Article

HMGB1 as a rheostat of chromatin topology and RNA homeostasis on the path to senescence

Konstantinos Sofiadis [1,2,§], Milos Nikolic[1,§], Anne Zirkel[1,§], Yulia Kargapolova[1,#], Natasa Josipovic[1,2,#], Antonis Papadakis[3], Eduardo G. Gusmao[1,2], Athanasia Mizi[1,2], Theodore Georgomanolis[1], Mirjam Koker[1,4], Roland T. Ullrich[1,4], Janine Altmüller[5], Peter Nürnberg[5], Andreas Beyer[3] & Argyris Papantonis[1,2,*]

[1]*Center for Molecular Medicine Cologne, University of Cologne, 50931 Cologne, Germany*

[2]*Institute of Pathology, University Medical Center Göttingen, 37075 Göttingen, Germany*

[3]*Cluster of Excellence Cologne for Age-associated Disease, University of Cologne, 50931 Cologne, Germany*

[4]*Clinic I of Internal Medicine and Center for Integrated Oncology, University Hospital Cologne, 50931 Cologne, Germany*

[5]*Cologne Center for Genomics, University of Cologne, 50931 Cologne, Germany*

[§,#]These authors contributed equally to this work

*Corresponding author; Tel.: +49 551 39 65734, Email: argyris.papantonis@med.uni-goettingen.de

Abstract

Spatial organization and gene expression of mammalian chromosomes are maintained and regulated in conjunction with cell cycle progression. This link is perturbed once cells enter senescence. The highly abundant HMGB1 protein, known to associate with bent and looped DNA, is actively depleted from senescent cell nuclei to act as an extracellular proinflammatory stimulus. Despite its physiological importance, we still lack understanding of the positioning and functional roles of HMGB1 on chromatin *in vivo*. To address this, we mapped HMGB1 binding genome-wide in different primary cells using a tailored protocol. We then integrated ChIP-seq and Hi-C data with a knot theory approach to uncover HMGB1 demarcation at the boundaries of particular topologically-associating domains (TADs). These TADs harbor genes involved in the key proinflammatory leg of the senescent transcriptional program. Moreover, we used sCLIP and siRNA-mediated knockdown to show that HMGB1 is a *bona fide* RNA-binding protein also affecting splicing choices. Together, our findings highlight a broader than hitherto assumed role for HMGB1 in chromatin homeostasis connected to cell cycle potency, and allow us to postulate a "rheostat" model for HMGB function with implications in cancer biology.

Keywords: 3D chromatin folding; replicative senescence; chromatin loop; RNA binding; splicing; single cell analysis; inflammation; senescence-associated secretory phenotype

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 (Štros, 2010). HMGB1, together with histone H1, are the next most abundant proteins after core histones, with one HMGB1 molecule for every ~10 nucleosomes (Thomas and Stott, 2012). Despite its documented abundance in cell nuclei, HMGB1 (also known as an "alarmin") has been predominantly studied *in vivo* as an extracellular signaling molecule (Lohani and Rajeswari, 2016; Bianchi et al, 2017).

HMGB1 is actively secreted by activated monocytes and macrophages to signal tissue damage, and passively released by necrotic and damaged cells. Once received by cells, HMGB1 molecules can be recognized with high affinity by RAGE receptors to potently signal inflammation (Scaffidi et al, 2002; Bonaldi et al, 2003; Orlova et al, 2007). In cells entering senescence, HMGB1 relocalizes from the nucleus to the cytoplasm and is then secreted to stimulate NF-κB activity via Toll-like receptor signaling. This relocalization and secretion controls and contributes to the senescence-associated secretory phenotype (SASP) of mammalian cells, representing a major paracrine contributor both *in vitro* and *in vivo* (Salminen et al, 2012; Acosta et al, 2013; Davalos et al, 2013).

Within proliferating cell nuclei, HMGB1 is thought to bind DNA in a nonspecific manner via its two HMGB-box domains, and to bend and contort the double helix – a mode of action that facilitates recruitment of various DNA-binding factors, like p53 (Štros et al, 2004; Štros, 2010; Rowell et al, 2012). HMGB1 has been studied in detail for its contribution to DNA repair (Ito et al, 2015; Mukherjee et al, 2016), V(D)J recombination (Little et al, 2013; Zagelbaum et al, 2016) or chromatin assembly (Bonaldi et al, 2002), but not so much for its transcriptional role (Calogero et al, 2009). Notably, cells lacking HMGB1 also 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; De Toma et al, 2014).

HMGB1 associates with its cognate DNA sites with characteristically high "on"/"off" rates, and its acidic tail is important for stabilizing binding (Pallier et al, 2003; Ueda et al, 2004; Štros, 2010; Blair et al, 2016). However, it is now understood that HMG-box DNA-binding domains are particularly sensitive to commonly-used fixatives, like formaldehyde; thus, capturing them on chromatin can be challenging (Pallier et al, 2003; Teves et al, 2016). As a result, there exist no genome-wide datasets describing HMGB binding repertoires in primary mammalian cells (see http://chip-atlas.org), and our appreciation of their on-chromatin roles remains vague. To address this, we employed here a tailored dual-crosslinking approach previously used to map HMGB2 binding genome-wide (Zirkel et al, 2018). We can now show that HMGB1 binding in proliferating primary endothelial and lung fibroblast cells is far from nonspecific, while also disparate to that by HMGB2. Once integrated with whole-genome chromosome conformation capture (Hi-C) data from proliferating and senescent cells, HMGB1-bound positions demarcate the boundaries of a considerable and specific fraction of topologically-associating domains (TADs; Dixon et al, 2012; Nora et al, 2017). This topological contribution is eliminated upon senescence entry, and knockdown experiments show that HMGB1 directly controls the expression of genes central to the senescent transcriptional program. Critically, as HMGB1 was proposed to also constitute a *bona fide* RNA-binding protein (Castello et al, 2016), we used sCLIP (Kargapolova et al, 2017) to show it does also influence RNA splicing. In summary, we use senescence as a model to comprehensively characterize the multifaceted *in nucleo* roles of HMGB1 and show how these converge to underpin transition into replicative arrest by linking chromatin homeostasis with cell cycle progression – a link with far-reaching implications for the attenuation of proliferation of cancer cells overexpressing *HMGB1*.

Results

HMGB1 nuclear loss marks senescence entry coinciding with chromatin changes

To investigate the *in nucleo* 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 (**Figure 1A**). Next, we used RNAseq data from proliferating and senescent HUVEC and IMR90 (from two different biological donors/isolates) to look into changing mRNA levels of chromatin-associated factors. Changes that were convergent between the two cell types involved genes encoding lamin B1, various histones, centrosomal proteins, cohesin and condensin complexes, as well as HMG family proteins, most of which were consistently suppressed upon senescence entry, also at the protein level (**Figure 1B,C**; Davalos et al, 2013; Shah et al, 2013; Rai et al, 2014; Zirkel et al, 2018). Here, we chose to focus on HMGB1 due to its high nuclear abundance (Thomas and Stott, 2012; **Figure 1D**) and key role in SASP induction (Davalos et al, 2013), but mostly due to its still elusive on-chromatin functions, especially as regards spatial chromatin folding.

HMGB1 immunodetection in early- and late-passage cells documented a >50% decrease in its total nuclear levels in senescence-entry cell populations of HUVECs or IMR90 (**Figure 1E**). However, HMGB1 nuclear depletion was most dramatic in the enlarged nuclei of senescent cells 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 than the average nucleus of a proliferating cell; **Figure S1A**). This showed that, within such heterogeneous populations, enlarged nuclei almost invariably represent senescent cells lacking HMGB1 and concomitantly exhibiting a drop in H3K27me3 levels that mark facultative heterochromatin – effects which would otherwise be masked (**Figure 1C** and **S1B,C**). Last, we showed that it is those larger cells that secrete HMGB1, but not HMGB2, into their growth medium to contribute to the SASP (**Figure S1D**).

Regulated *HMGB1* nascent transcription in single cells

Since senescence entry is idiosyncratic to each individual cell and occurs asynchronously in a given cell population, it is important to obtain a single-cell understanding of the transcriptional changes linked to the nuclear loss of HMGB1. To this end, we developed a new protocol for single-cell sequencing of nascent RNA. Typically, scRNA-seq relies on capturing and barcoding cellular mRNAs via reverse-transcription of their long poly(A) tails (See et al, 2018). To capture nascent RNA, we decided to add poly(A) tails to any RNA associated with an active site of transcription. We employed the previously-introduced "factory" RNA-seq protocol (Caudron-Herger et al, 2015) as a basis for isolating intact nuclei rich in nascent transcripts from both proliferating and senescent HUVEC. We next removed most chromatin not attached to active transcription sites, and polyadenylated nascent transcripts *in situ*, before a standard library preparation via a 10X Genomics platform (overview in **Figure S2A**). Using this new approach, and despite low rates of non-duplet capturing (see **Methods** for more details), we obtained >550 single nuclei with an average of 1,650 transcripts in each nucleus, >55% of which represented introns or intron-exon junctions. Unsupervised t-SNE clustering of single-nucleus transcriptomes grouped all senescent cells in a single cluster, while subdividing the larger proliferating HUVEC population into three clusters (**Figure S2B**). Reassuringly, genes differentially-expressed amongst proliferating and senescent clusters associated with GO terms relevant to senescence entry and SASP induction (**Figure S2C**).

Mapping nascent RNA levels of individual gene loci onto those clusters, showed that *HMGB1* and *HMGB2* are actively repressed not only in senescent cells, but also in numerous proliferating cells; the senescence-induced *HMGA1/A2* loci strongly produce nascent RNA in senescent cells, but are markedly upregulated also in proliferating cells showing lower *HMGB1/B2* transcription (e.g., in cells of the *blue* cluster; **Figure S2B**). As expected, SASP-related genes like *IL4R* and *MMP14* show strong transcription in cells repressing *HMGB1/B2*, while the p21-encoding locus, *CDKN1A*, has most nascent RNA detected in the senescent cluster (**Figure S2D**). Notably, correlating *HMGB1* and *HMGB2* levels in senescent cells confirms that switching off the latter locus is an early event on the path to senescence (as recently proposed; Zirkel et al, 2018), and this trend is already obvious in proliferating cells (**Figure S2E**). Similarly, correlating *HMGB1* levels with those of *CDKN1A*, *HMGA1* or *IL4R*, all of which are activated upon senescence entry, we observed that their hyperactivation mostly occurs in those cells where *HMGB1* nascent RNA production is diminished (**Figure S2E**).

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

Capturing HMGB proteins bound to chromatin has proven challenging, because their tandem HMG-box DNA-binding domains are not compatible with standard formaldehyde fixation (Pallier et al, 2003; Belmont et al, 2015; Teves et al, 2016). Here, we employed a tailored fixation strategy in ChIP assays to efficiently capture HMGB1 bound to its cognate sites genome-wide in both proliferating HUVECs and IMR90 (**Figures 2A** and **S3A**; see **Methods** for details). HMGB1 ChIP-seq peaks were predominantly found at the promoters and gene bodies of active genes (>75% and >65% of 810 and 593 peaks in HUVECs and IMR90, respectively; **Figure 2B**), overlapping regions marked by H3K27ac and oftentimes H2A.Z (**Figure 2C**). HMGB1 peaks were often found clustering along chromosomes of proliferating cells (**Figure S3B**), but as HMGB1 mostly binds active loci, and HUVECs have a gene expression program disparate to that of IMR90, the two repertoires only share 40 peaks (**Figure S3C**). It is noteworthy that none of these peaks mark SASP-related genes, despite HMGB1's role in proinflammatory stimulation.

De novo motif discovery in accessible "footprints" within DNase I-hypersensitive sites (derived from ENCODE data) overlapping HMGB1 peaks revealed rather specific CG-rich motifs, which do not however converge into a single consensus sequence (**Figure S3D**). We also surveyed these same accessible footprints for known transcription factor (TF) motifs to infer possible co-binding complexes or complexes that might replace HMGB1 on senescent chromatin. We found that HMGB1 binding sites are significantly enriched for E2F-family motifs, which are notably down-regulated upon senescence entry. Also, motifs for senescence-activated corepressors (e.g., REST and HEY2) and for the architectural ZBTB7B protein that is important for inflammatory gene activation (Nikopoulou et al, 2018) are also enriched therein (**Figure S3E**). Last, we combined RNA-seq with HMGB1 ChIP-seq data to catalogue differentially-expressed genes in either cell type that are directly regulated upon HMGB1 loss in senescence. In both HUVECs and IMR90 we find >2-fold more up- rather than downregulated genes bound by HMGB1 (**Figure S3F**). These senescence-induced genes are involved in ECM organization, regulation of apoptosis, as well as with NOTCH-/TGFβ-signalling that is now understood to represent a second "wave" of paracrine signalling in senescent cells (Hoare et al, 2017). On the other hand, downregulated genes in both cell types converge to cell growth and cell cycle regulation (**Figure S3F**). Together, this data demonstrates how HMGB1 nuclear eviction regulates a critical leg of the program defining senescence entry.

HMGB1 demarcates a particular subset of senescence-regulated TAD boundaries

The boundaries of TADs represent genomic sites of strong local insulation of spatial interactions, and are often marked by the presence of 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 this demarcation is lost upon senescence entry leading to the formation of senescence-induced CTCF foci (Zirkel et al, 2018). Based on this, we reasoned that HMGB1 might also function similarly Hi-C data from proliferating and senescent HUVEC to investigate this. First, we found that >20% of the 810 HMGB1 peaks reside at TAD boundaries (called at 25-kbp resolution; **Figure 2D,E**). Remarkably, and much unlike HMGB2, we identified TAD boundaries marked by HMGB1, but lacking CTCF (and the converse; **Figure 2E**). As expected, TADs of senescent nuclei lose this demarcation (**Figure 2F**), and HMGB1-marked boundaries exhibit an obvious loss of insulation and rearrangement of spatial interactions (**Figure 2G**). Next, we grouped TADs according to whether or not their boundaries change upon senescence entry. We found that the ~1,000 TADs remaining invariant in senescence, boundaries are more enriched for CTCF rather than HMGB1 binding. CTCF and HMGB1 demarcation is comparable in TADs shifting one boundary or collapsing into a larger TAD (~5,500 TADs in total; **Figure 2H**). Thus, we postulate that the loss of HMGB1 in senescence contributes to the remodeling of TADs. Last, we asked whether HMGB1 engages in long-range loop formation; the 810 HMGB1-bound genomic sites give rise to >600 interchromosomal loops, which appear to collapse upon senescence entry (**Figure 2I**).

Spatial TAD co-association reveals functional chromosome compartmentalization

It is now understood that human chromosomes undergo large-scale changes upon replicative arrest (Zirkel et al, 2018), which are accentuated in “deep” senescence (Criscione et al, 2016). Looking at TADs as building blocks of chromosomes (Sexton and Cavalli, 2015), one can ask how these might spatialuly associate to give rise to “meta-TADs” (Fraser et al, 2015) and ultimately to diverse functional topologies. To address this question, we employed a bias-free and unsupervised approach based on “topologically-intrinsic lexicographic ordering” (TiLO; Johnson, 2014), whereby TADs are treated as nodes in a clustered spatial network tested for robustness via iterative network slicing (see **Figure 3A** and **Methods** for details). By applying this to TADs derived from proliferating and senescent HUVEC Hi-C data, we found that there is a marked increase in clusters that include multiple consecutive TADs in senescence, indicative of TAD merging (**Figure 3B,C**). This is consistent with the observation of general spatial chromatin compaction in senescence (Criscione et al, 2016; Zirkel et al, 2018) and with the fact that ~75% of TADs merge in senescence (**Figure 2G**). However, we also identified individual (“singular”) TADs typically positioned between multi-TAD clusters (**Figure 3C**, *arrows*). Strikingly, the boundaries of singular TADs were strongly marked by HMGB1, compared to the extremities of clusters comprised of >3 consecutive TADs (**Figure 3D**). To assess the functional impact of this, we investigated expression level changes of genes embedded in different cluster types. Both singular TADs and TADs with >3 consecutive TADs harbor genes significantly more upregulated that those in all other TADs, but singular TADs also show significantly more modest gene downregulation (**Figure 3E**). Notably, these two subsets harbor most of the genes differentially-regulated upon senescence entry. What further discriminates these clusters functionally, is the fact that multi-TAD clusters uniquely harbor genes involved in the regulation of chromatin organization and conformation, while those unique to singular TADs associate with SASP production and its downstream effects (**Figure 3F**). Thus, TiLO identifies functional entities that involve spatial co-association of multiple TADs in *cis*, explaining the functional specialization of different genomic segments upon HMGB1 loss.

HMGB1 directly binds and regulates senescence-relevant transcripts

On top of its known ability to bind chromatin, HMGB1 was recently also classified as a *bona fide* RNA-binding protein (Castello et al, 2016; Trendel et al, 2018). Thus, its loss in senescence could also affect some aspect of RNA metabolism. Our suspicion was reinforced by cataloguing the proteins that co-immunoprecipitate with HMGB1 from proliferating IMR90 using mass spectrometry. This returned a diverse range of RNA-binding proteins and splicing regulators, in addition to the expected proteins involved in the regulation of chromatin conformation (**Figure 4A**).

To study HMGB1 as a direct RNA binder and regulator, we performed sCLIP (Kargapolova et al, 2017) in proliferating IMR90 in two well correlated biological replicates (**Figure S4A-C**). These two datasets provided a compendium of 1,773 binding peaks on 866 different transcripts (e.g., *ASH1L* and *CCNL2*; **Figure 4B** and **Table S1**), which reassuringly display <25% overlap to HMGB1-bound genes in ChIP-seq. On RNA, HMGB1 was mostly found bound to exons and 5'/3' UTRs, but also to a substantial number of non-coding RNAs (**Figure 4C**). These HMGB1 RNA-binding sites present a hexameric 5'-NMWGRA-3' (M=A/C, W=A/T, R=A/G) motif, irrespective of the predicted folding of the underlying RNA or the exclusion of repeat sequences from the genome build used for motif analysis (**Figures 4D** and **S4D**). Much like the trend observed in ChIP-seq data, HMGB1 was found bound to ~3-fold more transcripts that are up- rather than downregulated upon senescence (**Figure S4E**, *left*). Upregulated transcripts associated with senescence-relevant GO terms like "ECM organization", "wound healing" and "negative regulation of cell proliferation", while downregulated ones mostly converged on processes like RNA splicing, RNA-/miRNA-mediated gene silencing, or histone remodeling and deacetylation (**Figure S4E**, *right*). This suggests a feedback loop whereby transcripts encoding RNA regulators are affected by the senescence-induced loss of HMGB1, but also pointed to a more direct impact on splicing regulation.

We next examined how splicing patterns are altered upon senescence entry by IMR90 using Whippet (Sterne-Weiler et al, 2018). We documented significant changes involving ~4,000 transcripts, the majority of which concerned the use of alternative transcription start and polyadenylation sites (>80% of cases; **Figure 4E**). Interestingly, looking at the splicing changes involving 342 HMGB1-bound mRNAs, this trend remained invariant (**Figure 4E**). Finally, we asked whether transcripts undergoing different types of splicing changes are involved in different functional processes. GO term analysis revealed that transcripts with alternative TSS usage encoded regulators of ECM organization and growth control, whereas those with alternative poly(A)-site usage encoded mostly splicing and RNA processing factors; HMGB1-bound transcripts, however, associated with functions from the whole of this spectrum (**Figure 4F**). Thus, the nuclear loss of HMGB1 also directly affects processing of the cell's transcriptome.

HMGB1 depletion underlies induction of the senescence transcriptional program

It was previously demonstrated that transduction of human fibroblasts with shRNAs against *HMGB1* sufficed to induce senescence (Davalos et al, 2013). To avoid using lentiviral vectors, we first treated HUVECs with a pool of self-delivering siRNAs targeting *HMGB1*. This led to a ~2-fold reduction of HMGB1 protein and RNA levels within 72 h (**Figure S5A**), was accompanied by a doubling of β-gal and p21-positive cells in the knockdown population, but only small changes in nuclear size and BrdU incorporation were recorded (**Figure S5B-E**). We subsequently turned to IMR90, where standard siRNA transfections allowed for a >10-fold decrease in *HMGB1* levels, while also suppressing or inducing expression of known senescence-regulated genes (but only marginally affecting *HMGB2*; **Figure 5A**). "Deep" sequencing and analysis of mRNAs from *HMGB1*-knockdown and control IMR90 revealed ~900 up- and ~950 downregulated genes (**Figure 5B**). GO term and gene set enrichment analyses showed that genes that were upregulated associated with the SASP and proinflammatory signalling, while those that were downregulated with changes in chromatin organization, transcriptional silencing, and the p53 pathway (**Figure 5C,D**).

For a more precise understanding for the role of HMGB1 in these processes, we focused on genes differentially-regulated upon *HMGB1*-knockdown that also carried at least one HMGB1 ChIP-seq peak. This returned 44 up- and 56 downregulated genes constituting direct HMGB1 targets linked to NF-κB activation, and to chromatin organization or non-inflammatory signalling, respectively (**Figure S5F**). Nonetheless, this left >850 genes in either category that could not be directly linked to an HMGB1 chromatin-binding event. To bridge this gap, we repeated the above analysis using sCLIP binding events. We found 56 and 97 HMGB1-bound mRNAs up- and downregulated upon *HMGB1*-knockdown, respectively. Consistent to our previous observations, those upregulated could be linked to processes like ECM organization and proinflammatory activation, while downregulated ones to non-inflammatory signalling and the organization of chromatin (**Figure S5G**). Notably, splicing changes deemed significant upon *HMGB1*-KD displayed an trend in favour of alternative TSS and poly(A)-site usage as, essentially identical to that observed in senescence, with >60% of senescence splicing events also being recorded in knockdown cells (**Figure 5E**). These mRNAs encode factors linked to senescence-regulated processes like cell cycle and cell growth regulation or the p53 pathway (**Figure 5F**). Taken together, our data are in support of a model whereby HMGB1 acts both at chromatin loci and on RNA transcripts to regulate cellular functions, and this interplay is disrupted in senescent cells by its nuclear depletion.

An HMGB1-dependency for lung cancer cell proliferation

The nuclear depletion of HMGB1 and its ensuing transcriptional changes are senescence hallmarks and represent an inherent antitumorigenic mechanism (Rai and Adams, 2013). However, *HMGB1* is very often overexpressed in patient tumor specimens (**Figure 6A**), and it has been suggested that its increased titers might result in either the enhancement of cell survival or the regulation of cell death via diverse pathways that include inflammation, cell proliferation, and autophagy (Nagatani et al, 2001; Kang et al, 2013). We focused on non-small cell lung adenocarcinoma patient-derived lines, because high *HMGB1* expression therein is one predictor of poor patient survival (**Figure 6B**). We used three different lines, H1975, A549, and H1650, and found that cell proliferation and survival was significantly impaired via *HMGB1*-knockdown in 72 h in all three (**Figure 6C**). Interestingly, the extent of proliferative impediment in each line was almost proportional to its proliferation rate, with the faster H1975 cells arresting completely and the slower H1650 still exhibiting some population increase upon knockdown (**Figure 6C,D**). Nonetheless, *HMGB1*-knockdown in these lines showed convergent changes in the expression of senescence markers like *HMGB2*, *LMNB1*, *EZH2* or *SMC2* (**Figure 6E**), which are actually overexpressed in correlation to *HMGB1* levels in lung adenocarcinoma tumours (data from the TCGA cohort; **Figure 6F**). Since knocking down *HMGB1* in this context also decreased *HMGB2* levels, and we recently showed that reduced HMGB2 levels result in the formation of senescence-induced CTCF clusters (SICCs), we examined whether SICCs formed differentially in these lines. Indeed, we saw that the most affected H1975 cells exhibited larger and more prominent SICCs upon *HMGB1*-knockdown, compared to the least affected H1650 cells that displayed essentially no SICCs; the intermediately-affected A549 presented SICCs in both control and knockdown cells (**Figure S6A-C**). Thus, differences in proliferative capacity and SICC emergence correlate with the anti-proliferative extent of *HMGB1* depletion in these cancer lines, highlighting how targeting HMGB1 might need to be considered in therapeutic interventions for cancer treatment.

Discussion

Unlike its well documented proinflammatory role, the intracellular positioning and gene expression control exerted by HMGB1 on mammalian chromosomes is poorly understood. Here, we were able to assign a multifaceted role to HMGB1 — as an on-chromatin regulator of active gene loci, and as a *bona fide* RNA-binding protein recognizing a distinct subset of transcripts. As a result, we can deduce the following. First, we that the loss of HMGB1 from the nuclei of senescent cells mostly triggers upregulation of its previously-bound target loci and mRNAs, suggesting that HMGB1 tends to act a "buffering" factor thereon. Second, that ~1/5 of HMGB1-bound positions mark the boundary of a TAD, and many of these TADs specifically harbor SASP-related genes induced upon senescence entry. Last, that the loss of HMGB1 correlates essentially only with alternative usage of TSS and polyadenylation sites in its bound transcripts. These observations come to substantiate a previous hypothesis that low nuclear titers of HMGB1 are necessary for the fully-fledged development of the SASP (Davalos et al, 2013). This is because there is a need for alleviating the regulatory effects that HMGB1 exerts on active promoters, on mRNAs being processed, as well as on TAD boundaries. This is a rather unique example of a regulatory circuit, where the programmed deregulation in one cellular compartment (the nucleus) is in direct and quantitative crosstalk with the signaling deployed in another (in extracellular space to initiate paracrine activation). Thus, the senescent regulatory program has a strong, if not hierarchical, dependency on the nuclear events preceding SASP deployment.

Recently, we characterized the function of the sister protein to HMGB1, HMGB2, for the entry into replicative senescence (Zirkel et al, 2018). The loss of HMGB2 is an event preceding the loss of HMGB1, and leads to the formation of large senescence-induced CTCF clusters (SICCs). This has an apparent effect on the spatial architecture of chromosomes, and concomitantly on gene expression. Intriguingly, direct HMGB2 target loci are also typically upregulated when 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 immediate senescence entry. Also, HMGB1 and HMGB2 bind non-overlapping targets and also demarcate TADs in two distinct modes – HMGB2 marks the extremities of TADs that shift one boundary in senescence, while HMGB1 is mostly found at the boundaries of TADs that collapse together, in line with the overall compaction observed in senescent chromosomes (Criscione et al, 2016; Zirkel et al, 2018). Finally, it is important to note that the *HMGB1* knockdown does not reduce *HMGB2* levels in primary human cells, nor does the converse hold true, meaning that the two pathways these related factors control do not really cross one another, but are rather deployed in parallel.

Interestingly, the knockdown of *HMGB1* in primary lung fibroblasts exhibits differential gene expression patterns that are partially inversed in the same cells upon senescence entry (e.g., the negative regulation of RNAPII transcription is suppressed in the knockdown, but not in senescence). This hints towards a coordinated counter-regulation of HMGB1 effects by the rest of the program of cells entering senescence. This 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* (and often also *HMGB2*) in various cancer types (Li et al, 2014; Zirkel et al, 2018). This overexpression seems to favor increased cell proliferation (Li et al, 2014), and it is only reasonable to assume that its targeting might be an effective anti-cancer strategy. Here, we used three lung adenocarcinoma lines to show that indeed a simple siRNA-mediated inhibition of *HMGB1* suffices for replicative arrest and cell death. However, we observed that the response of each of these three lines correlated inversely to their proliferation rates, suggesting that higher rates come with stronger addiction to HMGB1 presence. Moreover, the formation (or strengthening) of SICCs in these cancer cells also aligned well with their response to *HMGB1* knockdown —i.e., obvious SICC emergence signified replicative arrest.

In summary, the above allows to propose a simple model by which HMGB1 titers can be seen as a “rheostat” of cell cycle potency of a given cell. Primary proliferating cells maintain normal nuclear HMGB1 levels, cells entering senescence arrest upon nuclear depletion of HMGB1, while aberrantly proliferating cancer cells actively overexpress HMGB1 and are addicted to it for propagation (**Figure 6G**). Thus, in a next step, elucidating the exact molecular dependencies of cancer cells to HMGB1 (and probably also HMGB2) nuclear overrepresentation for proliferation may lead to new ideas for combinatorial cancer treatments.

Materials and Methods

Primary cell culture and senescence markers

HUVECs from single, apparently healthy, donors (passage 2; 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/cm^2, except for late passages when they were seeded at ~20,000 cells/cm^2. 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 ml fresh medium plus 10 ml of a 12 mM MTT stock solution (Invitrogen), and cells are incubated at 37°C for 4 h. Subsequently, all but 25 mL of the medium is removed from the wells, and formazan dissolved in 50 mL 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. DNA methylation at six selected CpG islands (Franzen et al, 2017) was measured by isolating genomic DNA at different passages and performing targeted pyrosequencing (Cygenia GmbH). Finally, nascent DNA synthesis was monitored by EdU incorporation and subsequent labelling with Click-iT chemistry (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 MiToBo (version 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).

Protein extraction and western blotting

For assessing protein abundance at different passages, ~4 x 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 x g to pellet cell debris and collect the supernatant. The concentration of the 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. Resolved proteins were detected using the antisera mentioned above, plus a mouse monoclonal anti-H3K9me3 (1:200; Active motif 39286).

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 mg of a monoclonal HMGB1 antiserum (Developmental Studies Hybridoma Bank; PCRP-HMGB1-4F10) to ~30 mg 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 mapped to the reference human genome (hg19) using BWA (Li and Durbin, 2010), and the resulting .BAM files were processed using Picard tools (http://broadinstitute.github.io/picard/) before MACS2 software (Zhang et al, 2008) was used to identify signal enrichment over input. Thresholded HMGB1 ChIP-seq peaks per each cell type were annotated using Chipseeker (Yu et al, 2015) and are listed in **Table S2A**; .BAM files were used in *ngs.plot* (Shen et al, 2014) 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).

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

Hi-C data from proliferating and senescent HUVEC 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 20-kbp. For plotting insulation heatmaps and "loop-o-grams", normalized interactions values in the twenty 20-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 are available on request; HMGB1-anchored loops are listed in **Table S2B**.

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, 2014). In brief, topologically-intrinsic ordering was used to permutate the linear order of TADs (as a starting organization level in the Hi-C matrices) until a certain "robustly irreducible" topological condition is satisfied. Then, the "pinch ratio" algorithm is used (Heisterkamp and Johnson, 2013) is applied to heuristically slice the network at connections between TADs were local interaction minima are, 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.

Single-cell nascent RNA sequencing and analysis

Proliferating (p. 4) and senescent HUVEC (p. 16) were washed once in an isotonic near-physiological buffer (PB) that maintains the cells' transcriptional activity and subjected immediately to the first steps of the "factory RNA-seq" protocol (Melnik et al, 2016). In more detail, cell nuclei are gently isolated using PB+0.4% NP-40, DNase I-treated at 33oC for 25 min to detach most chromatin, pelleted and washed once in ice-cold PB, before polyadenylation of nascent RNA as described (Kargapolova et al, 2017). Next, ~2,500 cells from each state were loaded onto the Chromium 10X Genomics platform for encapsulation in oil droplets and generation of barcoded cDNA libraries from individual nuclei as per manufacturer's instructions. Despite the documented 0.8% chance of capturing a cell duplet on this platform, HUVEC nuclei are particularly prone to aggregation. As a result, 494 proliferating and 129 senescent single nuclei were efficiently captured and processed. Following sequencing on a HiSeq4000 platform (Illumina), and mapping to the reference genome (hg38) using STAR (Dobin et al, 2013) and filtering via UMI-tools (Smith et al, 2017), ~45,000 and ~60,000 reads were generated per each proliferating or senescent cell, respectively. Poor quality cells were excluded (i.e. cells with <300 or >5,000 expressed genes), as were genes expressed <10 cells. This returned 1,650 robustly captured transcripts per cell on average, with >55% of reads mapping to introns or exon-intron junctions, 575 genes being expressed in at least 25% of all cells, and with the 50 most-expressed genes taking over ~22% of all sequencing reads. Mapped and filtered data were then processed and visualized using a compilation of ZINB-WaVE (ver. 1.3.4; Risso et al, 2018) and Seurat (ver. 2.3.4; Butler et al, 2018) for clustering and t-SNE visualization. ZINB-WaVE was used to create a low-dimensional representation of the gene expression matrix, and factors inferred in the ZINB-WaVE model were added as one of the low-dimensional data representations in the Seurat object in order to identify cell subpopulations via a shared nearest neighbour (SNN) modularity optimization-based clustering algorithm as applied in the *FindClusters* function of Seurat. Visualization was performed using the t-SNE map by applying the *Rtsne* function on the ZINB-WaVE output. This map was integrated into the Seurat object and used to plot gene expression. For differential gene expression analysis between the clusters, we applied a combination of ZINB-WaVE and DESeq2 (Love et al, 2014), where the posterior probabilities that counts were generated from the negative binomial count component of the ZINB-WaVE model were used as observation-level weights in the estimation of regression parameters in DESeq2 (Van den Berge et al, 2018). Differentially-expressed genes identified via this method were filtered using a threshold of $\log_2$FC>±1 plus *P*-value<0.05 and are listed in **Table S3A**.

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; PCRP-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, ver. 78) and annotated using HOMER (http://homer.ucsd.edu), while Gene Ontology analysis was performed using Metascape (www.metascape.org). Finally, the final merged peak list was use for *de novo* motif analysis using ssHMM (Heller et al, 2017) and significantly enriched motifs were next compared to existing RBP motifs to predict proteins potentially recognizing similar sequences using Tomtom (http://meme-suite.org/tools/tomtom). All HMGB1-bound mRNAs are listed in **Table S1**.

siRNA-mediated *HMGB1* knockdown

HUVECs were seeded at ~20,000 cells/cm^2 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 the knockdown in lung adenocarcinoma lines and IMR90 cells, standard siRNA transfections were carried out using RNAiMAX (Invitrogen) as previously described (Zirkel et al, 2018).

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 selected as differentially-expressed and are listed in **Table S3B**. 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/pathway enrichment bar plots using Metascape (http://metascape.org/gp/index.html; Tripathi et al, 2015).

Statistical tests

P-values associated with Student's t-tests and Fischer's exact tests were calculated using GraphPad (https://graphpad.com/), and those associated with the Wilcoxon-Mann-Whitney test using the U test interface (https://www.socscistatistics.com/tests/mannwhitney/Default2.aspx). Unless otherwise stated, *P*-values <0.01 were deemed as statistically significant.

Data availability

The sequencing data generated in this study can be found in the NCBI Gene Expression Omnibus (GEO) repository under the accession number GSE98448, except HMGB1 sCLIP data that will be found under a GSE# accession number upon manuscript acceptance.

Supplementary Material contains Figures S1-S6 and Tables S1-S3.

Author contributions

SK, AZ, NS, NJ, YK, TG, and AM performed experiments; MN, NJ, YK, and EGG performed computational analyses; AP and AB analyzed single-cell RNA-seq data; MK and RTU provided cancer cell lines and performed knockdowns; JA and PN provided library preps and sequencing; AP conceived the study and wrote the manuscript with input from all coauthors.

Conflicts of interest

The authors declare that they have no conflict of interest.

Acknowledgements

We would like to thank members of all laboratories involved in this study for helpful discussions, and Leo Kurian for critical reading of this manuscript. We thank the CMMC FACS-sorting and CECAD Proteomics facilities for assistance. Work in the lab of AP was supported by CMMC Junior Research Group VIII core funding, the DFG (PA2456/4-1 and /5-1), and by an Else-Kroener-Fresenius-Stiftung "Key project" grant (2015_A125).

References

Acosta JC, Banito A, Wuestefeld T, Georgilis A, Janich P, Morton JP, Athineos D, Kang TW, Lasitschka F, Andrulis M, Pascual G, Morris KJ, Khan S, Jin H, Dharmalingam G, Snijders AP, Carroll T, Capper D, Pritchard C, Inman GJ, Longerich T, Sansom OJ, Benitah SA, Zender L, Gil J. A complex secretory program orchestrated by the inflammasome controls paracrine senescence. Nat Cell Biol. 2013 15:978-90.

Bianchi ME, Crippa MP, Manfredi AA, Mezzapelle R, Rovere Querini P, Venereau E. High-mobility group box 1 protein orchestrates responses to tissue damage via inflammation, innate and adaptive immunity, and tissue repair. Immunol Rev. 2017 280:74-82.

Blair RH, Horn AE, Pazhani Y, Grado L, Goodrich JA, Kugel JF. The HMGB1 C-terminal tail regulates DNA bending. J Mol Biol. 2016 428:4060-72.

Bonaldi T, Längst G, Strohner R, Becker PB, Bianchi ME. The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. EMBO J. 2002 21:6865-73.

Bonaldi T, Talamo F, Scaffidi P, Ferrera D, Porto A, Bachi A, Rubartelli A, Agresti A, Bianchi ME. Monocytic cells hyperacetylate chromatin protein HMGB1 to redirect it towards secretion. EMBO J. 2003 22:5551-60.

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018 36:411-20.

Calogero S, Grassi F, Aguzzi A, Voigtländer T, Ferrier P, Ferrari S, Bianchi ME. The lack of chromosomal protein Hmg1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. Nat Genet. 1999 22:276-80.

Castello A, Fischer B, Frese CK, Horos R, Alleaume AM, Foehr S, Curk T, Krijgsveld J, Hentze MW. Comprehensive identification of RNA-binding domains in human cells. Mol Cell. 2016 63:696-710.

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

Celona B, Weiner A, Di Felice F, Mancuso FM, Cesarini E, Rossi RL, Gregory L, Baban D, Rossetti G, Grianti P, Pagani M, Bonaldi T, Ragoussis J, Friedman N, Camilloni G, Bianchi ME, Agresti A. Substantial histone reduction modulates genome wide nucleosomal occupancy and global transcriptional output. PLoS Biol. 2011 9:e1001086.

Davalos AR, Kawahara M, Malhotra GK, Schaum N, Huang J, Ved U, Beausejour CM, Coppe JP, Rodier F, Campisi J. p53-dependent release of Alarmin HMGB1 is a central mediator of senescent phenotypes. J Cell Biol. 2013 201:613-29.

De Toma I, Rossetti G, Zambrano S, Bianchi ME, Agresti A. Nucleosome loss facilitates the chemotactic response of macrophages. J Intern Med. 2014 276:454-69.

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

El Gazzar M, Yoza BK, Chen X, Garcia BA, Young NL, McCall CE. Chromatin-specific remodeling by HMGB1 and linker histone H1 silences proinflammatory genes during endotoxin tolerance. Mol Cell Biol. 2009 29:1959-71.

Fraser J, Ferrai C, Chiariello AM, Schueler M, Rito T, Laudanno G, Barbieri M, Moore BL, Kraemer DC, Aitken S, Xie SQ, Morris KJ, Itoh M, Kawaji H, Jaeger I, Hayashizaki Y, Carninci P, Forrest AR; FANTOM

Consortium, Semple CA, Dostie J, Pombo A, Nicodemi M. Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation. Mol Syst Biol. 2015 11:852.

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

Heisterkamp D, Johnson J. Pinch ratio clustering from a topologically intrinsic lexicographic ordering. Proc SIAM Intl Conf Data Mining. 2013 560-8.

Heller D, Krestel R, Ohler U, Vingron M, Marsico A. ssHMM: extracting intuitive sequence-structure motifs from high-throughput RNA-binding protein data. Nucleic Acids Res. 2017 45:11004-18.

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

Johnson J. Topological graph clustering with thin position. Geometr Dedicata. 2014 169:165-73.

Kang R, Zhang Q, Zeh HJ 3rd, Lotze MT, Tang D. HMGB1 in cancer: good, bad, or both? Clin Cancer Res. 2013 19:4046-57.

Kargapolova Y, Levin M, Lackner K, Danckwardt S. 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. 2017 45:6074-6086.

Li Q, Li J, Wen T, Zeng W, Peng C, Yan S, Tan J, Yang K, Liu S, Guo A, Zhang C, Su J, Jiang M, Liu Z, Zhou H, Chen X. Overexpression of HMGB1 in melanoma predicts patient survival and suppression of HMGB1 induces cell cycle arrest and senescence in association with p21 up-regulation via a p53-independent, Sp1-dependent pathway. Oncotarget. 2014 5:6387-403.

Liao Y, Smyth GK, Shi W. featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014 30:923-30.

Little AJ, Corbett E, Ortega F, Schatz DG. Cooperative recruitment of HMGB1 during V(D)J recombination through interactions with RAG1 and DNA. Nucleic Acids Res. 2013 41:3289-301.

Lohani N, Rajeswari MR. Dichotomous life of DNA binding High Mobility Group Box 1 protein in human health and disease. Curr Protein Pept Sci. 2016 17:762-75.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014 15:550.

Melnik S, Caudron-Herger M, Brant L, Carr IM, Rippe K, Cook PR, Papantonis A. Isolation of the protein and RNA content of active sites of transcription from mammalian cells. Nat Protoc. 2016 11:553-65.

Nagatani G, Nomoto M, Takano H, Ise T, Kato K, Imamura T, Izumi H, Makishima K, Kohno K. Transcriptional activation of the human HMG1 gene in cisplatin-resistant human cancer cells. Cancer Res. 2001 61:1592-7.

Nikopoulou C, Panagopoulos G, Sianidis G, Psarra E, Ford E, Thanos D. The transcription factor ThPOK orchestrates stochastic interchromosomal interactions required for IFNB1 virus-inducible gene expression. Mol Cell. 2018 71:352-61.

Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG. Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. Cell. 2017 169:930-44.

Orlova VV, Choi EY, Xie C, Chavakis E, Bierhaus A, Ihanus E, Ballantyne CM, Gahmberg CG, Bianchi ME, Nawroth PP, Chavakis T. A novel pathway of HMGB1-mediated inflammatory cell recruitment that requires Mac-1-integrin. EMBO J. 2007 26:1129-39.

Pallier C, Scaffidi P, Chopineau-Proust S, Agresti A, Nordmann P, Bianchi ME, Marechal V. Association of chromatin proteins high mobility group box (HMGB) 1 and HMGB2 with mitotic chromosomes. Mol Biol Cell. 2003 14:3414-26.

Pellegrini L, Xue J, Larson D, Pastorino S, Jube S, Forest KH, Saad-Jube ZS, Napolitano A, Pagano I, Negi VS, Bianchi ME, Morris P, Pass HI, Gaudino G, Carbone M, Yang H. HMGB1 targeting by ethyl pyruvate suppresses malignant phenotype of human mesothelioma. Oncotarget. 2017 8:22649-61.

Rai TS, Adams PD. Lessons from senescence: chromatin maintenance innon-proliferating cells. Biochim Biophys Acta. 2013 1819:322-31.

Redmond AM, Byrne C, Bane FT, Brown GD, Tibbitts P, O'Brien K, Hill AD, Carroll JS, Young LS. Genomic interaction between ER and HMGB2 identifies DDX18 as a novel driver of endocrine resistance in breast cancer cells. Oncogene. 2015 34:3871-80.

Risso D, Ngai J, Speed T, Dudoit S. Normalization of RNA-seq data using factor analysis of control genes or samples. Nat Biotech. 2014 32:896-902.

Risso D, Perraudeau F, Gribkova S, Dudoit S, Vert JP. A general and flexible method for signal extraction from single-cell RNA-seq data. Nat Commun. 2018 9:284.

Rowell JP, Simpson KL, Stott K, Watson M, Thomas JO. HMGB1-facilitated p53 DNA binding occurs via HMG-Box/p53 transactivation domain interaction, regulated by the acidic tail. Structure. 2012 20:2014-24.

Salminen A, Kauppinen A, Kaarniranta K. Emerging role of NF-κB signaling in the induction of senescence-associated secretory phenotype (SASP). Cell Signal. 2012 24:835-45.

Scaffidi P, Misteli T, Bianchi ME. Release of chromatin protein HMGB1 by necrotic cells triggers inflammation. Nature. 2002 418:191-5.

See P, Lum J, Chen J, Ginhoux F. A Single-Cell Sequencing guide for immunologists. Front Immunol. 2018 9:2425.

Sexton T, Cavalli G. The role of chromosome domains in shaping the functional genome. Cell. 2015 160:1049-59.

Shah PP, Donahue G, Otte GL, Capell BC, Nelson DM, Cao K, Aggarwala V,Cruickshanks HA, Rai TS, McBryan T, Gregory BD, Adams PD, Berger SL. Lamin B1depletion in senescent cells triggers large-scale changes in gene expression and the chromatin landscape. Genes Dev. 2013 27:1787-99.

Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. 2017 27:491-9.

Sterne-Weiler T, Weatheritt RJ, Best AJ, Ha KCH, Blencowe BJ. Efficient and accurate quantitative profiling of alternative splicing patterns of any complexity on a laptop. Mol Cell. 2018 72:1097-2765.

Štros M, Muselíková-Polanská E, Pospísilová S, Strauss F. High-affinity binding of tumor-suppressor protein p53 and HMGB1 to hemicatenated DNA loops. Biochemistry. 2004 43:7215-25.

Štros M. HMGB proteins: interactions with DNA and chromatin. Biochim Biophys Acta. 2010 1799:101-13.

Tang D, Kang R, Livesey KM, Cheh CW, Farkas A, Loughran P, Hoppe G, Bianchi ME, Tracey KJ, Zeh HJ 3rd, Lotze MT. Endogenous HMGB1 regulates autophagy. J Cell Biol. 2010 190:881-92.

Thomas JO, Stott K. H1 and HMGB1: modulators of chromatin structure. Biochem Soc Trans. 2012 40:341-6.

Trendel J, Schwarzl T, Horos R, Prakash A, Bateman A, Hentze MW, Krijgsveld J. The human RNA-binding proteome and Its dynamics during translational arrest. Cell. 2018 doi:10.1016/j.cell.2018.11.004.

Tripathi S, Pohl MO, Zhou Y, Rodriguez-Frandsen A, Wang G, Stein DA, Moulton HM, DeJesus P, Che J, Mulder LC, Yángüez E, Andenmatten D, Pache L, Manicassamy B, Albrecht RA, Gonzalez MG, Nguyen Q, Brass A, Elledge S, White M, Shapira S, Hacohen N, Karlas A, Meyer TF, Shales M, Gatorano A, Johnson JR, Jang G, Johnson T, Verschueren E, Sanders D, Krogan N, Shaw M, König R, Stertz S, García-Sastre A, Chanda SK. Meta- and orthogonal integration of influenza "OMICs" data defines a role for UBR4 in virus budding. Cell Host Microbe. 2015 18:723-35.

Ueda T, Chou H, Kawase T, Shirakawa H, Yoshida M. Acidic C-tail of HMGB1 is required for its target binding to nucleosome linker DNA and transcription stimulation. Biochemistry. 2004 43:9901-8.

Van den Berge K, Perraudeau F, Soneson C, Love MI, Risso D, Vert JP, Robinson MD, Dudoit S, Clement L. Observation weights unlock bulk RNA-seq tools for zero inflation and single-cell applications. Genome Biol. 2018 19:24.

Yu G, Wang L, He Q. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015 31:2382-83.

Zagelbaum J, Shimazaki N, Esguerra ZA, Watanabe G, Lieber MR, Rothenberg E. Real-time analysis of RAG complex activity in V(D)J recombination. Proc Natl Acad Sci U S A. 2016 113:11853-8.

Zirkel A, Nikolic M, Sofiadis K, Mallm JP, Brackley CA, Gothe H, Drechsel O, Becker C, Altmüller J, Josipovic N, Georgomanolis T, Brant L, Franzen J, Koker M, Gusmao EG, Costa IG, Ullrich RT, Wagner W, Roukos V, Nürnberg P, Marenduzzo D, Rippe K, Papantonis A. HMGB2 loss upon senescence entry disrupts genomic organization and induces CTCF clustering across cell types. Mol Cell. 2018 70:730-44.

Main figures

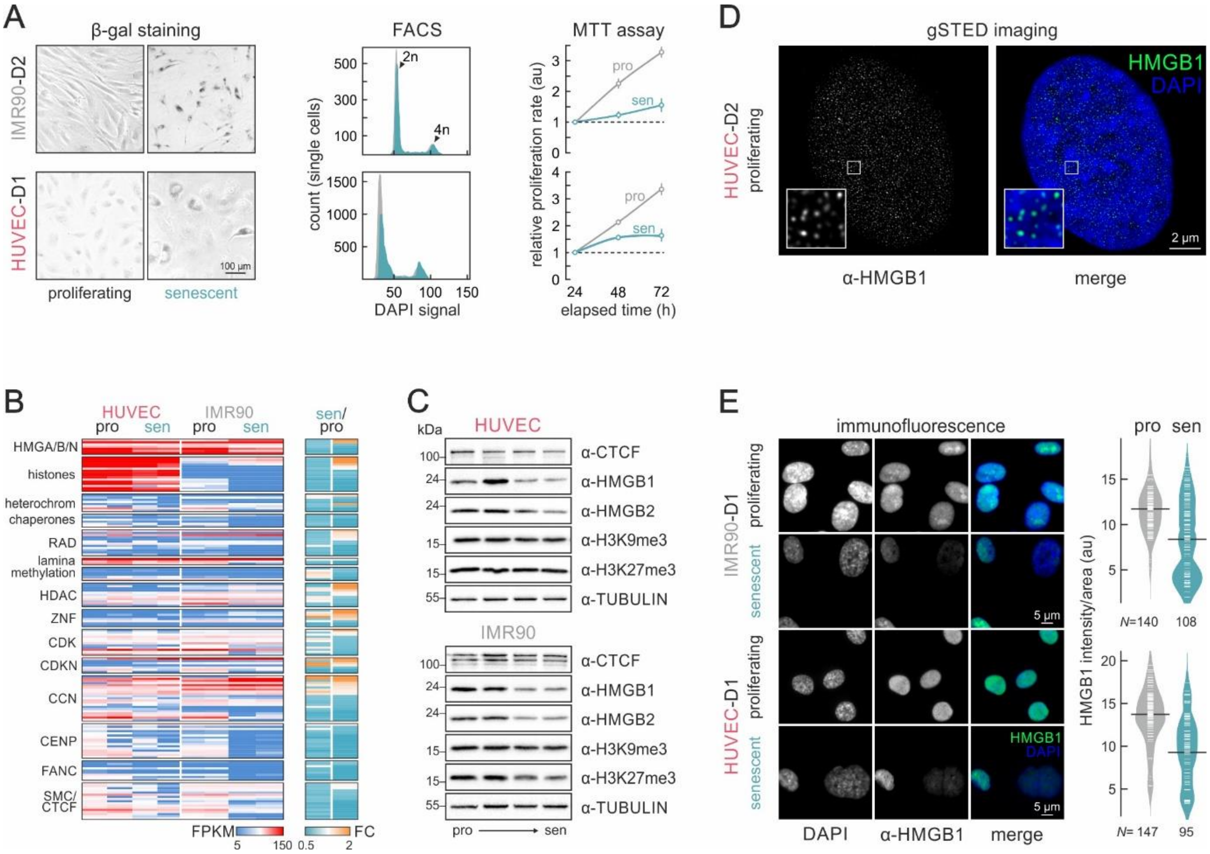
A
β-gal staining
IMR90-D2
HUVEC-D1
100 µm
proliferating
senescent
FACS
2n
4n
500
300
100
1500
1000
500
count (single cells)
50
100
150
DAPI signal
MTT assay
pro
sen
relative proliferation rate (au)
24
48
72
elapsed time (h)
D
gSTED imaging
HUVEC-D2
proliferating
HMGB1
DAPI
2 µm
α-HMGB1
merge
B
HUVEC
pro
sen
IMR90
pro
sen
sen/
pro
HMGA/B/N
histones
heterochrom
chaperones
RAD
lamina
methylation
HDAC
ZNF
CDK
CDKN
CCN
CENP
FANC
SMC/
CTCF
FPKM
5
150
0.5
2
FC
C
kDa
HUVEC
100
24
24
15
15
55
α-CTCF
α-HMGB1
α-HMGB2
α-H3K9me3
α-H3K27me3
α-TUBULIN
IMR90
pro
sen
E
immunofluorescence
IMR90-D1
HUVEC-D1
proliferating
senescent
5 µm
HMGB1
DAPI
DAPI
α-HMGB1
merge
pro
sen
HMGB1 intensity/area (au)
15
10
5
N=140
108
20
N= 147
95

Figure 1. Senescence entry affects chromatin-binding factors in primary human cells.

(A) Proliferating and senescent IMR90 and HUVECs were tested for β-galactosidase activity (*left*), cell cycle progression via FACS (*middle*), and proliferation rates via MTT assays (*right*).

(B) Heatmaps show gene expression levels (FPKM; *left*) and fold-change upon senescence ($\log_2$; *right*) of selected genes encoding chromatin-associated factors.

(C) Western blots show changing protein levels in IMR90 and HUVECs on the path to senescence.

(D) Super-resolution (gSTED) imaging of HMGB1 distribution in proliferating HUVEC nuclei. Bar: 2µm.

(E) Representative immunofluorescence images of IMR90 and HUVECs show reduced HMGB1 levels in enlarged senescent nuclei (*left*), and bean plots quantify this reduction (*right*; *N* indicates the number of cells analyzed per each condition/cell type). Bars: 5 µm.

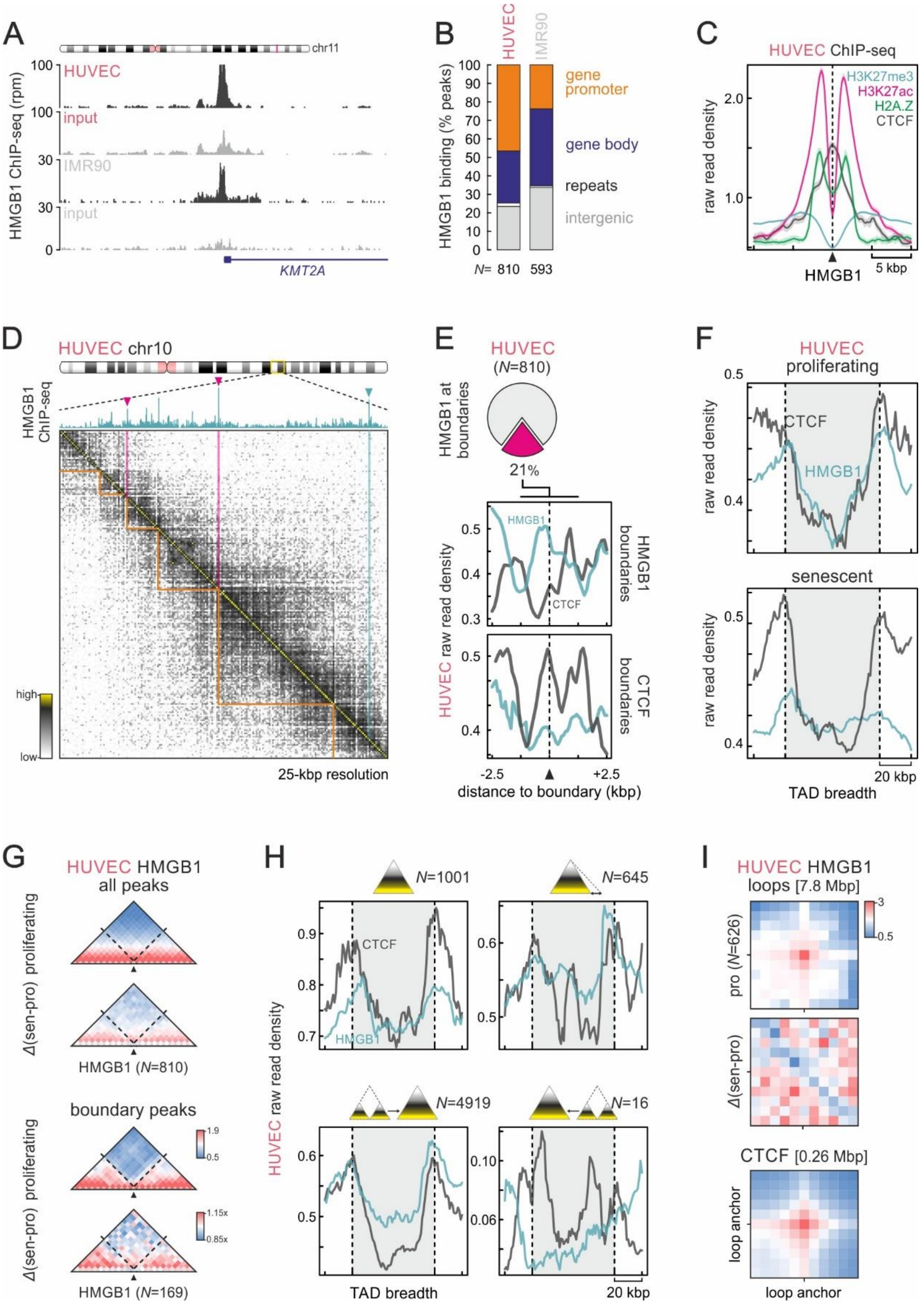

Figure 2. HMGB1 binds active genes and demarcates a subset of TAD boundaries.

(A) Genome browser views showing HMGB1 ChIP-seq signal at the *KMT2A* promoter in HUVEC and IMR90 (*black*); input tracks (*grey*) provide a negative control.

(B) Bar graphs show the genomic distribution of HMGB1 binding peaks in HUVEC and IMR90. The number of peaks (*N*) analyzed per each cell type is indicated.

(C) Line plots show the distribution of H3K27ac (*magenta*), H3K27me3 (*light blue*), H2A.Z (*green*), and CTCF ChIP-seq signal (*grey*) in the 10 kbp around all HUVEC HMGB1 peaks.

(D) Exemplary Hi-C heatmap for a sub region of HUVEC chromosome 10 aligned to HMGB1 ChIP-seq shows signal peaks (*magenta*) at TAD boundaries (*orange triangles*).

(E) Pie chart (*top*) shows the fraction of HUVEC HMGB1 peaks residing at TAD boundaries. Line plots (*below*) show the mean distribution of HMGB1 (*green*) and CTCF ChIP-seq signal (*grey*) in the 5 kbp around HMGB1- or CTCF-marked boundaries.

(F) Line plots show the mean distribution of HMGB1 (*green*) and CTCF ChIP-seq signal (*grey*) along all TADs ±20 kbp from proliferating or senescent HUVECs (*left*) and IMR90 (*right*).

(G) Heatmaps show mean Hi-C signal from proliferating (*top row*) or senescent/proliferating HUVEC (*bottom row*) in the 0.4 Mbp around all HMGB1 peaks (*left*) or those at TAD boundaries (*right*). *N* indicates the number of peaks in each subgroup.

(H) As in panel C, but for TADs that remain unchanged (*top left*), shift one boundary (*top right*), merge together (*bottom left*), or break up upon senescence entry (*bottom right*). *N* indicates the number of TADs in each subgroup.

(I) Heatmaps show focal Hi-C interactions between HMGB1-bound peaks from proliferating HUVEC (*left*) and their loss upon senescence entry (*middle*); HUVEC CTCF loops provide a positive control (*right*). Median sizes of CTCF and HMGB1 loops are shown in square brackets.

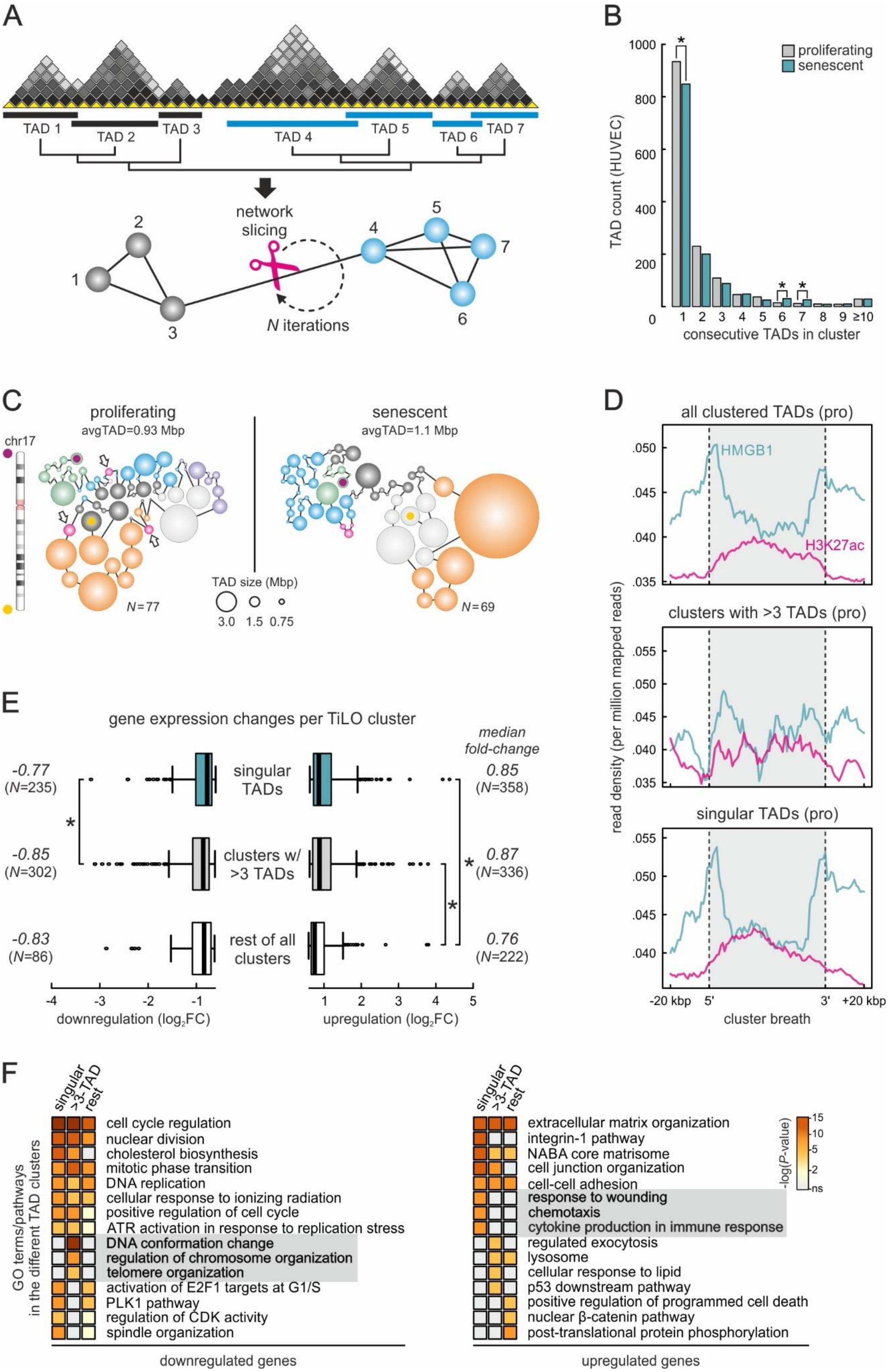
A
TAD 1
TAD 2
TAD 3
TAD 4
TAD 5
TAD 6
TAD 7
network slicing
N iterations
1
2
3
4
5
6
7
B
1000
800
600
400
200
0
TAD count (HUVEC)
proliferating
senescent
*
1 2 3 4 5 6 7 8 9 ≥10
consecutive TADs in cluster
C
proliferating
avgTAD=0.93 Mbp
senescent
avgTAD=1.1 Mbp
chr17
N=77
N=69
TAD size (Mbp)
3.0 1.5 0.75
D
all clustered TADs (pro)
HMGB1
H3K27ac
.050
.045
.040
.035
clusters with >3 TADs (pro)
.055
singular TADs (pro)
read density (per million mapped reads)
-20 kbp 5' 3' +20 kbp
cluster breath
E
gene expression changes per TiLO cluster
median fold-change
-0.77 (N=235)
-0.85 (N=302)
-0.83 (N=86)
singular TADs
clusters w/ >3 TADs
rest of all clusters
0.85 (N=358)
0.87 (N=336)
0.76 (N=222)
-4 -3 -2 -1
downregulation (log2FC)
1 2 3 4 5
upregulation (log2FC)
F
singular
>3-TAD
rest
GO terms/pathways in the different TAD clusters
cell cycle regulation
nuclear division
cholesterol biosynthesis
mitotic phase transition
DNA replication
cellular response to ionizing radiation
positive regulation of cell cycle
ATR activation in response to replication stress
DNA conformation change
regulation of chromosome organization
telomere organization
activation of E2F1 targets at G1/S
PLK1 pathway
regulation of CDK activity
spindle organization
downregulated genes
extracellular matrix organization
integrin-1 pathway
NABA core matrisome
cell junction organization
cell-cell adhesion
response to wounding
chemotaxis
cytokine production in immune response
regulated exocytosis
lysosome
cellular response to lipid
p53 downstream pathway
positive regulation of programmed cell death
nuclear β-catenin pathway
post-translational protein phosphorylation
upregulated genes
-log(P-value)
15
10
5
2
ns

Figure 3. HMGB1 marks individual TADs harboring SASP-related genes.

(A) Overview of TiLO. TADs in each HUVEC chromosome are treated as nodes in an interacting network and inter-TAD interaction strength is used to infer network connections. These connections are then sliced and network robustness is assessed iteratively; the final network clustering is determined once the minimum number of TAD clusters is reached.

(B) Bar plots show the number of TADs contained in clusters of 1 or more TADs derived using TiLO on proliferating (*grey*) and senescent HUVEC Hi-C data (*green*). *: *P*<0.01; Fisher's exact test.

(C) Illustration of TAD clusters identified using proliferating (*left*) and senescent HUVEC Hi-C data (*right*) from chromosome 17. Each sphere represents one TAD; the most 5'/3' TADs are indicated by purple and yellow dots, respectively; arrows indicate "singular" TADs.

(D) Line plots show the mean distribution of HMGB1 (*green*), CTCF (*grey*), and H3K27ac ChIP-seq signal (*magenta*) at the extremities of all clustered TADs (±20 kbp; *top*), of clusters of >3 TADs (*middle*), or of "singular" TADs (*bottom*; as those indicated in panel C) from proliferating HUVECs.

(E) Box plots show the ($\log_2$) fold-change in mRNA levels of senescence-regulated genes embedded in the three cluster categories from panel D. The median fold-changes in expression and the numbers of genes in each plot are also indicated. *: *P*<0.01; Wilcoxon-Mann-Whitney test.

(F) Heatmaps show GO terms/pathways associated with differentially-expressed genes in the three cluster categories from panel D. Chromatin organization- and SASP-related GO terms are highlighted (*grey rectangles*).

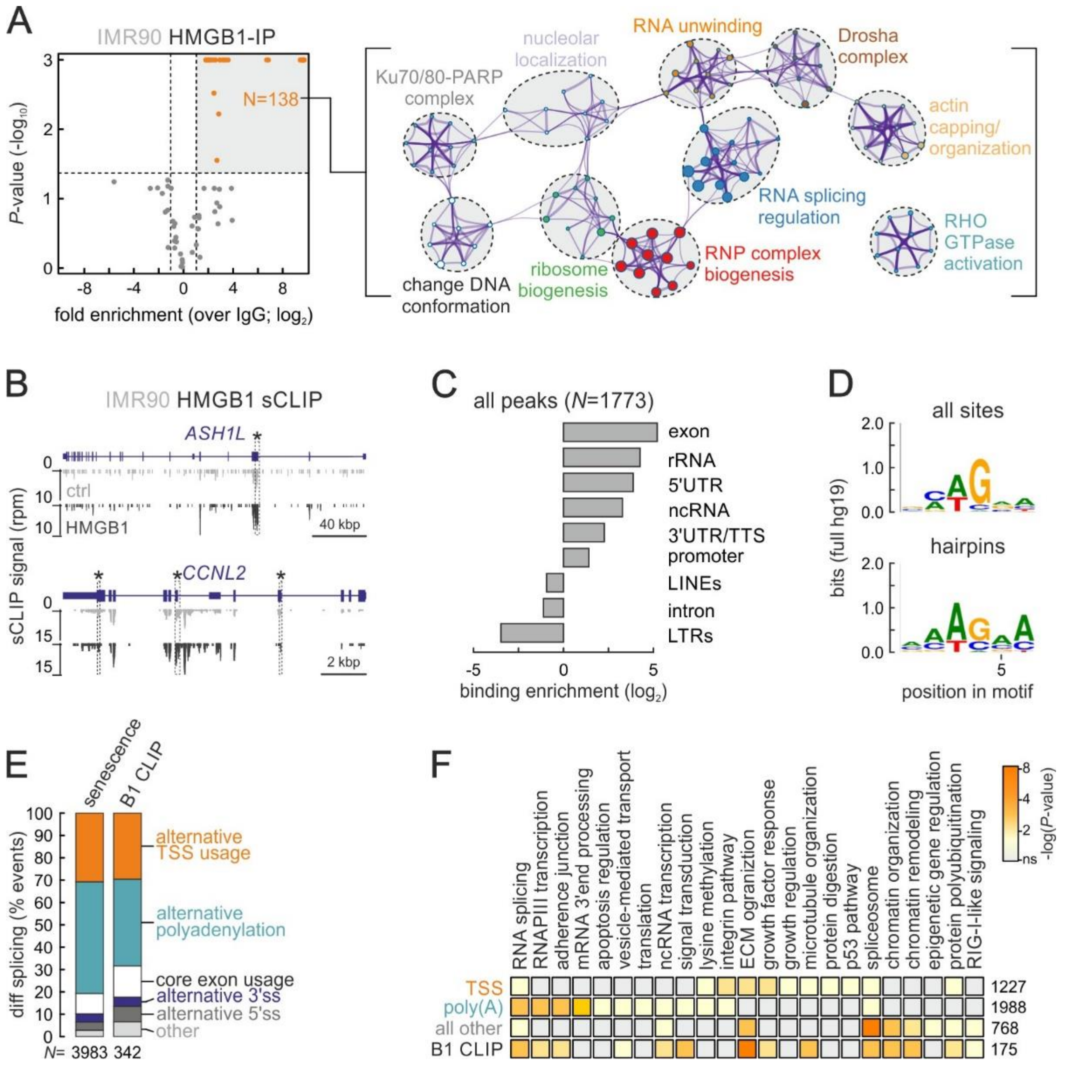

Figure 4. HMGB1 binds specific mRNAs and affects splicing choices.

(A) Scatter plot (*left*) shows proteins from proliferating IMR90 co-immunoprecipitating with HMGB1 (*orange dots*) compared to IgG controls. HMGB1 interactors associate with GO terms/pathways illustrated in the network analysis (*right*; node size refelcts the number of proteins it includes).

(B) Genome browser views showing HMGB1 sCLIP signal (*black*) along the *ASH1L* and *CCNL2* IMR90 loci; input tracks (*grey*) provide a negative control. *: significantly-enriched peak.

(C) Bar graphs show genomic distribution of HMGB1 RNA-bound peaks ($\log_2$ enrichment) from proliferating IMR90.

(D) Logos show consensus hexameric motifs deduced from all (*top*) or hairpin-embedded HMGB1 sCLIP peaks (*bottom*) using ssHMM analysis.

(E) Bar plots show relative occurrence of differential-splicing events in IMR90 undergoing senescence or only for those mRNAs bound by HMGB1. The number of bound mRNAs (*N*) analyzed is indicated.

(F) Heatmaps show GO terms/pathways associated with differentially-spliced mRNAs from three categories from panel E (TSS, poly(A), all other) or with all HMGB1-bound mRNAs (B1 CLIP).

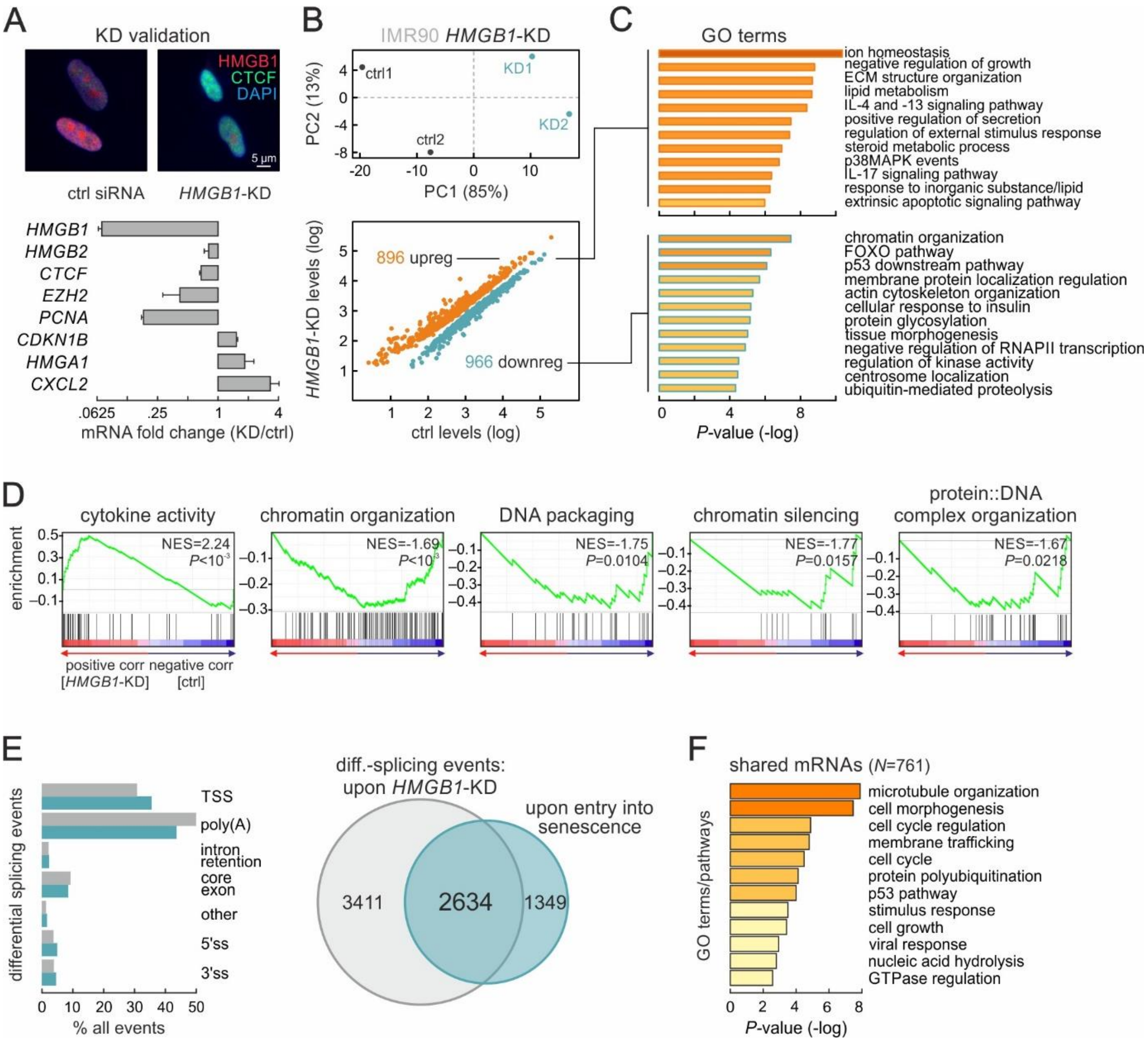

Figure 5. *HMGB1* knockdown induces senescence-specific gene expression changes.

(A) Immunofluorescence (*top*) and RT-qPCR analyses (*bottom*; mean fold-change ±S.D., *N*=3) confirm *HMGB1* knockdown in IMR90 at the protein and mRNA levels, respectively. Bar: 5 μm.

(B) PCA analysis plot (*top*) of the two control (*black*) and *HMGB1*-knockdown replicates (*green*). Scatter plot (*bottom*) shows genes significantly up- (>0.6 $\log_2$-fold change; *orange*) or downregulated (<-0.6 $\log_2$-fold change; *green*) upon *HMGB1* knockdown.

(C) Bar plots show the most enriched GO terms associated with the gene sets in panel B and their enrichment *P*-values (*right*).

(D) Gene set enrichment analysis (GSEA) of gene expression changes in *HMGB1*-KD IMR90. Normalized enrichment scores (NES) and *P*-values are indicated for each correlation.

(E) Bar plots (*left*) show occurrence of significant splicing events upon senescence (grey) or *HMGB1*-KD (*green*). Venn diagram (*right*) shows the overlap of these events.

(F) Bar plots show the most enriched GO terms associated with the 761 shared differentially-spliced mRNAs in panel E and their enrichment *P*-values (*right*)

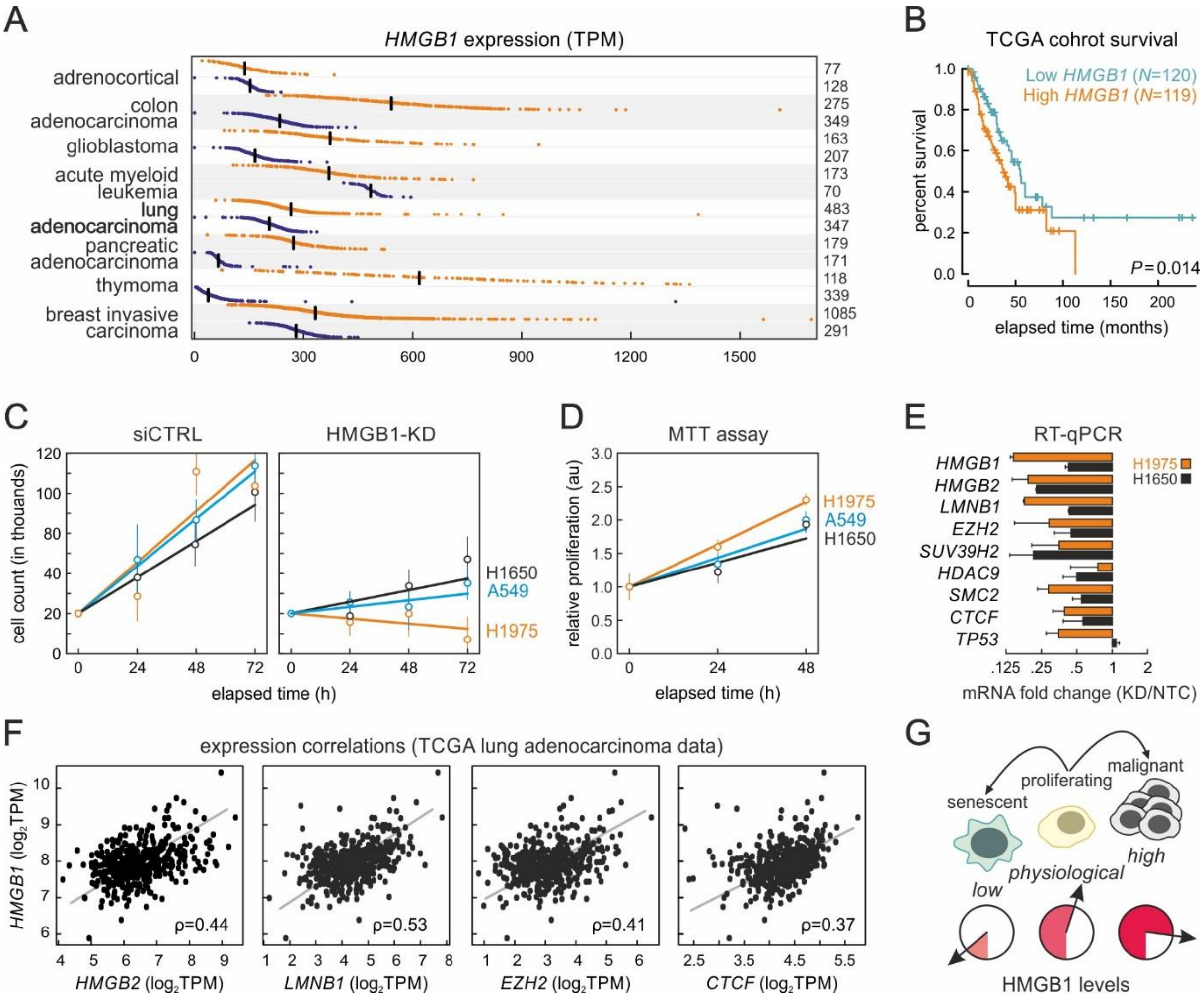

Figure 6. Lung cancer cell proliferation is dependent on HMGB1.

(A) Dot plots show *HMGB1* expression levels (in transcripts per million) across individual patient samples from 8 cancer types (*orange*) matched to levels in normal tissue (*blue*; all data from the TCGA). Mean expression (*black line*) and number of samples per each type (*right*) are shown.

(B) Line plot shows differences in per cent survival of lung adenocarcinoma patients with low (*green*) and high *HMGB1* levels (*orange*) from the TCGA cohort; the associated log-rank *P*-value is shown.

(C) Data points show mean changes (±S.E.M.; *N*=3) in cell counts of control or *HMGB1*-knockdown H1975 (*orange*), A549 (*blue*), and H1650 lung adenocarcinoma cells (*black*) over 72 h; trend lines are fitted using least mean square regression.

(D) As in panel C, but for relative proliferation rates of wild-type cells in MTT assays over 48 h.

(E) Bar graphs show mean fold-change (±S.E.M.; *N*=3) in mRNA levels of senescence-relevant genes upon *HMGB1* knockdown in H1975 or H1650 cells.

(F) Correlation of *HMGB1* expression to genes from panel E (data from the lung adenocarcinoma TCGA cohort); Pearson's correlation coefficients (ρ) and trendlines (*grey*) are shown.

(G) Cartoon describing a simple model of cell proliferation dependency on intracellular HMGB1 levels, with senescent cells exhibiting the lowest and malignant cells the highest ones.

Supplementary figures

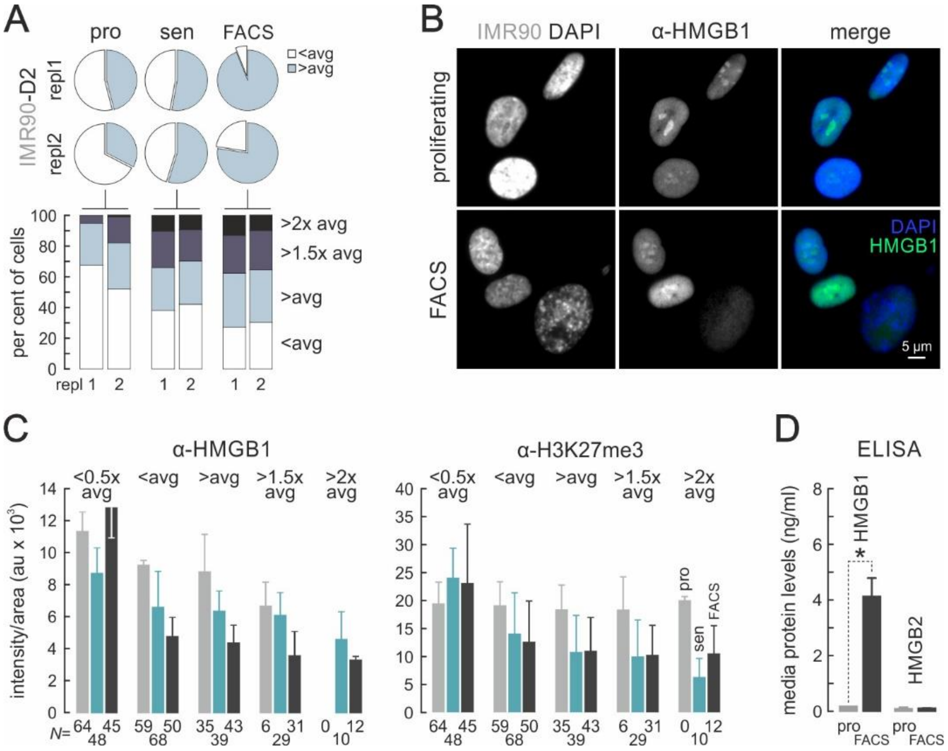

Figure S1. HMGB1 depletion mostly occurs in enlarged nuclei.

(A) Pie charts (*top*) show senescent IMR90 cell populations containing consistently larger nuclei, which can be enriched via FACS sorting. Bar graphs (*bottom*) show stratification of population by increasing nuclear size (compared to the population average).

(B) Representative immunofluorescence images of proliferating (*top row*) and senescent FACS-sorted IMR90 (*bottom row*) show reduced HMGB1 levels in enlarged nuclei. Bar: 5 µm.

(C) Bar graphs show reduced mean HMGB1 (*left*) and H3K27me3 levels ± 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. *N* indicates the number of cells analyzed in each subgroup.

(D) Bar graphs show mean HMGB1/B2 protein levels ± S.D. in the growth medium of proliferating (*grey*) or senescent IMR90 (*green*). *: *P*<0.01; unpaired two-tailed Student's t-test (*N*=2).

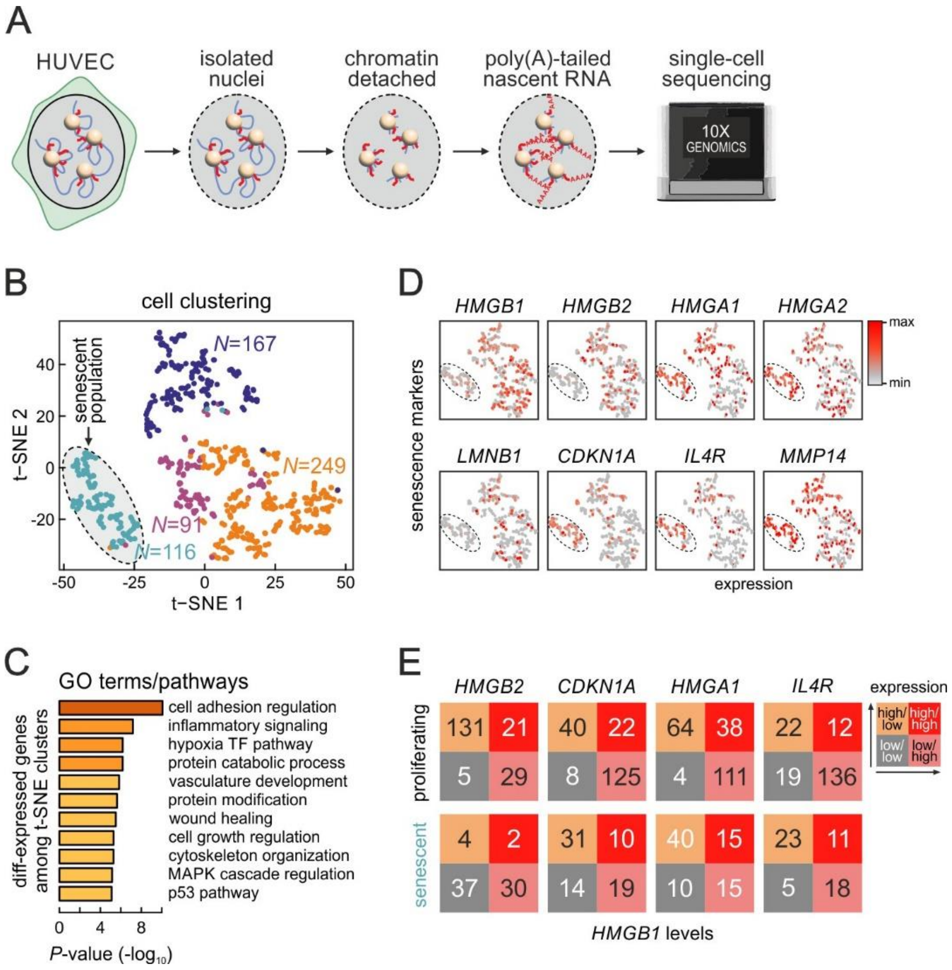

Figure S2. Single-cell sequencing of nascent RNA.

(A) Strategy for nascent scRNA-seq involves isolation of intact HUVEC nuclei, detachment of non-transcribed chromatin (*blue*) via DNase I digestion, and *in situ* polyadenylation of nascent RNA (*red*) at transcription sites (*spheres*), before processing on a 10X Genomics platform.

(B) t-SNE clustering of proliferating and senescent cell populations (*grey oval*) revealed 4 distinct clusters (*blue*, *red*, *purple*, and *green*; *N* indicates the number of cells per cluster).

(C) Bar graphs show (-log) enrichment *P*-values for GO terms associated with differentially-expressed genes amongst clusters in panel B.

(D) Heatmaps show normalized expression of exemplary senescence marker genes in single cells.

(E) Correlation of *HMGB1* expression to that of genes from panel D in single proliferating (*top row*) or senescent HUVEC (*bottom row*; low: ≤1, high: >1 transcript per million).

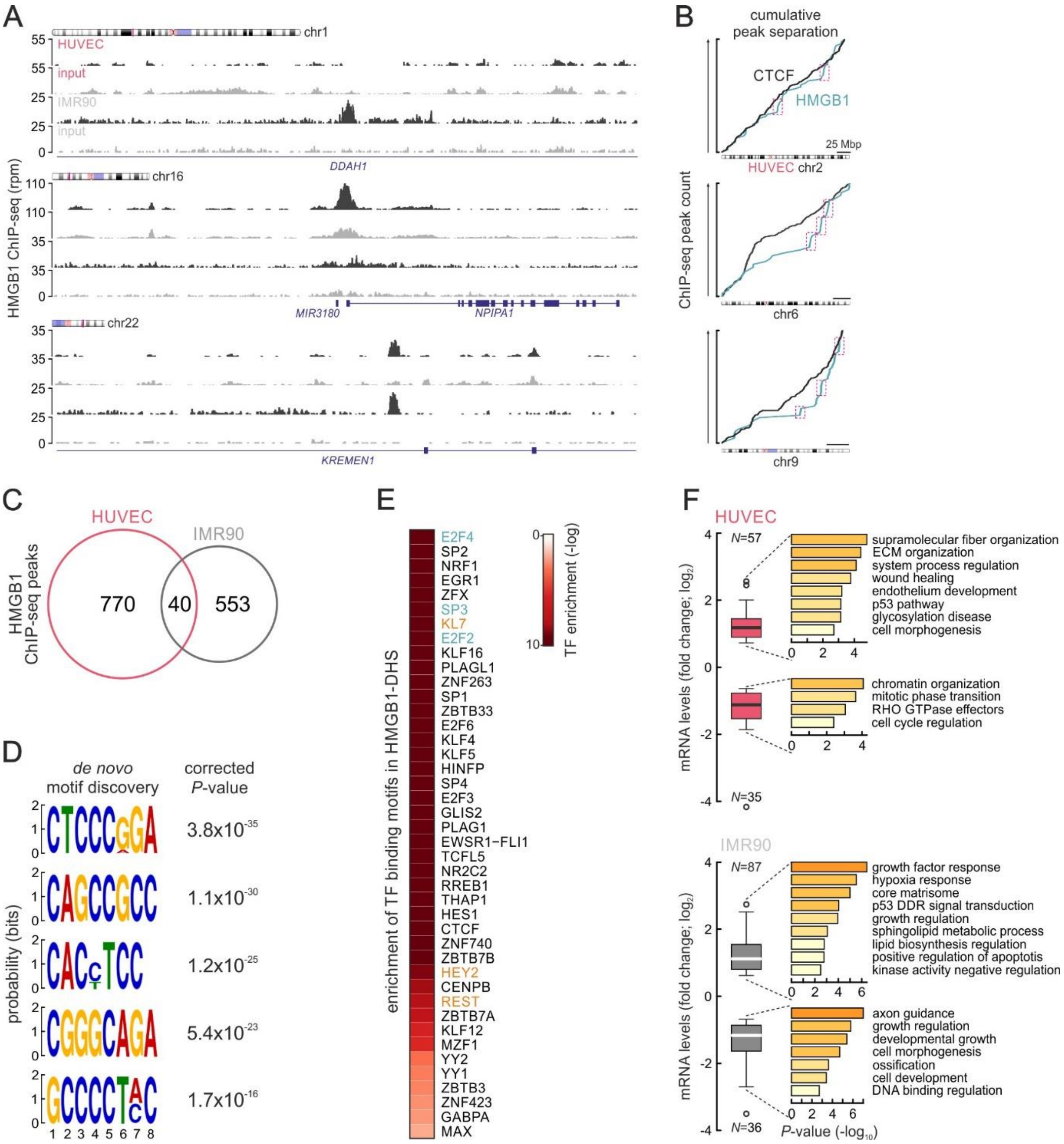

Figure S3. HMGB1 chromatin-binding features in proliferating cells.

(A) Genome browser views showing HMGB1 ChIP-seq signal from loci in chromosomes 1, 16, and 22 of both HUVEC and IMR90 (*black*); input tracks (*grey*) provide a control.

(B) Plots show the cumulative distribution of HMGB1 binding peaks along HUVEC chromosomes 2, 6 and 9 (*green*); the distributions of CTCF peaks (*black*) provide a reference.

(C) Venn diagrams show minimal overlap between HMGB1 binding peaks from HUVEC and IMR90.

(D) Logos for the five most enriched *de novo* motifs discovered under accessible HMGB1 peaks.

(E) Heatmap shows most enriched TF binding motifs under accessible HMGB1 peaks; up- (>0.6 log$_2$-fold change; *orange*) or downregulated TFs (<-0.6 log$_2$-fold change; *green*) are indicated.

(F) Box plots show the fold-change in mRNA levels (log$_2$) of genes differentially-expressed upon senescence and directly bound by HMGB1 in HUVECs (*top*) and IMR90 (*bottom*). The most enriched GO terms associated with each subgroup and their enrichment *P*-values are shown. The number of peaks (*N*) analyzed is indicated below each bar.

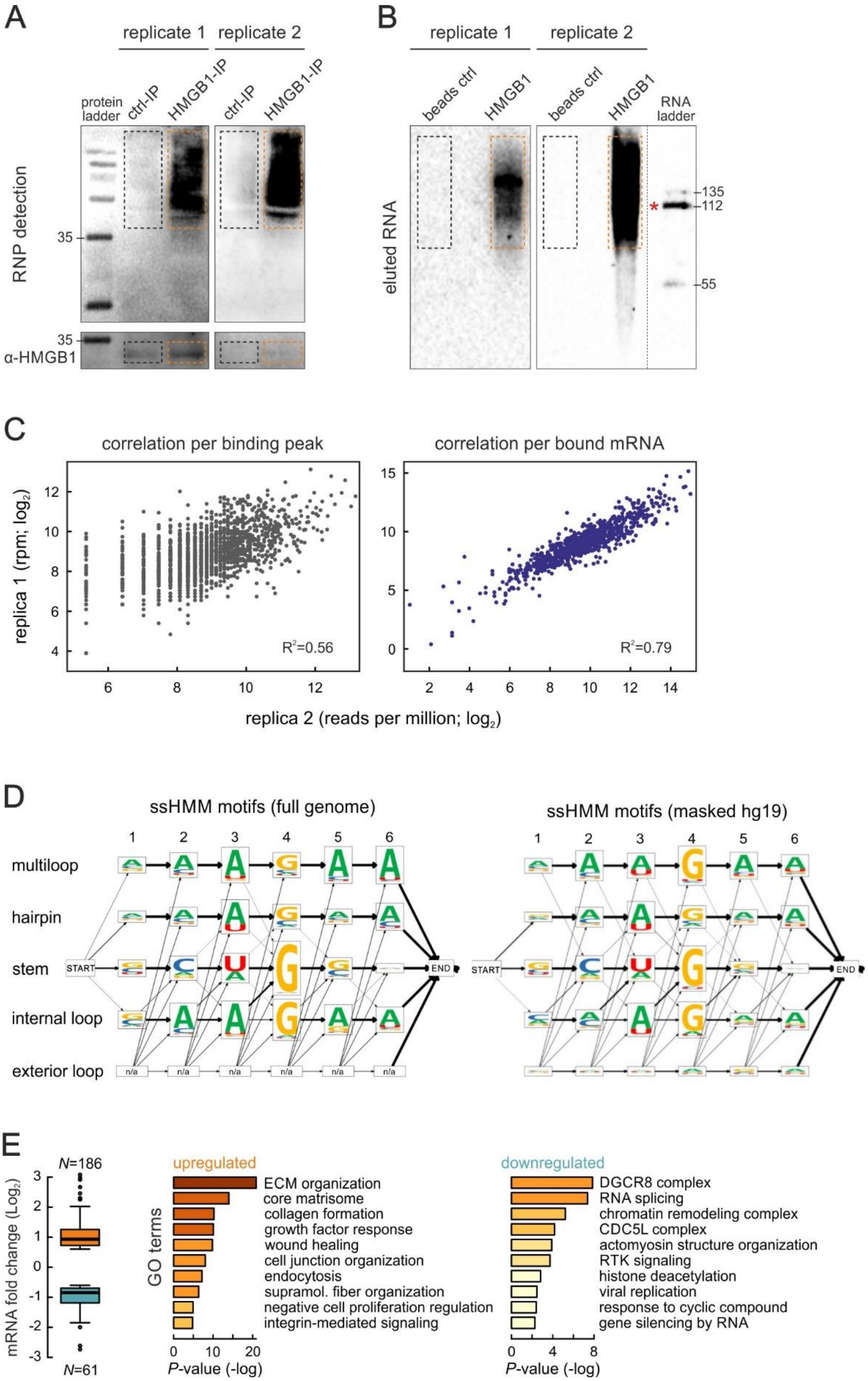

Figure S4. HMGB1 sCLIP controls and analysis.

(A) Electrophoretic profiles of control (beads only; *black dotted square*) and HMGB1 IP (*orange dotted rectangle*) probed for RNA (*top*) or the HMGB1 protein (*bottom*) in both sCLIP replicates. The 35-kDa band of the molecular mass ladder is indicated.

(B) Electrophoretic profiles of RNA eluted from control (beads only; *black dotted square*) or HMGB1 IP (*orange dotted rectangle*) in both sCLIP replicates. The 112-nt band of the molecular mass ladder is indicated (*red star*) and corresponds to ~144 pg of RNA.

(C) Scatter plots show the correlation of sCLIP data from two independent biological replicates compared per binding peak (*left*) or per bound mRNA normalized read count (*right*). Spearman correlation values (R^2) are indicated.

(D) Output of ssHMM motif analyses of sCLIP data showing sequence probabilities for HMGB1-bound hexameric motifs predicted to form different structures using either the complete (*top*) or masked reference human genome (*bottom*; hg19).

(E) Box plots show fold-change in expression ($\log_2$) of mRNAs differentially-regulated upon senescence and directly bound by HMGB1 in IMR90 sCLIP. The most enriched GO terms associated with each subgroup and their enrichment *P*-values are shown. The number of peaks (*N*) analyzed is indicated below each bar.

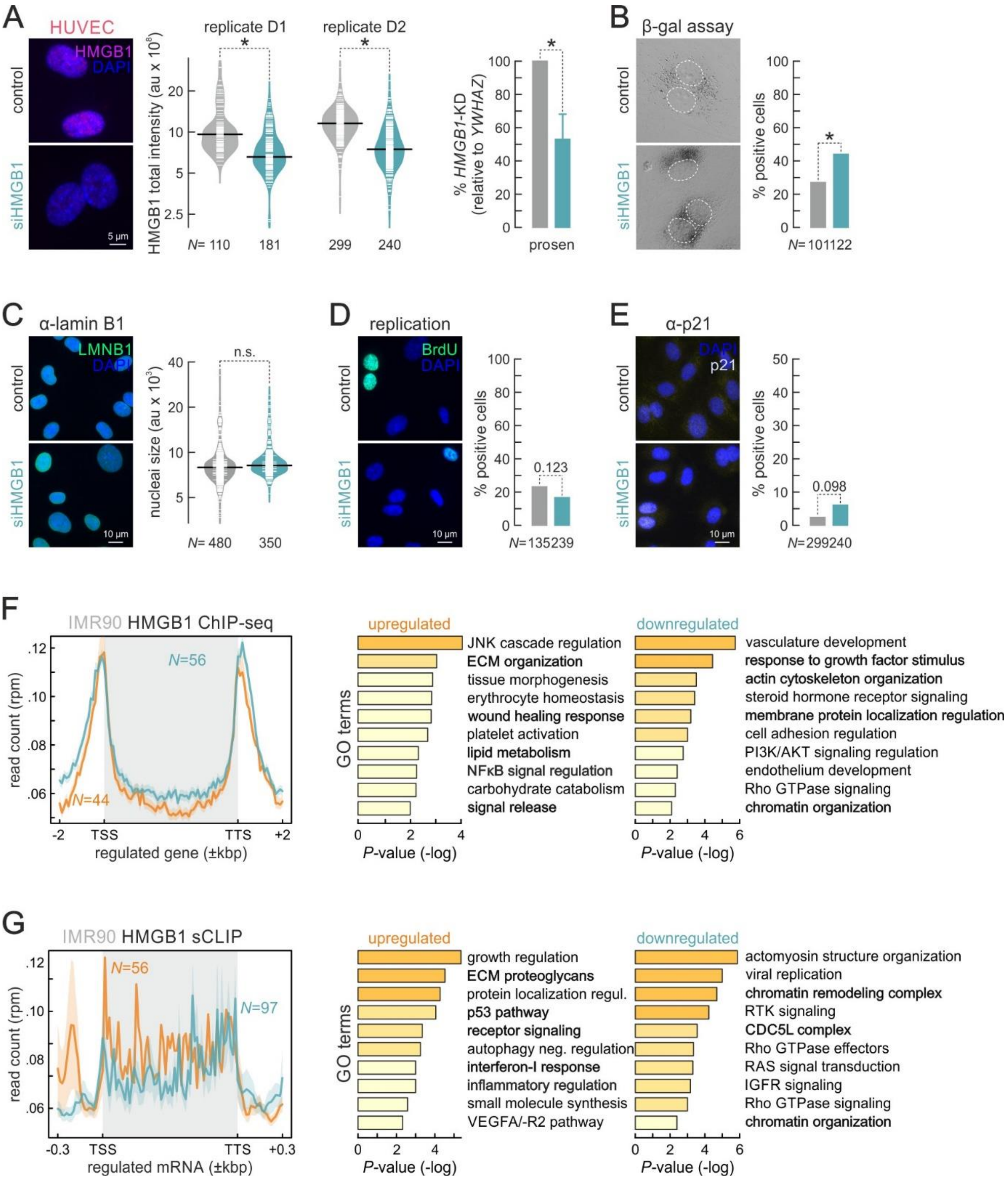

Figure S5. Effects of *HMGB1* knockdown in proliferating HUVECs and IMR90.

(A) Representative immunofluorescence images of siHMGB1-treated HUVECs show reduced HMGB1 levels compared to control cells (*left*), and bean plots quantify knockdown efficiency (*middle*; *N* indicates the number of cells analyzed). Bar plots (*right*) show normalized *HMGB1* mRNA levels (± S.D.; *N*=2) in knockdown compared to control cells. Bar: 5 µm. *: *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 show elevated β-galactosidase activity levels compared to control cells (*left*), and bar plots 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 (*left*), and bean plots quantify this (*right*; *N* indicates the number of cells analyzed). Bar: 10 µm. No statistically significant difference (n.s.); Wilcoxon-Mann-Whitney test.

(D) As in panel A, but for BrdU incorporation (*left*), and bar plots quantify BrdU-positive cells (*right*; *N* indicates the number of cells analyzed). Bar: 10 µm. *P*=0.123; Fisher's exact test.

(E) As in panel A, but for p21 levels (*left*), and bar plots quantify p21-positive cells (*right*; *N* indicates the number of cells analyzed). Bar: 10 µm. *P*=0.098; Fisher's exact test.

(F) Line plots (*left*) show the mean distribution of IMR90 HMGB1 ChIP-seq signal along gene bodies of genes up- (*orange*) or downregulated upon *HMGB1* knockdown (*green*; *N* indicates the number of directly HMGB1-bound genes). Bar plots show the most enriched GO terms associated with either gene set and their enrichment *P*-values (*middle*/*right*), and those relevant to senescence are highlighted (*bold*).

(G) As in panel C, but using IMR90 HMGB1 sCLIP data

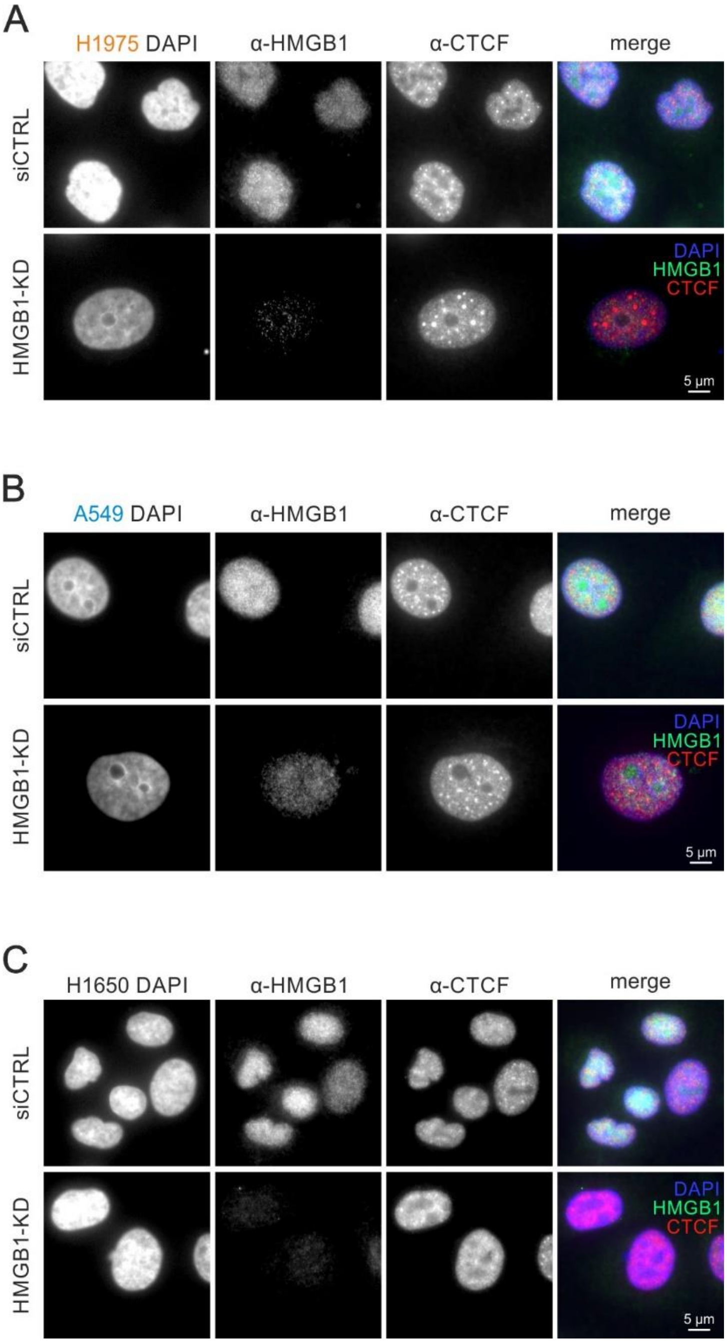
A
H1975 DAPI
α-HMGB1
α-CTCF
merge
siCTRL
HMGB1-KD
DAPI
HMGB1
CTCF
5 µm
B
A549 DAPI
α-HMGB1
α-CTCF
merge
siCTRL
HMGB1-KD
DAPI
HMGB1
CTCF
5 µm
C
H1650 DAPI
α-HMGB1
α-CTCF
merge
siCTRL
HMGB1-KD
DAPI
HMGB1
CTCF
5 µm

Figure S6. Formation of CTCF clusters in lung adenocarcinoma lines.

(A) Representative immunofluorescence images of control (siCTRL) and *HMGB1*-knockdown H1975 cells (HMGB1-KD) showing HMGB1 (*green*) and CTCF (*red*) distribution, and counterstained with DAPI (*blue*). Bars: 5 µm.

(B) As in panel A, but for A549 cells.

(C) As in panel A, but for H1650 cells.