mechanisms to reconstruct genome compartmentalization. *Genes Dev* 35: 65–81

Zirkel A, Nikolic M, Sofiadis K, Mallm J-P, Brackley CA, Gothe H, Drechsel O, Becker C, Altmüller J, Josipovic N, *et al* (2018) HMGB2 Loss upon Senescence Entry Disrupts Genomic Organization and Induces CTCF Clustering across Cell Types. *Mol Cell* 70: 730-744.e6

Zirkel A & Papantonis A (2018) Detecting Circular RNAs by RNA Fluorescence In Situ Hybridization. In *Circular RNAs*, Dieterich C & Papantonis A (eds) pp 69–75. New York, NY: Springer New York

Zuin J, Franke V, IJcken WFJ, Sloot A, Krantz ID, Reijden MIJA, Nakato R, Lenhard B & Wendt KS (2014) A cohesin-independent role for NIPBL at promoters provides insights in CdLS. *PLoS Genet* 10: 1004153

Nataša Josipović

Address: Robert-Koch-Str. 40, 37075 Göttingen

Email: natasajosipovic0@gmail.com

Phone: +49 (0) 1525 4838229

Education

- July 2016 – present **PhD candidate** at International Max Planck Research School for Genome Science, Georg-August-Universität, Göttingen, Germany
Former affiliation: Center for Molecular Medicine Cologne, University of Cologne, Germany
“Chromatin reorganization, signaling responses and non-coding RNA regulators in homeostasis and senescence”
- 2015-2016 MSc in Molecular biology, module Neurobiology
Faculty of Biology, University of Belgrade, Serbia
“Changes in expression and activity of ecto-5'-nucleotidase in rat spinalcord during the experimental autoimmune encephalomyelitis”
- 2011-2015 BSc in Biology, module Molecular biology
Faculty of Biology, University of Belgrade, Serbia
-

Publications

1. Sofiadis K[#], **Josipovic N^{**}**, Nikolic M[#], Kargapolova Y, Zirkel A, Papadakis A, Papadionysiou I, Loughran G, Keanes J, Michel A, et al (2020) HMGB1 coordinates SASP-related chromatin folding and RNA homeostasis on the path to senescence. bioRxiv: 540146 *revised manuscript in print*
2. Zhang S[#], Übelmesser N[#], **Josipovic N^{**}**, Forte G, Slotman JA, Chiang M, Gothe H, Gusmao EG, Becker C, Altmüller J, et al (2020) RNA polymerase II is required for spatial chromatin reorganization following exit from mitosis. bioRxiv: 2020.10.27.356915 *under revision*
3. Kargapolova Y, Rehim R, Kayserili H, Brühl J, Sofiadis K, Zirkel A, Palikyras S, Mizi A, Li Y, Yigit G, et al (2021) Overarching control of autophagy and DNA damage response by CHD6 revealed by modeling a rare human pathology. Nat Commun 12: 3014
4. Wallis R, **Josipovic N**, Mizen H, Robles-Tenorio A, Tyler EJ, Papantonis A & Bishop CL (2021) Isolation methodology is essential to the evaluation of the extracellular vesicle component of the senescence-associated secretory phenotype. J Extracell Vesicles 10: e12041
5. Casa V, Moronta Gines M, Gade Gusmao E, Slotman JA, Zirkel A, **Josipovic N**, Oole E, van IJcken WFJ, Houtsmuller AB, Papantonis A, et al (2020) Redundant and specific roles of cohesin STAG subunits in chromatin looping and transcriptional control. Genome Res 30: 515–527
6. Goranci-Buzhala G, Mariappan A, Ricci-Vitiani L, **Josipovic N**, Pacion S, Gottardo M, Ptok J, Callaini G, Rajalingam K, Dynlacht B, et al (2020) Cilium induction triggers differentiation of glioma stem cells. bioRxiv: 2020.12.23.424140

7. Gothe HJ, Bouwman BAM, Gusmao EG, Piccinno R, Petrosino G, Sayols S, Drechsel O, Minneker V, **Josipovic N**, Mizi A, et al (2019) Spatial Chromosome Folding and Active Transcription Drive DNA Fragility and Formation of Oncogenic MLL Translocations. *Mol Cell* 75: 267-283.e12
8. Lalioti M-E, Arbi M, Loukas I, Kaplani K, Kalogeropoulou A, Lokka G, Kyrousi C, Mizi A, Georgomanolis T, **Josipovic N**, et al (2019) GemC1 governs multiciliogenesis through direct interaction with and transcriptional regulation of p73. *J Cell Sci* 132
9. Zirkel A, Nikolic M, Sofiadis K, Mallm J-P, Brackley CA, Gothe H, Drechsel O, Becker C, Altmüller J, **Josipovic N**, et al (2018) HMGB2 Loss upon Senescence Entry Disrupts Genomic Organization and Induces CTCF Clustering across Cell Types. *Mol Cell* 70: 730-744.e6
10. Adzic M, Stevanovic I, **Josipovic N**, Laketa D, Lavrnja I, Bjelobaba IM, Bozic I, Jovanovic M, Milosevic M & Nedeljkovic N (2017) Extracellular ATP induces graded reactive response of astrocytes and strengthens their antioxidative defense in vitro. *J Neurosci Res* 95: 1053–1066

Conferences and meetings

- 2019 RNA control and regulation conference in Cold Springs Harbor Laboratory, USA
 “Regulating the anti-inflammatory potency of primary human cells via two circRNAs”, poster presentation
- 2018 Ernst Klenk Symposium in Molecular Medicine, Germany
- 2018 Cologne Spring meeting, "RNA: beyond its genetic code", Germany
- 2015 Belgrade International Molecular Life Science Conference for Students, Serbia
 “Astrocytes activate antioxidant defense system in response to extracellular ATP *in vitro*”, poster presentation

Workshops

- 2019 GGNB Professional Skills Course: Drug Discovery and Project Management in the Pharmaceutical Industry
- 2018 Crispr/Cas9 genome editing – CEPLAS
- 2017 Bioinformatics (Bioinf), Prof. Dr. Röbbbe Wünschiers
- 2016 Advanced scientific writing: Writing Publishable Research Articles, BioScript

Languages

English	Advanced proficiency
German	Limited proficiency
Serbian	Native proficiency