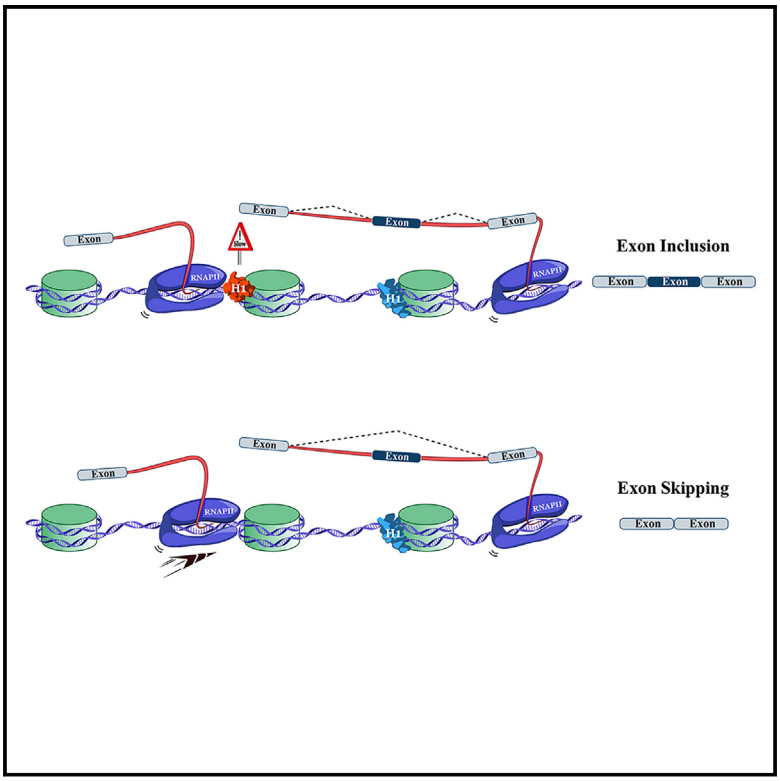
Molecular Cell

Article

Human histone H1 variants impact splicing outcome by controlling RNA polymerase II elongation

Graphical abstract



Highlights

- Histone H1 variants differ in binding patterns across exons and introns and GC content
- H1 variants are major regulators of splicing
- H1 variants influence splicing through fine-tuning of RNA polymerase II elongation

Authors

Corina Pascal, Jonathan Zonszain, Ofir Hameiri, ..., Albert Jordan, Eran Meshorer, Gil Ast

Correspondence

gilast@tauex.tau.ac.il

In brief

Pascal et al. characterized genomic binding profiles of six H1 variants by creating cell lines lacking up to five H1 subtypes. Most H1s bind to specific sites, influenced by factors like GC content. H1 histones regulate mRNA splicing by modulating the elongation rate of RNA polymerase II.

Pascal et al., 2023, Molecular Cell *83*, 3801–3817
November 2, 2023 © 2023 Elsevier Inc.
https://doi.org/10.1016/j.molcel.2023.10.003

CellPress

Molecular Cell

CellPress

Article

Human histone H1 variants impact splicing outcome by controlling RNA polymerase II elongation

Corina Pascal,[1,5] Jonathan Zonszain,[1,5] Ofir Hameiri,[1] Chen Gargi-Levi,[1] Galit Lev-Maor,[1] Luna Tammer,[1] Tamar Levy,[1] Anan Tarabeih,[1] Vanessa Rachel Roy,[1] Stav Ben-Salmon,[1] Liraz Elbaz,[1] Mireille Eid,[1] Tamar Hakim,[1] Salima Abu Rabe'a,[1] Nana Shalev,[1] Albert Jordan,[2] Eran Meshorer,[3,4] and Gil Ast[1,6,*]

[1]Department of Human Molecular Genetics and Biochemistry, Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel
[2]Instituto de Biologia Molecular de Barcelona (IBMB-CSIC), Carrer de Baldiri Reixac, 15, 08028 Barcelona, Spain
[3]Department of Genetics, The Alexander Silberman Institute of Life Sciences, Jerusalem 91904, Israel
[4]Edmond and Lily Center for Brain Sciences (ELSC), The Hebrew University of Jerusalem, Jerusalem 91904, Israel
[5]These authors contributed equally
[6]Lead contact
*Correspondence: gilast@tauex.tau.ac.il
https://doi.org/10.1016/j.molcel.2023.10.003

SUMMARY

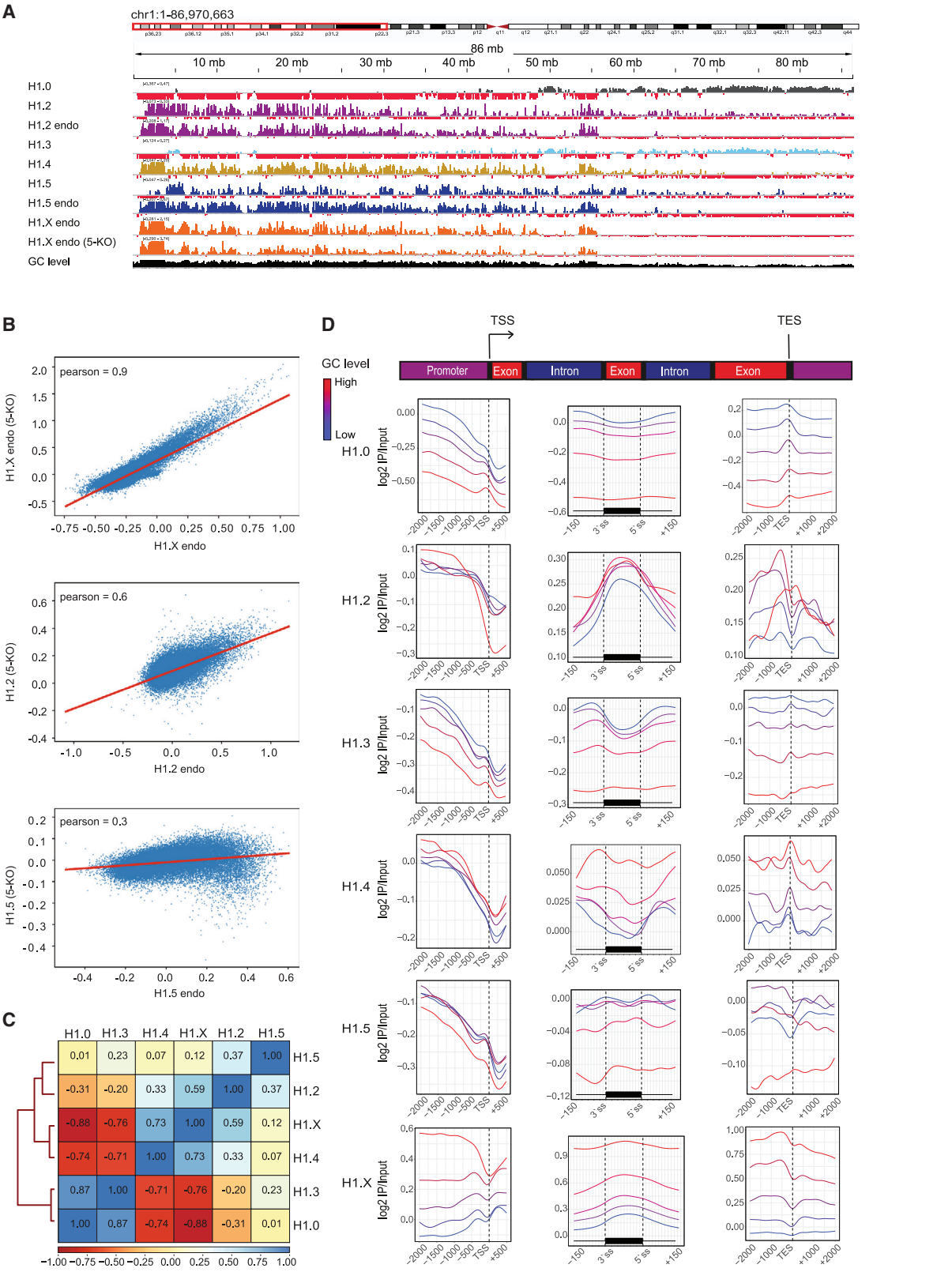
Histones shape chromatin structure and the epigenetic landscape. H1, the most diverse histone in the human genome, has 11 variants. Due to the high structural similarity between the H1s, their unique functions in transferring information from the chromatin to mRNA-processing machineries have remained elusive. Here, we generated human cell lines lacking up to five H1 subtypes, allowing us to characterize the genomic binding profiles of six H1 variants. Most H1s bind to specific sites, and binding depends on multiple factors, including GC content. The highly expressed H1.2 has a high affinity for exons, whereas H1.3 binds intronic sequences. H1s are major splicing regulators, especially of exon skipping and intron retention events, through their effects on the elongation of RNA polymerase II (RNAPII). Thus, H1 variants determine splicing fate by modulating RNAPII elongation.

INTRODUCTION

Chromatin organization controls all aspects of cellular processes and functions.[1–3] In eukaryotes, the higher-order chromatin organization is the result of DNA wrapping around a nucleosome, which consists of four core histone proteins, H2A, H2B, H3, and H4, and the linker histone H1.[4] H1 binds DNA sequences close to sites at which the DNA enters or exits the nucleosome.[4] H1 is highly conserved in mammals, and functionally related proteins are found in protists and bacteria.[5] This evolutionary conservation of H1 implies important functions. Eleven mammalian H1 variants have been identified: seven are expressed in somatic cells, including five in a replication-dependent manner (H1.1, H1.2, H1.3, H1.4, and H1.5) and two that replicate independently (H1.0 and H1.X); four are germ cell specific (H1t, H1T2, HILS1, and H1oo).[6–8] H1s bind dynamically, and only around 80% of the nucleosomes in human cells associate with H1s.[9] H1s have a rather minor effect on gene expression,[10–12] but H1 variants interact with many proteins involved in transcription, DNA methylation, and other cellular functions, suggesting that they have additional cellular activities.[13]

Alternative splicing (AS) of mRNA precursors is a fundamental process that leads to a high level of genomic complexity.[14,15] Splicing is regulated on multiple levels, including transcription, chromatin organization, and epigenetic factors.[3,16–25] There are higher levels of nucleosome occupancy on exon sequences than on flanking introns, and there are higher levels of histone modifications, such as H3K36me3, on nucleosomes that bind exon sequences.[16,22,26–30] Since splicing proceeds co-transcriptionally[3,31] with the upstream 5′ splice site (SS) of a pre-mRNA bound to RNA polymerase II (RNAPII) during intron synthesis, the RNAPII elongation rate influences the levels of exon inclusion and intron retention (IR).[32–38] H1.2 interacts with elongating RNAPII,[39] implying its involvement in transcription. H1 variants have also been associated with certain splicing factors, implicating them in splicing regulation.[40] We recently observed that H1.5 influences SS selection in fibroblast-derived normal human lung cells (IMR90).[41]

Depleting or knocking out the somatic variants of histone H1 (H1s) impacts the survival, proliferation, and development of various cell types and organisms, including humans and mice.[12,42–44] To date, the maximum number of H1s knocked out simultaneously in mammals is three, and this had an embryonic lethal phenotype in mice, but the knockout (KO) of a single H1 is non-lethal, suggesting that H1 subtypes can compensate for each other.[6,44,45] In IMR90 cells, small differences in the distributions of H1.2, H1.3, H1.4, and H1.5 along the genome were found; however, the resolution of these experiments was relatively low.[46] Analyses of recombinant H1s tagged with hemagglutinin (HA) in human T47D breast cancer cells demonstrated



(legend on next page)

that H1.2 is primarily distributed in intergenic regions, whereas H1.0 and H1.X are bound to gene-rich regions.[47] Interestingly, the various H1s appear to be distributed differently in different cell types, suggesting tissue-specific functions.[40,46,48–50]

Here, we hypothesized that H1 variants function in mRNA-splicing regulation. We generated a human genomic binding map of the six histone H1 variants expressed in the somatic human embryonic kidney (HEK293) cell line. We generated viable HEK293 cells lacking five somatic H1 variants (5-KO) by knocking out H1-encoding genes one by one in a stepwise fashion. The 5-KO cells are viable but have an impaired G2/M checkpoint and nucleosome spacing that differs from that of wild-type (WT) cells, confirming the importance of H1s in cell cycle regulation and for maintaining nucleosome positioning. Chromatin immunoprecipitation sequencing (ChIP-seq) analyses of endogenously expressed H1 proteins and exogenously expressed HA-tagged H1 proteins revealed that most H1s bind to specific sites, and binding depends on multiple factors, including guanine-cytosine (GC) content. H1s are major regulators of AS, especially of exon skipping (ES) and IR events, through their effects on the elongation of RNAPII. Deletion of H1.2 in the N- and C-terminal domains influences its ability to affect splicing and elongation and indicates that H1.2 interacts with RNAPII, implying the mechanism of co-transcriptional splicing. Therefore, H1s play an important role in the specific regulation of co-transcriptional splicing via direct and/or indirect association with RNAPII.

RESULTS

Generation of cell lines lacking specific histone H1 somatic variants

H1 subtype diversity is evolutionarily conserved, suggesting that variants have unique features,[50] although some H1 variants can compensate for each other.[13,51] This H1 compensation is apparently incomplete, since the depletion of multiple H1s has a significant effect on cell functioning.[10,52] To examine H1 subtype-specific occupancy of chromatin without interference or competition with other H1 somatic subtypes, we generated a series of H1 KO HEK293 cells and cells in which H1 histones were transfected back (rescued). We used CRISPR-Cas9 with single guide RNAs (sgRNAs) to target individual H1-encoding genes (Table S1). We analyzed six H1 variants expressed in somatic cells: H1.0, H1.2, H1.3, H1.4, H1.5, and H1.X. We generated a cell line lacking expression of H1.2, which is ubiquitously expressed in most human tissues.[53] We then knocked out genes encoding additional H1s in turn to obtain a cell line with five H1s KO that only expresses H1.X. The cell lines we created are listed in Table S2. It is important to note that in the KO, the mRNA of the targeted H1 gene is expressed, but the H1 protein is not translated. Loss of protein expression was validated by western blot (WB) and DNA sequencing (Figure S1A).

Cell division was abnormal in the H1 KO cells (Figures S1B and S1C), as expected since the lack of H1s alters cell cycle progression and nucleosome spacing.[10,12,54] The replication rate gradually decreased from one to five H1 KO, implying the importance of functional H1 for efficient cell division (Figure S1Biii). Cells lacking H1s took longer to clear the G2/M checkpoint than WT cells (Figures S1B and S1C). Transfection (rescue) of H1.2 variant in 1-KO and H1.2 with H1.3 to 2-KO cells abolished negative cell cycle-related effects (Figures S1Bii, S1Cvi, and S1Cvii). In H1 KO cells, there were shorter linker DNA/spacers between nucleosomes than in WT cells (Figure S1D) implying the importance of H1 variants for maintaining the nucleosome spacer length.

Next, the mRNA levels of the targeted H1 genes were evaluated, and H1.2 was present at the highest levels among the examined H1-encoding mRNAs, especially in 5-KO cells (Figure S1E). H1.1 was excluded from this analysis since its mRNA expression in HEK293 was almost undetectable and decreased upon KO of other variants, as shown by reverse-transcriptase PCR performed on WT, 4-KO, and 5-KO cells (Figure S1F). In order to observe the possible upregulation of non-somatic H1 variants, the H1t testis-specific variant was evaluated, which is not expressed in somatic cells at the protein level[55]; although its mRNA expression was detected in the control group, it decreased substantially in the 4-KO and 5-KO cells (Figure S1F). In summary, H1 KO leads to robust dysregulation of variable cell processes.

H1 variants have both shared and unique chromatin binding patterns

To study H1 genome-wide localization and variant-specific characteristics, we conducted ChIP-seq analysis on H1 KO HEK293 cells, enclosing both endogenous and exogenous H1 proteins. We analyzed endogenous H1.2, H1.5, and H1.X and HA-tagged H1.0, H1.2, H1.3, H1.4, and H1.5 transfected to the 5-KO cells (Figure 1A). We detected common distribution patterns shared by subsets of H1s. For example, H1.0 and H1.3 had a low binding affinity at the start of chromosome 1, where the GC content is relatively high, whereas H1.2, H1.4, H1.5, and H1.X have a

Figure 1. The human genomic landscape of somatic H1 variant binding sites

(A) Integrative Genomics Viewer (IGV) snapshot: H1 variants binding on human chromosome 1 (86 Mb, shown schematically above the snapshot). H1 input-subtracted ChIP-seq signals are plotted. Samples represent HA-tagged H1 immunoprecipitated (IP) from KO cell lines (exogenous) or WT using specific H1 antibodies (when available, marked endo). Negative values are colored in red. The GC level along the genome segment is also shown.

(B) Scatter plots of Pearson correlation between $\log_2$ IP/input ChIP-seq signal for H1.X, H1.2, and H1.5 the endogenous (endo) and exogenous (HA-tagged), WT, and 5-KO background. $p < 0.001$ for all panels, permutation test.

(C) Heatmap of Pearson correlation coefficients and hierarchical clustering of the $\log_2$ IP/input ChIP-seq signals for H1 variants expressed in 5-KO background. HA-tagged (H1.0, H1.2, H1.3, H1.4, and H1.5) and H1.X endogenous. $p < 0.01$ for all correlation coefficients, permutation test.

(D) Binding profiles of H1s, same samples as in (C), plotted as mean $\log_2$ IP/input ChIP-seq signal at a single base resolution in the regions from 2,000 nt upstream to 500 downstream of TSSs (left), across intron-exon-intron structures including 75 nt of sequences from both ends of the exons and 150 nt of their flanking introns, excluding SS signals (middle), and in the 2,000 nt on either side of TESs (right). Color indicates GC-content level: red represents high GC content for leftmost and rightmost panels and leveled exon-intron architecture for middle panels; blue represents low GC content for leftmost and rightmost panels and differential exon-intron architecture for the middle panels. Exons are depicted as boxes and introns as lines.

high signal in the same region (Figure 1A). In contrast, in the following segment of chromosome 1, where the GC content is lower, we observed an inverted binding profile of these two groups of variants (Figure 1A). Global coverage patterns of H1s over chromosomes are shown in Figure S1G. The binding patterns of these two groups of H1s are related to the GC level: H1.2, H1.4, and H1.X tend to bind preferentially to regions with high GC content, whereas H1.0 and H1.3 are associated with low GC content, and H1.5 has peaks in regions with both high and low GC content (Figures 1A, 1D, and S2A).

To assess H1 redistribution resulting from KO of other H1 variants, we performed a Pearson correlation analysis comparing ChIP-seq results between WT and KO cells (Figure 1B). There was a very strong correlation ($r = 0.9$) for H1.X between the WT and the 5-KO cells, indicating that H1.X binds to specific sites even when most other H1s are depleted from the cells. The binding pattern of HA-tagged H1.2 transfected to the 5-KO cells was also strongly correlated to the pattern of endogenous H1.2 in the WT cells ($r = 0.6$). The moderate correlation ($r = 0.3$) between HA-tagged H1.5 transfected to the 5-KO cells and endogenous H1.5 in WT cells suggests reduced specificity when other H1 variants are absent (Figure 1B), possibly due to anti-H1.5 antibody affinity or that H1.5 incorporation is sensitive to the HA tag. We used hierarchical clustering to evaluate the genomic distribution similarities of the HA-tagged H1s (H1.0, H1.2, H1.3, H1.4, and H1.5) that were exogenously introduced and endogenous H1.X, all in the 5-KO cells. This clustering revealed two major groups, the first containing H1.2, H1.4, H1.5, and H1.X and the second containing H1.0 and H1.3 (Figure 1C). H1.0 and H1.3 binding patterns were correlated ($r > 0.8$), but the signals of these two variants were negatively correlated ($r \leq -0.7$) with H1.4 and H1.X, members of the other cluster. This indicates that the two groups have complementary signal distributions. H1.0 and H1.3 preferentially bind intronic sequences; over 40% of the binding peaks of these two variants were assigned to introns (Figure S2B). H1.4 displays a strong preference for binding within 1 kb of promoters, with 21.64% of the peaks in these regions (Figure S2B). Taken together, these results demonstrate that H1s have unique and shared binding regions.

To investigate the biological functions of each H1 variant, we generated binding profiles along gene bodies. We classified a dataset of protein-coding genes into five groups according to their GC level. Our results support the previous report that H1s are depleted from the transcription start site (TSS) regions[46,47] (Figure 1D, left). Given our KO cell results, the reduction in binding around the TSS appears independent of the presence of other H1s. There was an enrichment in signal for H1.0, H1.3, and H1.4, and mainly a depletion of H1.2, H1.5, and H1.X at transcript end site (TES) regions (Figure 1D, right).

We previously reported that intron length correlates with GC-content architecture across exon-intron structures.[56] To examine whether different GC-content architectures affect H1 variant binding profiles, we built a dataset of unique internal exons and their flanking introns.

We visualized the H1 signal along the intron-exon-intron regions, according to the five GC-content architecture groups (Figure 1D middle), as described by Amit et al. and Tammer et al.[56,57] The H1.0 signal slightly increased in the introns in regions of exon-intron junctions of low GC architecture. In contrast, H1.2 exhibited a bell-shaped curve with a strong peak at the exons and the edges of the flanking introns, regardless of the GC-content architecture. The profile for H1.3 shifted gradually from a low signal in high GC-content architecture with an almost flattened pattern to a peak of signal in the intronic part around the SS region of low GC-content architecture. H1.4 binding was higher in introns and at 5′ SSs in regions with low GC content, and at 3′ SSs in regions with high GC content. In accordance with our previous reports, binding of the H1.5 variant was higher around SSs, especially those with low GC content,[41] whereas H1.X displayed a moderate enrichment over exons and a preference for high GC-content architectures, with a gradient of binding signal from high to low GC-content genes (Figure 1D).

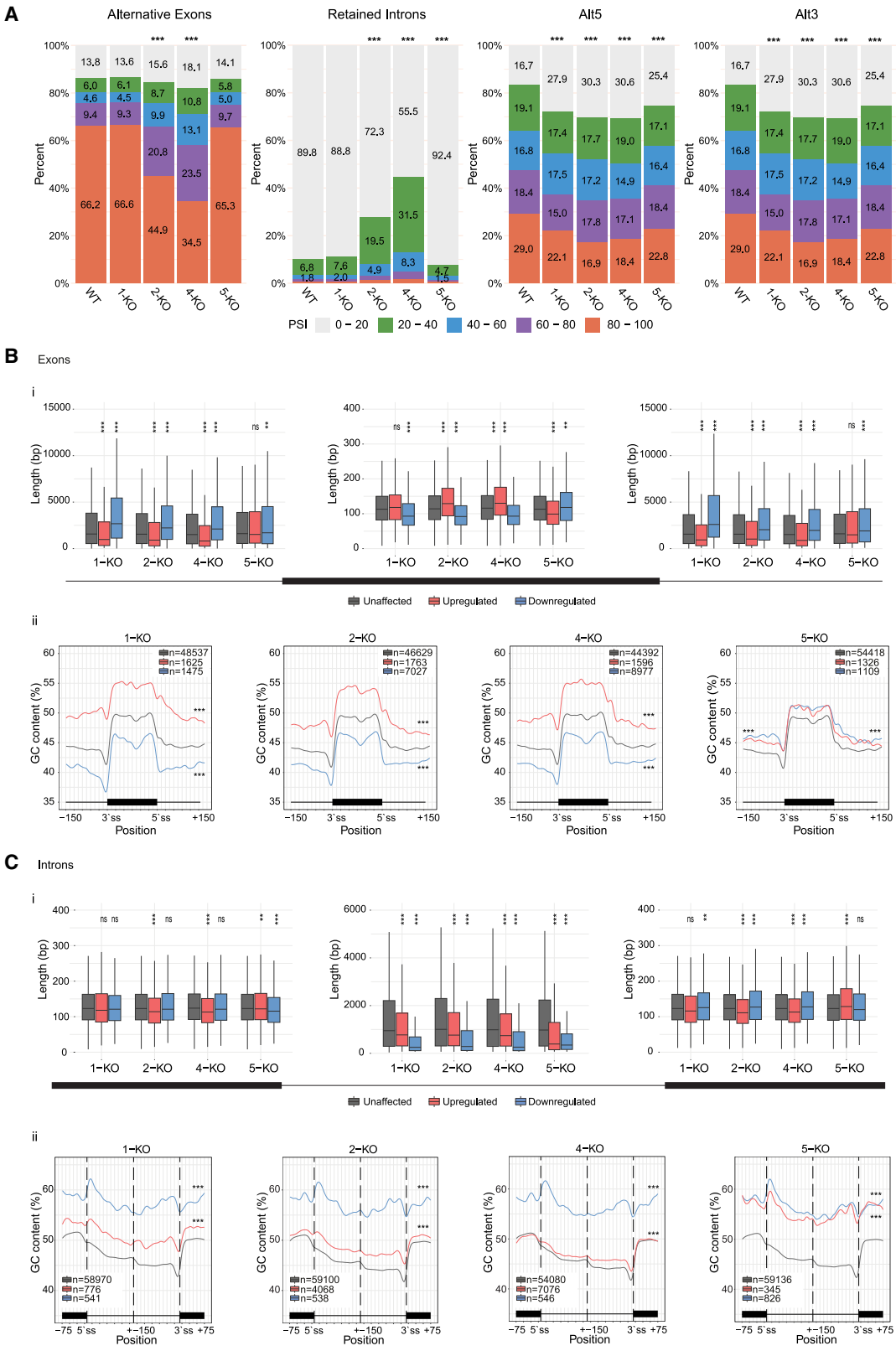
H1 KOs significantly affect splicing

In order to evaluate the effect of H1 on gene expression and splicing regulation, we performed RNA sequencing (RNA-seq) on all the KO cell lines (Table S2). We observed effects on the global gene expression patterns in H1 KO cell lines (Figure S2C). Spearman correlation analysis of gene expression between WT and KO samples revealed values that were at least/over 0.8, thus changes were not drastic (Figure S2C).

Recently, H1 was shown to preferentially bind poorly expressed genes due to the increase in nucleosome repeat length in regions where expression is repressed.[58] We categorized genes expressed in WT HEK293 cells into five groups according to their expression levels and examined H1 binding in the 5-KO cells transfected with H1-HA. For almost all the H1 variants, signals were higher over poorly expressed genes than those expressed at the highest level. The exception was H1.X, which bound genes expressed at high and low levels (Figure S2D). H1.2 and H1.5 had occupancy levels that gradually decreased from the poorly to highly expressed gene groups (Figure S2D). Thus, our results demonstrate that the H1 subtypes display specific binding characteristics with implications for their functions in gene expression and/or mRNA processing.

Analysis of the RNA-seq revealed that H1s are major splicing regulators. We compared splicing efficiency differences between WT and KO cells (1-KO, 2-KO, 4-KO, and 5-KO), assessing splicing events' specific inclusion levels (Figure 2A). H1 variants primarily affect exon and intron splicing, with minor influence on alternative 5′ and 3′ SSs (Alt5 and Alt3, respectively; Figure 2A). The quantities of altered exons and introns gradually increased from 1-KO to 4-KO, and inclusion differences were similar between 1-KO and 5-KO cells (Figure 2).

The general effect of H1 variant loss was to reduce the inclusion of the affected exons and increase the inclusion levels of introns (Figure 2A). To validate the effect of H1 KO on splicing, we conducted RT-qPCR analyses on exons of nine genes detected in the RNA-seq analysis (events with $\geq$10% absolute differential inclusion between WT and 4-KO cells) (Figure S3A). To check whether the effect of H1 KO on splicing was the direct outcome of the specific H1 variant depletion, we rescued the lacking variants in the 1-KO and 2-KO cells. We selected exons and introns splicing events passing quality filters (see STAR Methods) in all samples (WT, 1-KO, 1-rescue, 2-KO, and 2-rescue) and plotted their percent spliced-in (PSI) levels. Heatmaps depict inclusion



(legend on next page)

levels of these common events across all samples, highlighting the restoration of inclusion patterns upon reintroducing absent H1 in KO sample. This effect was particularly notable in 2-KO, which showed significant changes in global splicing efficiency (Figure S3B).

We recently demonstrated that two main exon-intron architectures arose during evolution. One, termed the differential architecture, has an overall low level of GC, but GC content is distributed differentially between the exons and the long flanking introns. On the other hand, the leveled architecture is composed of short introns with high GC content and similar GC content in the exons and flanking introns.[56] The leveled architecture is enriched in the nuclear center and is spliced via intron definition, whereas the differential architecture is predominantly located in the nuclear periphery and is spliced via exon definition.[57] H1s can either elevate or decrease the inclusion level of exons and introns, termed upregulated and downregulated, respectively. The exons with downregulated inclusion rates in 1-KO, 2-KO, and 4-KO cell lines are shorter than the average human exon, flanked by relatively long introns, and display characteristics consistent with those of the differential architecture (Figure 2B), implying that they can be subjected to the exon definition mode of splicing. Conversely, the upregulated exons in 1-KO, 2-KO, and 4-KO cells have shorter flanking introns but retain similarities to the differential architecture (Figure 2B). The events affected in the 5-KO resembled differential GC profile architecture, whereas its upregulated exons are shorter and downregulated exons are longer than in the unaffected group (Figure 2B).

Additionally, we incorporated micrococcal nuclease digestion with deep sequencing (MNase-seq) data on nucleosome positioning in HEK293 cells (GEO accession GSM1624500) in the groups of affected exons and introns. This analysis revealed that upregulated and downregulated exons have higher nucleosome occupancy than their flanking introns (Figure S3C). Additionally, the downregulated exons in all H1 KO cells (except in the 2-KO cells) have lower nucleosome occupancy compared with unaffected exons. Notably, the genomic regions corresponding to exon events upregulated in 5-KO cells have generally higher nucleosome occupancy in WT cells (Figure S3C).

The introns affected by H1 KO tend to be short (Figure 2Ci). In 2-KO and 4-KO, the flanking exons of the upregulated introns are shorter, whereas in 1-KO, there are no differences compared with the unaffected group, and in 5-KO, the exons are longer (Figure 2Ci). Intronic regions affected by H1 KO have generally higher nucleosome occupancy in WT cells and, in most cases, lower occupancy on at least one flanking exon (Figure S3D). The group of downregulated introns was characterized by the leveled GC architecture, whereas all except 5-KO cells upregulated introns had the differential architecture (Figure 2Cii). In the groups of upregulated exons and downregulated introns, 3′ SSs and 5′ SSs exhibit weaker strengths (Figures S3E and S3F).

We then aimed to see the extent of overlap of the affected events detected in the H1 KO samples (Figures S3G and S3H). The affected events that differ between 4-KO and 5-KO (addition of H1.4 KO) overlap to a great extent with events that are shared between 2-KO and 4-KO. The 5-KO sample shows a significant reduction in abnormal splicing compared with 4-KO (Figure 2A). Thus, the KO of H1.4 in 5-KO appears to rescue aberrant splicing events observed in 2-KO and 4-KO. This suggests that events are altered progressively with each KO, thus implying a directed and specific role of H1s both in exons and introns splicing (Figures S3Gi and S3Hi). Overlapping 5-KO samples with 1-KO events yields 23% and 18% overlap with affected exons and introns, respectively, and even less with 2-KO and 4-KO cells (Figures S3Gii and S3Hii). Therefore, events that differed in 5-KO cells are not the same as those that appeared after 1-KO, highlighting the uniqueness of each H1. Thus, defined events arise from the loss of each H1, and each variant has specific roles in splicing. Taken together, these results indicate that H1 variants are a major regulator of splicing outcomes.

Distinct splicing outcomes are associated with H1-specific binding, GC-content profile, and SS strength

The RNA-seq analysis (Figure 2) revealed a considerable effect of H1 KO on the splicing pattern. We next analyzed the association between the H1 binding around the SSs and the effect on exon and intron splicing. For this purpose, we integrated RNA-seq from the H1 KO, rescue (samples in which H1 was transfected back after KO), and WT HEK293 cell lines with ChIP-seq data and evaluated H1 variant binding in a window of ±75 nt across the SSs. We compared WT, KO, and rescue samples (see STAR Methods) and generated groups of exons and introns affected by the presence of a specific H1. To select events impacted by a specific H1 when it was KO in combination with another H1, we evaluated the events affected after both KO and rescue of the same H1 variant. We then matched affected splicing events with potential H1 target genes to maximize the likelihood of the events being directly influenced. Figure 3 describes the exon and intron groups and their inclusion dynamics while the specific H1 subtype is present. Wherever there are less than 20 exons or introns in the group with upregulated or

Figure 2. H1 variants influence splicing

(A) Stacked bar plots showing the distributions of percent spliced-in (PSI) values for alternative exons (n = 14,393), retained introns (n = 13,643), alternative 3′ SS (Alt3) (n = 861), and alternative 5′ SS (Alt5) (n = 1,064) in WT and H1 KO cell lines. The percentage of AS events within specified PSI ranges is displayed on each bar. Chi-square tests were conducted between each H1 KO sample and the WT.

(B) (Bi) Boxplots showing the average lengths of exons and flanking introns derived from groups of exons with altered inclusion levesl or unaffected after H1 KOs. (Bii) Mean GC-content (%) profiles calculated in a 75-nt window taken from the exon borders and the adjacent 150 nt of the flanking intronic sequences, excluding SS signals.

(C) (Ci) Boxplots showing the average lengths of introns and the adjacent exons derived from groups of introns with altered inclusion levels or unaffected by H1 KOs. (Cii) Mean GC-content (%) profiles calculated in a 150-nt window taken from the intron borders and the adjacent 75 nt of the flanking exonic sequences, excluding SS signals. (B and C) Exons are depicted as boxes, and introns as lines. The number of members in each group is shown in the panel. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; Wilcoxon test for all the panels except (A).

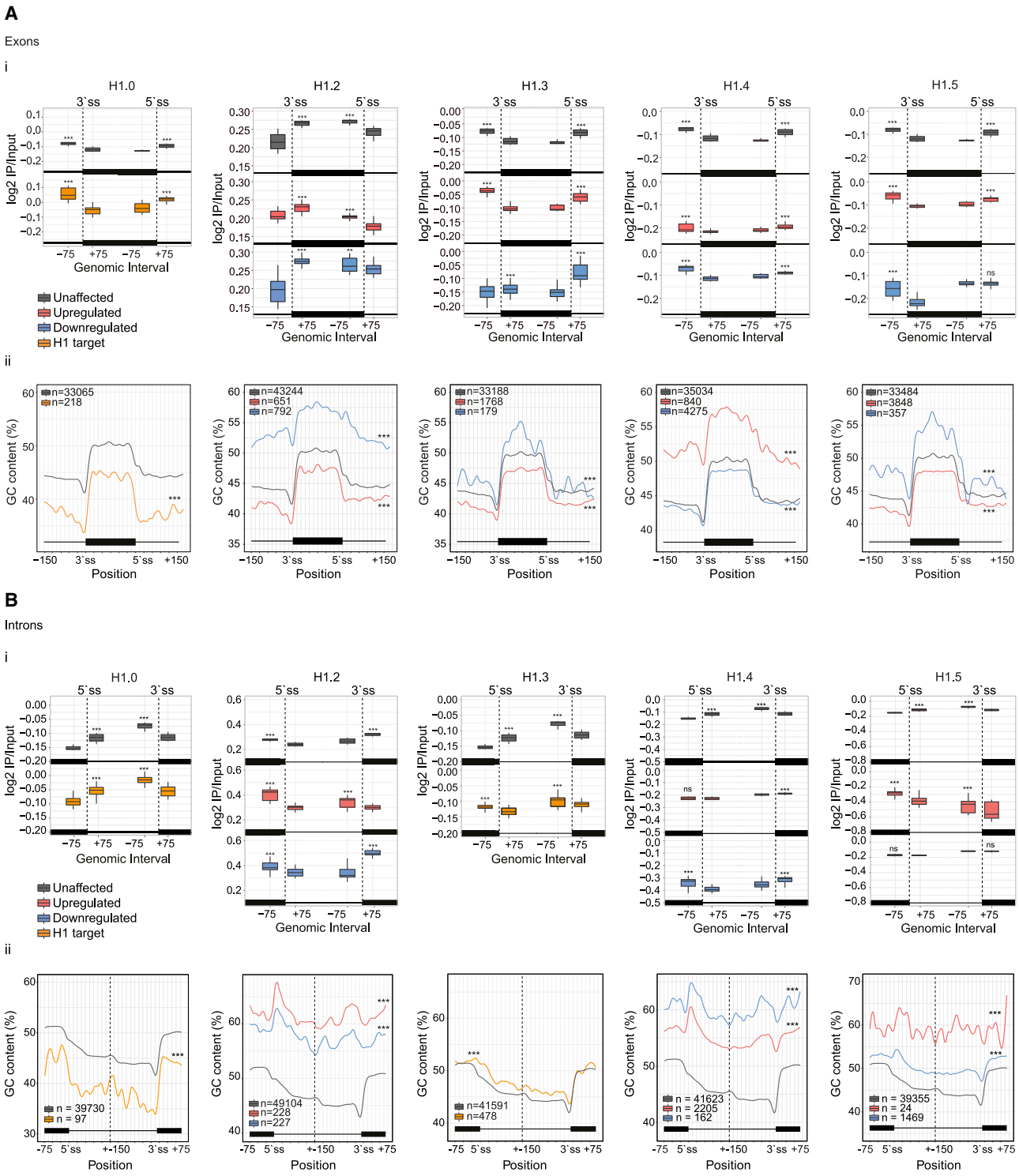


Figure 3. Differential binding of H1 variants across SSs affects exon and intron inclusion rates

(A) (Ai) Boxplots showing average log_2 IP/input ChIP-seq signal for exon groups of the indicated H1 in a window of 75 nt taken from the exon borders and the adjacent 75 nt of the flanking introns, excluding SS signals. Paired Wilcoxon test to examine the statistical significance of signal difference in intron and exon sides of each SS; asterisks mark significantly higher signals. (Aii) Mean GC-content (%) profiles calculated in 75-nt windows for sequences taken from exon borders and 150 nt of the adjacent flanking introns, excluding SS signals.

(legend continued on next page)

downregulated inclusion levels, these are merged into one group called H1 target.

In Figure 1D, we observed the tendency of histones to possess a specific mode of binding, as the most distinctive was H1.2 with its elevated signal in exons and H1.3 with peaks of signal in introns (Figure 1D). We thus examined how the mode of H1 binding to groups of splicing events affected by H1 presence is different from the general mode examined in Figure 1D. This will reveal the ability of H1 to alter the splicing outcome by shifting its positioning.

In all exons and introns whose splicing was unaffected, the signal pattern matches that presented in Figure 1D. Significantly, there were alterations in the H1.2, H1.3, and H1.5 binding modes across the SSs of exons, with decreased inclusion rates compared with the unaffected and upregulated groups. In contrast, the H1.0 and H1.4 binding patterns across SSs of exons with altered inclusion levels did not differ from those unaffected (Figure 3Ai). We compared the H1 signal intensity between the unaffected and affected groups and observed that all the H1s have higher occupancy levels with at least one of the affected event groups (Figure S4A). The SSs of all the downregulated exons affected by H1.2, H1.3, and H1.5 are weaker than those of unaffected exons (Figure S4Bi). Overall, the SSs regions of exons with higher inclusion levels (upregulated) in H1.4 presence have higher GC content, whereas the exons with lower inclusion rates (downregulated) tend to be in lower GC-content regions with differential architectures. The opposite tendency was detected for H1.2, H1.3, and H1.5 (Figure 3Aii). However, H1.3 and H1.5 are associated with upregulated levels of exon inclusion. H1.4 has a higher signal in the group of downregulated exons, highlighting its role in promoting ES (Figure S4Ai). These results imply that H1 variants regulate the inclusion rate of exons due to a specific binding profile across the SSs, GC-content environment, and SSs strengths, which are also contributing factors.

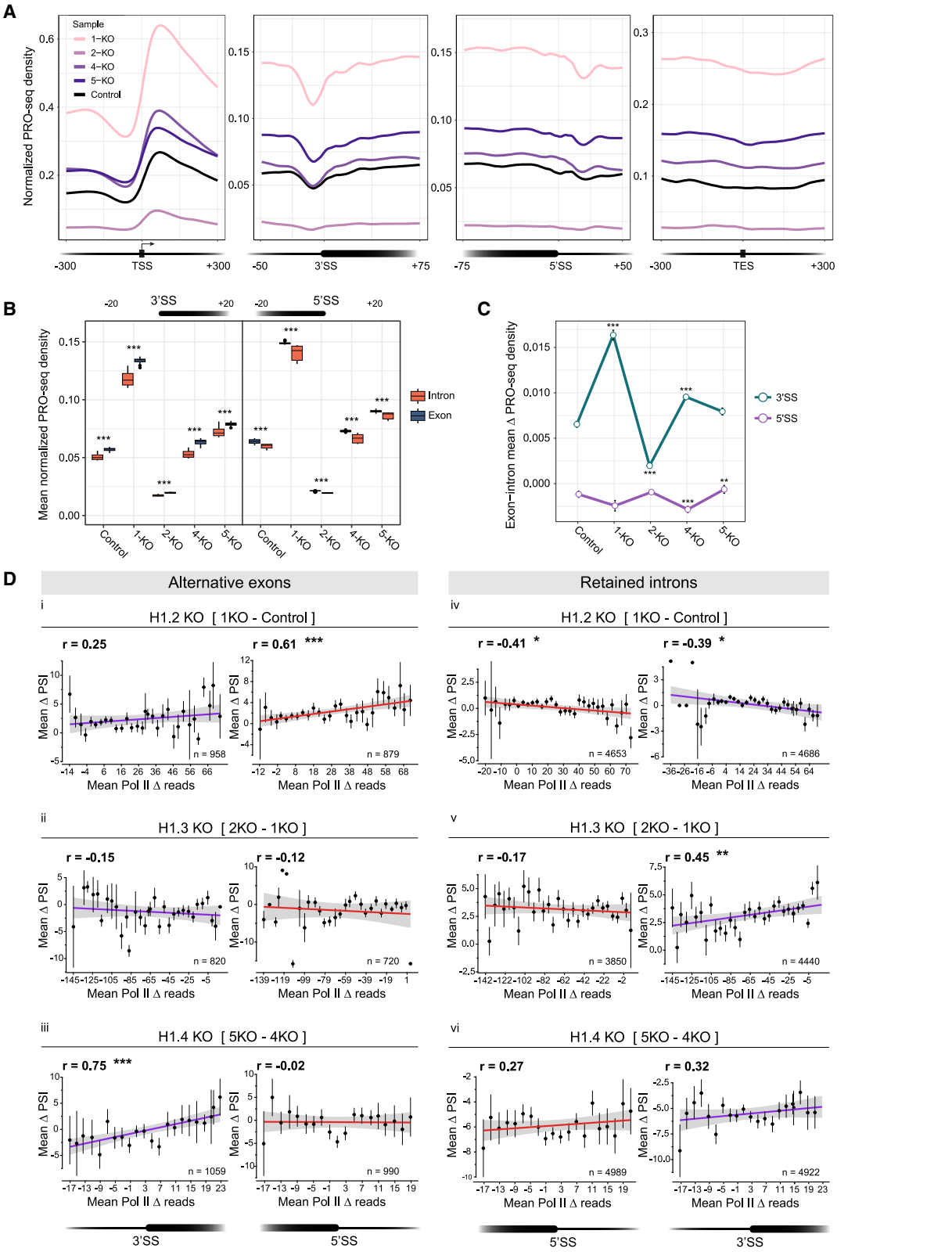
The binding pattern of all H1 variants except H1.0 differed in SS regions of introns, with retention levels altered by fluctuations in H1 availability compared with unaffected introns (Figure 3B). Although the association of H1.0 with affected splicing events is significantly stronger than with the unaffected group, indicating its influence on splicing, this effect does not appear to result from a shift in its binding pattern to SSs (Figure S4A). H1.4, which usually binds intron sequences (Figure 1D), was enriched in the exons flanking affected introns (Figure 3Bi). In addition, H1.2 binding was enriched on the 3′ SS intronic side of the upregulated introns, whereas H1.3 binding was enriched on the exon side of the affected introns' 5′ SS, implying that when H1.2 and H1.3 bind to sites outside of their preferential ones (exons for H1.2 and introns for H1.3; Figure 1D), they affect inclusion levels of introns. H1.3 predominantly affected the intron inclusion rate negatively, with only 7 of 478 affected introns upregulated in the presence of H1.3. H1.2 and, to some extent, H1.3 have distinctly higher signals on SS regions of the affected intron group (Figure S4Aii). The SS score of the analyzed introns is either lower or does not differ significantly from the unaffected group for any H1 variants (Figure S4Bii). Analysis of the GC-content environment of the affected introns showed that H1.0 and H1.3 affect the ones located in low GC-content regions with differential architecture, whereas H1.2 and H1.4 affect introns that reside in high GC-content regions with leveled architecture (Figure 3Bii). H1.5 downregulates retention of introns with relatively differential architecture while upregulating inclusion of introns from the leveled.

We aimed to elucidate the functionality of genes whose exon inclusion rates are affected by H1 presence. Gene ontology (GO) analyses revealed biological processes like DNA repair, regulation of transcription, and protein transport to be shared among most H1s (Figure S4C). Regulation of exon inclusion levels by H1.2 and H1.4 affects genes involved in protein phosphorylation, whereas H1.3 controls splicing of genes that participate in peptidyl-serine phosphorylation. These proteins control a wide variety of cell processes, such as signal transduction, cell cycle regulation, RNAPII elongation control, and more.[59,60] Overall, these results show that each H1 variant regulates splicing of genes with unique features.

H1s regulate splicing by modulating the RNAPII elongation pattern

Splicing predominantly occurs co-transcriptionally, and fluctuations in RNAPII elongation rate have previously been shown to influence AS.[34,61–63] Considering H1 variant-specific effects on splicing, we investigated their potential involvement in RNAPII elongation modulation. We conducted precision run-on sequencing (PRO-seq) experiments on each of our H1 KO samples and WT cells. The stepwise KOs enabled us to distinguish the effect of H1s on the occupancy of transcriptionally engaged RNAPII along gene bodies, including TSSs, exon-intron junctions, and TESs. Peaks in PRO-seq signal align with RNAPII pausing[62,64] and diminished PRO-seq densities denote acceleration,[62] thus the observed changes in RNAPII density in H1 KO samples compared with WT reflect definitive shifts in RNAPII elongation rates (Figure 4A). Our results revealed that H1 variants are not functionally redundant. The deletion of different combinations of H1 variants can lead to either an increase or a decrease in RNAPII elongation rate, with varying intensities of effect. The most noticeable difference was observed between samples 1-KO and 2-KO. The single KO of H1.2 drastically reduced the elongation rate of RNAPII, whereas cells lacking both H1.2 and H1.3 had a higher RNAPII elongation rate than WT cells. The 4-KO sample, which lacks H1.0, H1.2, H1.3, and H1.5, showed a PRO-seq density pattern across exon-intron junctions similar to that of WT cells. We conducted a PRO-seq metagene analysis and examined the impact of H1 KOs on promoter-proximal pausing and pausing index, which is the ratio of PRO-seq reads in the initiation region compared with reads in the gene body[65–68] (Figures S5A and S5B). The pausing index provides insights into the frequency and duration of transcriptional

(B) (Bi) Boxplots showing the same analysis as described in (A) for intron groups. (Bii) Mean GC-content (%) profiles for intron groups as described in (A). Values were calculated in 75-nt windows of the adjacent exons and 150 nt from intron borders, excluding SS signals. (A and B) Exons are illustrated as black boxes, and introns are depicted as lines. The number of members in each group is shown in the panel. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; Wilcoxon test.



(legend on next page)

pauses by polymerases. A higher pausing index typically correlates with lower gene activity.[65] Our findings indicate that the deletion of H1.3 in the 2-KO sample and H1.4 in the 5-KO sample reduced RNAPII pausing (Figure S5B), implying that these specific H1 variants may regulate gene activity.

Plotting the RNAPII density profile on selected highly expressed housekeeping genes revealed a non-uniform elongation behavior across the gene structure, with a unique pattern observed per gene (Figure S5C). In sum, maintaining a precise transcription elongation pattern requires a delicate interplay between specific H1 variants.

Previous PRO-seq analyses reported fluctuations in RNAPII densities between the 3′ and the 5′ SSs and their upstream regions, with the latter having no significant change.[62] We thus analyzed RNAPII elongation activity at intron-exon junctions in H1 KO cell lines and WT cells. Focusing on a 40-nt window spanning each SS (20 nt each of the exon and intron), we detected significant dips in PRO-seq densities in both the 3′ and the 5′ SSs in all samples (Figure 4B). To determine whether H1s control the magnitude of these dips, we calculated the mean exon-intron differential PRO-seq coverage of over 200,000 exons, comparing each KO sample with WT cells (Figure 4C). H1s differentially regulate RNAPII elongation at the 3′ and 5′ SSs. In 1-KO, 2-KO, and 4-KO cells, deceleration of RNAPII (shown in Figure 4A) was accompanied by larger PRO-seq dips at 3′ SS and vice versa (Figure 4C, $p < 0.0001$). The effect of H1s on the RNAPII rate at 5′ SSs was the opposite, albeit more subtle. For instance, we found that in the 5-KO cells, a reduction in RNAPII rate was linked to an increase in PRO-seq dips (Figure 4C, $p < 0.01$). When the 3′ SSs are stronger, the elongation rate is more rapid across intron-exon junctions in all samples (Figure S5D), implying the importance of the polypyrimidine tract on the rate of RNAPII.

To examine H1-mediated effects on RNAPII elongation as a regulatory mechanism impacting co-transcriptional splicing decisions, we integrated PRO-seq and RNA-seq data. We then correlated shifts in PRO-seq read densities across SSs with alterations in exon and intron inclusion levels resulting from the lack of H1s. Comparing pairs of samples that differed by a single KO, we isolated the impact of H1.2, H1.3, and H1.4 (Figure 4D). Specific variants are adapted for the co-transcriptional splicing regulation of exons and/or introns, affecting either a single SS or both. For example, H1.2 KO-induced changes in RNAPII rate significantly affect the inclusion of both exons and introns. Specifically, H1.2 affects the inclusion of exons through the 5′ SS with a strong positive correlation ($r = 0.61$, $p < 0.001$; Figure 4Di), whereas its effect on IR involved both SSs with relatively moderate, yet significant, negative correlations ($r = -0.41$, $p < 0.05$ for the 5′ SS and $r = -0.39$, $p < 0.05$ for the 3′ SS; Figure 4Dii). The shifts in RNAPII rate that result from H1.3 and H1.4 KOs are associated with the alterations in the splicing of introns and exons through the 3′ SSs ($r = 0.45$, $p < 0.001$ for H1.3 and $r = 0.75$, $p < 0.0001$ for H1.4; Figures 4Diii and 4Dv). These data establish distinct roles for different H1 variants in modulating elongation rates, thereby regulating co-transcriptional splicing.

RNAPII elongation and splicing are regulated by interaction with H1.2

To gain insights into the mechanism by which H1s control the elongation of RNAPII and splicing, we made a deletion in either the C- or the N-terminal domain of H1.2. The H1.2 variant was selected as it is the most highly expressed H1 subtype in HEK293 cells (Figure S1E) and has the most robust effect on the RNAPII elongation and splicing of the H1s evaluated (Figures 2, 3, and 4). The terminal domain sequences are not conserved across H1s and may result in variant-specific functionalities.[13,69] We performed co-immunoprecipitation of RNAPII and HA-H1.2 with and without deletions in its terminal domains (Figure 5A). H1.2 interacts with elongating serine 2-phosphorylated RNAPII.[70] This interaction is independent of the functional C- and N-terminals of H1.2 (Figures 5A and 5B). These results imply that RNAPII binding can be performed by both tails or the globular domain of H1.2.

We aimed to extend our understanding of H1-mediated effects on RNAPII elongation and splicing. We focused on six genes where RNA-seq and PRO-seq analyses identified alterations in RNAPII density and exon inclusion levels but did not show major variation in global transcription level following H1.2 KO. Additionally, H1.2 is present on these exons by ChIP-seq. We conducted RT-qPCR in H1.2 KO cells and HA-H1.2-expressing rescue cells with three primer sets on the same exon: one spanning exon-intron junctions to measure expression changes, another spanning exon-exon junctions for exon inclusion, and a third for exon exclusion analysis. Expression differences were assessed using 5,6-dichlorobenzimidazole 1-β-d-ribofuranoside (DRB) to block RANPII transition from initiation to elongation. We restored RNAPII activity simultaneously in all cells by washing out of DRB after 3 h. For all except one (*PCGF2*), there were statistically significant differences in expression in cells lacking H1.2 (Figure 5C). Efficiencies of splicing of exons in *PCGF2*, *TOX2*, and *SLC46A* also differed in the presence and absence of H1.2 (Figure 5D). These results indicate that H1.2 directly affects RNAPII density and splicing of specific exons. Next, we performed the experiment

Figure 4. PRO-seq analyses of the H1-related effect on transcription elongation rate and splicing

(A) Mean PRO-seq 3′ end coverage (single nucleotide resolution) along gene bodies: TSSs, exon-intron junctions (75 nt into the exon and 50 nt into the adjacent intron), and TESs in H1 KOs and WT.

(B) Boxplots describing the mean PRO-seq density comparing exonic and intronic sequences around 3′ SSs (left) and 5′ SSs (right) (20 nt into the intron and exon, indicated above). Wilcoxon rank sum tests.

(C) Line plot with mean and standard error of delta exon-intron PRO-seq density around the 3′ SS and the 5′ SS, determined by subtracting the mean read coverage of 20 nt in each intron from that of the adjacent exon (n = 205,561). Wilcoxon rank sum test was applied between each H1s KO sample and the WT.

(D) The relative effects of H1.2, H1.3, and H1.4 (comparing two consecutive KO samples or the WT cells) on the inclusion levels of alternative exons (left) and retained introns (right). Pearson's correlation tests (two-tailed) were performed between the mean delta PSI values and mean delta PRO-seq densities within 3′ SSs and 5′ SSs (10 nt into the exon and 20 nt into the adjacent intron). Regression lines with standard errors are shown in purple and red for the 3′ SS and 5′ SS, respectively, with n number of events. (A–D) Spike-in normalization was applied to all samples. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

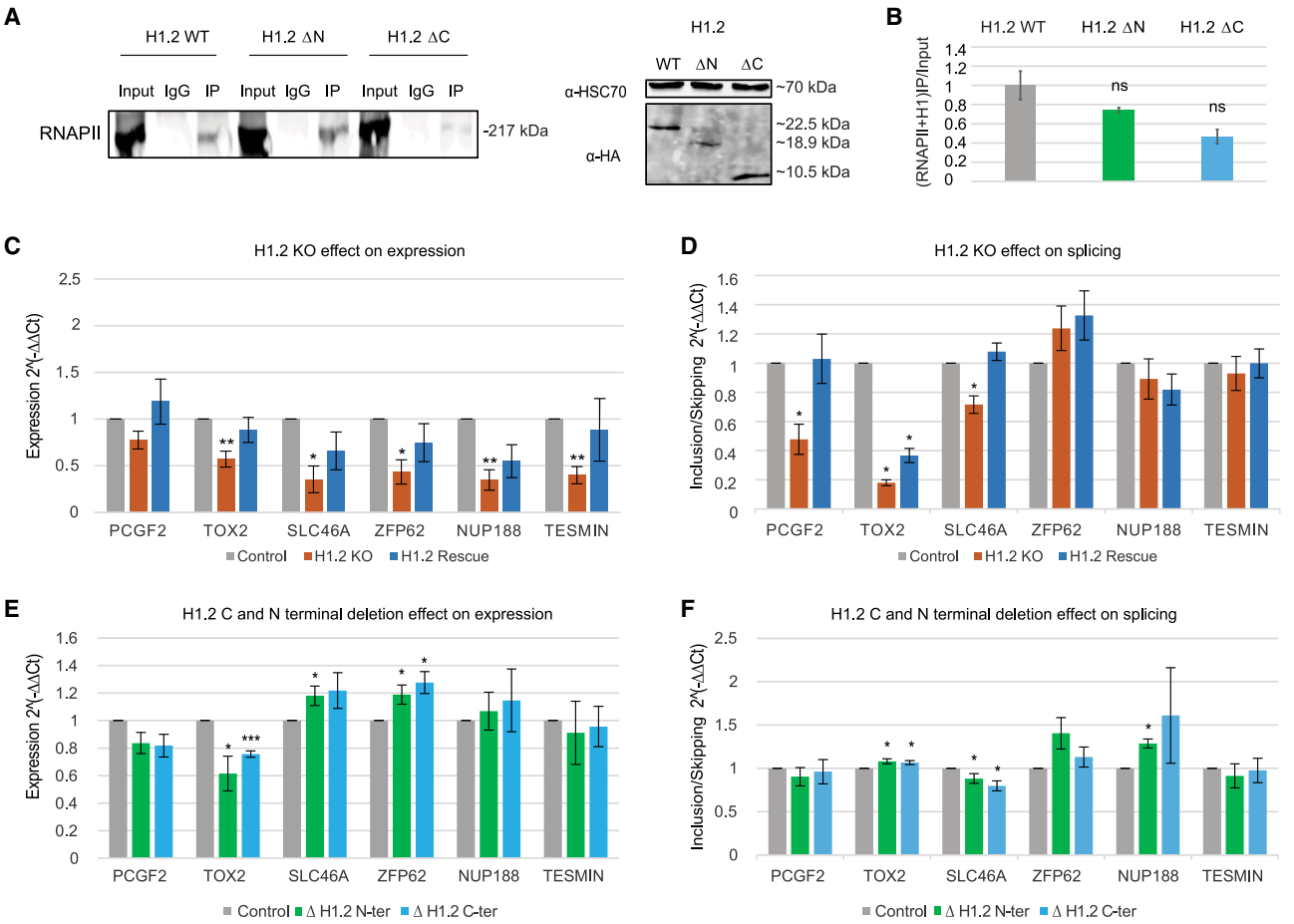


Figure 5. H1.2 associates with RNAPII and affects elongation and splicing

(A) Left: representative western blot (WB) analysis of serine 2 (Ser2) phosphorylated RNAPII co-immunoprecipitated with an antibody against the HA tag of WT H1.2-HA, and H1.2-HA with a deletion in the N terminus (H1.2 ΔN) or C terminus (H1.2 ΔC), IgG as a negative control. Input corresponds to cell lysate. Right: WB analysis of H1.2 WT, H1.2 ΔN, and H1.2 ΔC with anti-HSC-70, and anti-HA antibodies. 40 micrograms of total protein loaded per lane.

(B) Quantification of RNAPII precipitated with indicated H1.2 relative to input based on WB from (A). Plotted are means ± SEM; n = 3. Two-tailed t test between the WT H1.2 and H1.2 with deletions (p = 0.48; p = 0.22, respectively).

(C–F) Bar plots displaying gene expression and splicing ratios (exon inclusion/exon skipping) of representative endogenous genes from PRO-seq and RNA-seq analysis. (C and D) upon KO of H1.2 cells and during rescue by transfection of H1.2 into H1.2 KO cells, compared with the WT HEK293. (E and F) upon expression of H1.2 with deletion in N- (33 aa) or C-terminal (102 aa) in H1.2 KO cells, compared with the control (H1.2 KO cells transfected with H1.2). Gene expression level is normalized to the housekeeping gene *RPLO*. p values correspond to two-tailed t tests: ***p < 0.001; **p < 0.01; *p < 0.05.

with H1.2 KO cells that express H1.2 proteins that lacked a functional N- or C-terminal domain. Both domains are important for regulation of RNAPII elongation (Figure 5E), and both influence inclusion levels of the exons in the same genes (Figure 5F).

In summary, the putative mechanism of H1.2 influence on RNAPII elongation is mediated by H1.2 direct association with serine 2-phosphorylated form of the RNAPII. There is a link between the effect of H1.2 on elongation of RNAPII and splicing. Moreover, the C- and N-terminal domains of H1.2 are both important for this regulation.

DISCUSSION

We interrogated the previously elusive roles of histone H1 variants in the regulation of RNAPII elongation and splicing. The changes in RNAPII elongation that result from the loss of H1s were strongly associated with changes in splicing. Our PRO-seq analyses showed that H1 variants strongly influence the elongation process, presumably via direct interactions between RNAPII and the H1s.[39,50] H1s have been found to interact with multiple splicing factors. For example, H1.0 directly interacts with SF2/ASF and U2AF65,[40] which may be loaded onto pre-mRNA during transcription.[71–73]

During exon transcription, there is a gradual increase in Ser5 phosphorylation on RNAPII's C-terminal domain. The exon-to-intron transition reduces Ser5 phosphorylation, making Ser2 phosphorylation dominant during intron synthesis, indicating a shift from RNAPII deceleration to acceleration.[74] In support of this, we showed that during exon synthesis, RNAPII slows slightly but that after passing the 5′ SS, elongation accelerates.

In cells that lack both H1.2 and H1.3, this trend was almost abolished.

The influence of H1s on elongation by RNAPII, along with specific binding to the SS regions, underlies the effects of H1s on co-transcriptional splicing. Chromatin architecture may influence the RNAPII elongation rate and thus influence AS.[2,75] H1 variants may force 3D architectural changes in chromatin through regulation of compaction, and proper H1 balance is needed to maintain the chromatin structure.[10,76] According to the kinetic model, the modulation of RNAPII elongation may affect splicing via the differential formation of specific pre-mRNA secondary structures.[77] Recent findings showed that RNAPII activity affects AS of exons with adjacent pausing architectures.[78] H1-mediated chromatin compaction may thus influence the splicing decision by modulating chromatin organization, which in turn affects the elongation rate of RNAPII and thus SS recognition during AS. This suggests a possible mechanism by which H1 position shift may provoke the distinct splicing outcome.

Despite KO of multiple H1 variants, H1.2 and H1.X nuclear distributions remained preserved compared with WT cells. In contrast, H1.5 showed relative redistribution, possibly due to the tag addition. However, it has been shown that H1.5 is a dynamic histone with a distribution that is dependent on the differentiation state of the cell.[79,80] Therefore, it is possible that H1.5 binds to additional regions in the absence of other H1 variants. H1 variants were also found to be grouped according to overall affinity to GC level. We found the general patterns of H1 binding in HEK293 cells, with H1.X and H1.2 being enriched in the exonic portion spanning the SS region and H1.0, H1.3, and H1.4 being preferentially positioned on the intron portion, with H1.5 being more distinctly placed in the SS regions themselves. We observed the association between the shift in global H1 mode of binding and the effect on splicing observed while these H1 variants are KO. Meaning, H1 can alternate its binding mode for defined groups of splicing events, modulating the splicing outcome. H1.0 exhibited a strong association with affected exons and introns, all the while retaining its binding mode. This suggests that H1.0 might influence splicing outcomes by recruiting splicing factors.[81] H1 variants displayed elevated signals within the splicing events impacted by the H1 presence, emphasizing their involvement in the observed inclusion alterations.

The level of splicing events altered by the absence of H1 variants gradually increased in 1-KO, 2-KO, and 4-KO cells, with thousands of events altered in the 4-KO cells. Each H1 regulates specific splicing events, and in KO cells, compensation by other variants is limited. The loss of H1 variants had effects on splicing that impacted cell division, cell cycle progression, and DNA repair. Recent studies showed that cell cycle progression defects, among others, are associated with abnormal splicing.[82,83] Restored expression of H1 variants that were knocked out was sufficient to abrogate both splicing and cell cycle-related effects that arose due to the lack of specific H1s.

The most intriguing finding is the spontaneous recovery of the aberrant splicing when five H1 were KO. Although the exact reasons for this are challenging to determine, we have several possible explanations. Our results indicate that there is no complete overlap between the events affected in the 5-KO and 1-KO cells. The majority of events with significantly altered inclusion levels between 4-KO and 5-KO cells predominantly overlap with the shared events that emerged after the KO of two histones (2-KO) and four histones (4-KO). Therefore, we suggest that the effect seen in the 5-KO cells on splicing may be a consequence of dysregulation of chromatin architecture provoked by the KO of the five H1 variants or the unique function of H1.4 histone in splicing, the one that was KO in the shift from 4-KO to 5-KO.

We demonstrated the association between the higher signal of H1.4 within SS regions and downregulation of exon inclusion. Among the examined H1, this histone possesses a significant number of signal peaks within promoter regions. Additionally, H1.4 KO led to alterations in RNAPII elongation, which was associated with aberrant splicing. The phosphorylation status of H1.4 is known to impact the interaction of splicing factors with nucleosomes, potentially influencing its catalytic activity.[84] Previous research has shown the association of H1.4 with RNAPII and RNAPI.[50,85,86] Considering all this evidence, H1.4's role in co-transcriptional splicing may be the key trigger for the changes that occurred in the shift from 4-KO to 5-KO.

Another possible explanation for splicing restoration may be the lack of H1s in shaping the splicing outcome. KO of H1 variants may lead to a more "primitive" form of splicing, which allows RNAPII and the splicing machinery to detect and recognize exons and introns even in the absence of five H1s, for example, via marking of exons by nucleosomes.[16,28,30,87] However, 5-KO cells are severely compromised in terms of their viability, replicability, and other biological characteristics, indicating that the cell cannot effectively sustain the loss of five H1 variants.

We also examined the importance of the non-conserved H1.2 N- and C-terminal domains on RNAPII elongation and splicing. Evidence suggests that these domains fine-tune the binding affinity of H1 to the nucleosome and DNA.[88,89] The H1 N- and C-terminal domains are subjected to various post-translational modifications that are in turn involved in the regulation of chromatin packaging, apoptosis, and transcription.[90–92] We showed that both N- and C-terminal domains of H1.2 were important for controlling elongation and splicing in most of the analyzed genes. Therefore, a possible regulatory mechanism of H1.2 on RNAPII elongation involves both physical interference and interactions initiated by H1-disordered tails.

Recent studies report that H1 distribution is very complex and depends on histone linker-specific modifications, cell differentiation stage, and cell type.[50] Thus, H1s are possible regulators of cell identity. H1s may promote epigenetic modifications or recruit splicing factors in a variant-specific fashion. Our results indicate that H1 histones are major regulators of RNAPII elongation and that these effects in turn influence AS patterns in thousands of genes. In regions spanning the SSs, there are significant changes in the GC-content level,[56] which impacts H1 binding affinities. We propose that two main mechanisms underlie H1 variant-specific co-transcriptional splicing regulation: direct interaction of the H1 variants with RNAPII and H1-mediated recruitment of splicing-related factors.

Limitations of the study

ChIP-seq experiments were performed on H1 histones with an HA tag that was exogenously introduced into cells. Thus, we cannot rule out the possibility that the addition of the tag to H1

influenced its binding preferences. Additional research will be required to elucidate the relatively lesser impact on splicing in the 5-KO condition.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Cell maintenance
  - Cloning of sgRNAs
  - CRISPR/Cas9-mediated generation of H1-KO and rescue cell lines
- METHOD DETAILS
  - Protein extraction
  - Western blot and antibodies
  - Co-immunoprecipitation
  - MNase digestion
  - Cell cycle and replication rate evaluation
  - cDNA synthesis and qPCR
  - RNA-seq analyses
  - Nucleosome occupancy
  - ChIP-seq analyses
  - Dataset preparation and grouping
  - Profile plot construction
  - Integration between RNA-seq and ChIP-seq
  - Plotting GC levels over the chromosomes
  - PRO-seq library preparation
  - Pre-processing and analyses of PRO-seq experiments
  - RT-PCR and DRB treatment
  - Gene ontology
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2023.10.003.

ACKNOWLEDGMENTS

The research was funded by the Israel Science Foundation (ISF 671/18, ISF 142/13, and ISF 2417/20); the Israel Cancer Research Foundation (ICRF PG-18-105, PG-20-104); and the United States – Israel Binational Science Foundation (BSF 2017086). V.R.R. was supported by Edmond J. Safra Bioinformatics Center fellowship at Tel Aviv University. We thank K. Adelman from Blavatnik Institute, Harvard Medical School, Boston, USA, for the PRO-seq analysis protocols. We also thank I. Ulitsky from the Weizman Institute of Science, Rehovot, Israel, for consultations on the PRO-seq analysis. We thank R. Shayevitch for the PRO-seq protocol design.

AUTHOR CONTRIBUTIONS

C.P., O.H., and S.B.-S. designed and performed bioinformatic analyses. J.Z., C.G.-L., L.T., G.L.-M., S.A.R., T.L., M.E., N.S., L.E., A.T., and T.H. designed and performed biological experiments. A.J. and E.M. advised and guided the design of performed analyses. C.P., J.Z., O.H., V.R.R., and G.A. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: March 21, 2023
Revised: August 17, 2023
Accepted: October 5, 2023
Published: November 2, 2023

REFERENCES

1. Perino, M., and Veenstra, G.J.C. (2016). Chromatin control of developmental dynamics and plasticity. Dev. Cell *38*, 610–620. https://doi.org/10.1016/j.devcel.2016.08.004.
2. Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. Cell *128*, 707–719. https://doi.org/10.1016/j.cell.2007.01.015.
3. Kornblihtt, A.R., Schor, I.E., Allo, M., and Blencowe, B.J. (2009). When chromatin meets splicing. Nat. Struct. Mol. Biol. *16*, 902–903. https://doi.org/10.1038/nsmb0909-902.
4. Harshman, S.W., Young, N.L., Parthun, M.R., and Freitas, M.A. (2013). H1 histones: current perspectives and challenges. Nucleic Acids Res. *41*, 9593–9609. https://doi.org/10.1093/nar/gkt700.
5. Kasinsky, H.E., Lewis, J.D., Dacks, J.B., and Ausió, J. (2001). Origin of H1 linker histones. FASEB J. *15*, 34–42. https://doi.org/10.1096/fj.00-0237rev.
6. Izzo, A., Kamieniarz, K., and Schneider, R. (2008). The histone H1 family: specific members, specific functions? Biol. Chem. *389*, 333–343. https://doi.org/10.1515/BC.2008.037.
7. Happel, N., and Doenecke, D. (2009). Histone H1 and its isoforms: contribution to chromatin structure and function. Gene *431*, 1–12. https://doi.org/10.1016/j.gene.2008.11.003.
8. Parseghian, M.H., Newcomb, R.L., and Hamkalo, B.A. (2001). Distribution of somatic H1 subtypes is non-random on active vs. inactive chromatin II: distribution in human adult fibroblasts. J. Cell. Biochem. *83*, 643–659. https://doi.org/10.1002/jcb.1224.
9. Bustin, M., Catez, F., and Lim, J.-H. (2005). The dynamics of histone H1 function in chromatin. Mol. Cell *17*, 617–620. https://doi.org/10.1016/j.molcel.2005.02.019.
10. Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultchi, A.I. (2005). Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. Cell *123*, 1199–1212. https://doi.org/10.1016/j.cell.2005.10.028.
11. Shen, X., and Gorovsky, M.A. (1996). Linker histone H1 regulates specific gene expression but not global transcription in vivo. Cell *86*, 475–483. https://doi.org/10.1016/S0092-8674(00)80120-8.
12. Sancho, M., Diani, E., Beato, M., and Jordan, A. (2008). Depletion of human histone H1 variants uncovers specific roles in gene expression and cell growth. PLoS Genet. *4*, e1000227. https://doi.org/10.1371/journal.pgen.1000227.
13. Hergeth, S.P., and Schneider, R. (2015). The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. EMBO Rep. *16*, 1439–1453. https://doi.org/10.15252/embr.201540749.
14. Ast, G. (2004). How did alternative splicing evolve? Nat. Rev. Genet. *5*, 773–782. https://doi.org/10.1038/nrg1451.

15. Irimia, M., and Blencowe, B.J. (2012). Alternative splicing: decoding an expansive regulatory layer. Curr. Opin. Cell Biol. *24*, 323–332. https://doi.org/10.1016/j.ceb.2012.03.005.

16. Schwartz, S., Meshorer, E., and Ast, G. (2009). Chromatin organization marks exon-intron structure. Nat. Struct. Mol. Biol. *16*, 990–995. https://doi.org/10.1038/nsmb.1659.

17. Naftelberg, S., Schor, I.E., Ast, G., and Kornblihtt, A.R. (2015). Regulation of alternative splicing through coupling with transcription and chromatin structure. Annu. Rev. Biochem. *84*, 165–198. https://doi.org/10.1146/annurev-biochem-060614-034242.

18. Allemand, E., Myers, M.P., Garcia-Bernardo, J., Harel-Bellan, A., Krainer, A.R., and Muchardt, C. (2016). A broad set of chromatin factors influences splicing. PLoS Genet. *12*, e1006318. https://doi.org/10.1371/journal.pgen.1006318.

19. Tilgner, H., and Guigó, R. (2010). From chromatin to splicing: RNA-processing as a total artwork. Epigenetics *5*, 180–184. https://doi.org/10.4161/epi.5.3.11319.

20. Iannone, C., Pohl, A., Papasaikas, P., Soronellas, D., Vicent, G.P., Beato, M., and ValcáRcel, J. (2015). Relationship between nucleosome positioning and progesterone-induced alternative splicing in breast cancer cells. RNA *21*, 360–374. https://doi.org/10.1261/rna.048843.114.

21. Luco, R.F., and Misteli, T. (2011). More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. Curr. Opin. Genet. Dev. *21*, 366–372. https://doi.org/10.1016/j.gde.2011.03.004.

22. de Almeida, S.F., and Carmo-Fonseca, M. (2012). Design principles of interconnections between chromatin and pre-mRNA splicing. Trends Biochem. Sci. *37*, 248–253. https://doi.org/10.1016/j.tibs.2012.02.002.

23. Alló, M., Schor, I.E., Muñoz, M.J., de la Mata, M., Agirre, E., Valcárcel, J., Eyras, E., and Kornblihtt, A.R. (2010). Chromatin and alternative splicing. Cold Spring Harb. Symp. Quant. Biol. *75*, 103–111. https://doi.org/10.1101/sqb.2010.75.023.

24. Bonnal, S.C., López-Oreja, I., and Valcárcel, J. (2020). Roles and mechanisms of alternative splicing in cancer - implications for care. Nat. Rev. Clin. Oncol. *17*, 457–474. https://doi.org/10.1038/s41571-020-0350-x.

25. Iannone, C., and Valcárcel, J. (2013). Chromatin's thread to alternative splicing regulation. Chromosoma *122*, 465–474. https://doi.org/10.1007/s00412-013-0425-x.

26. Chen, W., Luo, L., and Zhang, L. (2010). The organization of nucleosomes around splice sites. Nucleic Acids Res. *38*, 2788–2798. https://doi.org/10.1093/nar/gkq007.

27. Nahkuri, S., Taft, R.J., and Mattick, J.S. (2009). Nucleosomes are preferentially positioned at exons in somatic and sperm cells. Cell Cycle *8*, 3420–3424. https://doi.org/10.4161/cc.8.20.9916.

28. Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C., and Komorowski, J. (2009). Nucleosomes are well positioned in exons and carry characteristic histone modifications. Genome Res. *19*, 1732–1741. https://doi.org/10.1101/gr.092353.109.

29. Spies, N., Nielsen, C.B., Padgett, R.A., and Burge, C.B. (2009). Biased chromatin signatures around polyadenylation sites and exons. Mol. Cell *36*, 245–254. https://doi.org/10.1016/j.molcel.2009.10.008.

30. Tilgner, H., Nikolaou, C., Althammer, S., Sammeth, M., Beato, M., Valcárcel, J., and Guigó, R. (2009). Nucleosome positioning as a determinant of exon recognition. Nat. Struct. Mol. Biol. *16*, 996–1001. https://doi.org/10.1038/nsmb.1658.

31. Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nat. Rev. Mol. Cell Biol. *18*, 637–650. https://doi.org/10.1038/nrm.2017.63.

32. Kornblihtt, A.R. (2006). Chromatin, transcript elongation and alternative splicing. Nat. Struct. Mol. Biol. *13*, 5–7. https://doi.org/10.1038/nsmb0106-5.

33. Godoy Herz, M.A., and Kornblihtt, A.R. (2019). Alternative splicing and transcription elongation in plants. Front. Plant Sci. *10*, 309. https://doi.org/10.3389/fpls.2019.00309.

34. Dujardin, G., Lafaille, C., de la Mata, M., Marasco, L.E., Muñoz, M.J., Le Jossic-Corcos, C., Corcos, L., and Kornblihtt, A.R. (2014). How slow RNA polymerase II elongation favors alternative exon skipping. Mol. Cell *54*, 683–690. https://doi.org/10.1016/j.molcel.2014.03.044.

35. Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., Diener, K., Jones, K., Fu, X.-D., and Bentley, D.L. (2014). Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. Genes Dev. *28*, 2663–2676. https://doi.org/10.1101/gad.252106.114.

36. Howe, K.J., Kane, C.M., and Ares, M. (2003). Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA *9*, 993–1006. https://doi.org/10.1261/rna.5390803.

37. Jimeno-González, S., Payán-Bravo, L., Muñoz-Cabello, A.M., Guijo, M., Gutierrez, G., Prado, F., and Reyes, J.C. (2015). Defective histone supply causes changes in RNA polymerase II elongation rate and cotranscriptional pre-mRNA splicing. Proc. Natl. Acad. Sci. USA *112*, 14840–14845. https://doi.org/10.1073/pnas.1506760112.

38. Leader, Y., Lev Maor, G., Sorek, M., Shayevitch, R., Hussein, M., Hameiri, O., Tammer, L., Zonszain, J., Keydar, I., Hollander, D., et al. (2021). The upstream 5′ splice site remains associated to the transcription machinery during intron synthesis. Nat. Commun. *12*, 4545. https://doi.org/10.1038/s41467-021-24774-6.

39. Kim, K., Lee, B., Kim, J., Choi, J., Kim, J.-M., Xiong, Y., Roeder, R.G., and An, W. (2013). Linker histone H1.2 cooperates with Cul4A and PAF1 to drive H4K31 ubiquitylation-mediated transactivation. Cell Rep. *5*, 1690–1703. https://doi.org/10.1016/j.celrep.2013.11.038.

40. Kalashnikova, A.A., Winkler, D.D., McBryant, S.J., Henderson, R.K., Herman, J.A., DeLuca, J.G., Luger, K., Prenni, J.E., and Hansen, J.C. (2013). Linker histone H1.0 interacts with an extensive network of proteins found in the nucleolus. Nucleic Acids Res. *41*, 4026–4035. https://doi.org/10.1093/nar/gkt104.

41. Glaich, O., Leader, Y., Lev Maor, G., and Ast, G. (2019). Histone H1.5 binds over splice sites in chromatin and regulates alternative splicing. Nucleic Acids Res. *47*, 6145–6159. https://doi.org/10.1093/nar/gkz338.

42. Jedrusik, M.A., and Schulze, E. (2001). A single histone H1 isoform (H1.1) is essential for chromatin silencing and germline development in Caenorhabditis elegans. Development *128*, 1069–1080. https://doi.org/10.1242/dev.128.7.1069.

43. Wierzbicki, A.T., and Jerzmanowski, A. (2005). Suppression of histone H1 genes in Arabidopsis results in heritable developmental defects and stochastic changes in DNA methylation. Genetics *169*, 997–1008. https://doi.org/10.1534/genetics.104.031997.

44. Fan, Y., Nikitina, T., Morin-Kensicki, E.M., Zhao, J., Magnuson, T.R., Woodcock, C.L., and Skoultchi, A.I. (2003). H1 linker histones are essential for mouse development and affect nucleosome spacing in vivo. Mol. Cell. Biol. *23*, 4559–4572. https://doi.org/10.1128/MCB.23.13.4559-4572.2003.

45. Fan, Y., Sirotkin, A., Russell, R.G., Ayala, J., and Skoultchi, A.I. (2001). Individual somatic H1 subtypes are dispensable for mouse development even in mice lacking the H1(0) replacement subtype. Mol. Cell. Biol. *21*, 7933–7943. https://doi.org/10.1128/MCB.21.23.7933-7943.2001.

46. Izzo, A., Kamieniarz-Gdula, K., Ramírez, F., Noureen, N., Kind, J., Manke, T., van Steensel, B., and Schneider, R. (2013). The genomic landscape of the somatic linker histone subtypes H1.1 to H1.5 in human cells. Cell Rep. *3*, 2142–2154. https://doi.org/10.1016/j.celrep.2013.05.003.

47. Millán-Ariño, L., Islam, A.B.M.M.K., Izquierdo-Bouldstridge, A., Mayor, R., Terme, J.M., Luque, N., Sancho, M., López-Bigas, N., and Jordan, A. (2014). Mapping of six somatic linker histone H1 variants in human breast cancer cells uncovers specific features of H1.2. Nucleic Acids Res. *42*, 4474–4493. https://doi.org/10.1093/nar/gku079.

48. Torres, C.M., Biran, A., Burney, M.J., Patel, H., Henser-Brownhill, T., Cohen, A.-H.S., Li, Y., Ben-Hamo, R., Nye, E., Spencer-Dene, B., et al. (2016). The linker histone H1.0 generates epigenetic and functional intratumor heterogeneity. Science *353*, aaf1644. https://doi.org/10.1126/science.aaf1644.

49. Mayor, R., Izquierdo-Bouldstridge, A., Millán-Ariño, L., Bustillos, A., Sampaio, C., Luque, N., and Jordan, A. (2015). Genome distribution of replication-independent histone H1 variants shows H1.0 associated with nucleolar domains and H1X associated with RNA polymerase II-enriched regions. J. Biol. Chem. *290*, 7474–7491. https://doi.org/10.1074/jbc.M114.617324.

50. Prendergast, L., and Reinberg, D. (2021). The missing linker: emerging trends for H1 variant-specific functions. Genes Dev. *35*, 40–58. https://doi.org/10.1101/gad.344531.120.

51. Allan, J., Hartman, P.G., Crane-Robinson, C., and Aviles, F.X. (1980). The structure of histone H1 and its location in chromatin. Nature *288*, 675–679. https://doi.org/10.1038/288675a0.

52. Izquierdo-Bouldstridge, A., Bustillos, A., Bonet-Costa, C., Aribau-Miralbés, P., García-Gomis, D., Dabad, M., Esteve-Codina, A., Pascual-Reguant, L., Peiró, S., Esteller, M., et al. (2017). Histone H1 depletion triggers an interferon response in cancer cells via activation of heterochromatic repeats. Nucleic Acids Res. *45*, 11622–11642. https://doi.org/10.1093/nar/gkx746.

53. Doenecke, D., Albig, W., Bode, C., Drabent, B., Franke, K., Gavenis, K., and Witt, O. (1997). Histones: genetic diversity and tissue-specific gene expression. Histochem. Cell Biol. *107*, 1–10. https://doi.org/10.1007/s004180050083.

54. Maresca, T.J., and Heald, R. (2006). The long and the short of it: linker histone H1 is required for metaphase chromosome compaction. Cell Cycle *5*, 589–591. https://doi.org/10.4161/cc.5.6.2581.

55. Cole, K.D., Kandala, J.C., and Kistler, W.S. (1986). Isolation of the gene for the testis-specific H1 histone variant H1t. J. Biol. Chem. *261*, 7178–7183. https://doi.org/10.1016/S0021-9258(17)38371-0.

56. Amit, M., Donyo, M., Hollander, D., Goren, A., Kim, E., Gelfman, S., Lev-Maor, G., Burstein, D., Schwartz, S., Postolsky, B., et al. (2012). Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. Cell Rep. *1*, 543–556. https://doi.org/10.1016/j.celrep.2012.03.013.

57. Tammer, L., Hameiri, O., Keydar, I., Roy, V.R., Ashkenazy-Titelman, A., Custódio, N., Sason, I., Shayevitch, R., Rodríguez-Vaello, V., Rino, J., et al. (2022). Gene architecture directs splicing outcome in separate nuclear spatial regions. Mol. Cell *82*, 1021–1034.e8. https://doi.org/10.1016/j.molcel.2022.02.001.

58. Dombrowski, M., Engeholm, M., Dienemann, C., Dodonova, S., and Cramer, P. (2022). Histone H1 binding to nucleosome arrays depends on linker DNA length and trajectory. Nat. Struct. Mol. Biol. *29*, 493–501. https://doi.org/10.1038/s41594-022-00768-w.

59. Manning, G., Plowman, G.D., Hunter, T., and Sudarsanam, S. (2002). Evolution of protein kinase signaling from yeast to man. Trends Biochem. Sci. *27*, 514–520. https://doi.org/10.1016/s0968-0004(02)02179-5.

60. Bartkowiak, B., and Greenleaf, A.L. (2011). Phosphorylation of RNAPII: to P-TEFb or not to P-TEFb? Transcription *2*, 115–119. https://doi.org/10.4161/trns.2.3.15004.

61. Aslanzadeh, V., Huang, Y., Sanguinetti, G., and Beggs, J.D. (2018). Transcription rate strongly affects splicing fidelity and cotranscriptionality in budding yeast. Genome Res. *28*, 203–213. https://doi.org/10.1101/gr.225615.117.

62. Reimer, K.A., Mimoso, C.A., Adelman, K., and Neugebauer, K.M. (2021). Co-transcriptional splicing regulates 3′ end cleavage during mammalian erythropoiesis. Mol. Cell *81*, 998–1012.e7. https://doi.org/10.1016/j.molcel.2020.12.018.

63. Braunschweig, U., Gueroussov, S., Plocik, A.M., Graveley, B.R., and Blencowe, B.J. (2013). Dynamic integration of splicing within gene regulatory pathways. Cell *152*, 1252–1269. https://doi.org/10.1016/j.cell.2013.02.034.

64. Kwak, H., Fuda, N.J., Core, L.J., and Lis, J.T. (2013). Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. Science *339*, 950–953. https://doi.org/10.1126/science.1229386.

65. Core, L.J., Waterfall, J.J., and Lis, J.T. (2008). Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science *322*, 1845–1848. https://doi.org/10.1126/science.1162228.

66. Hunter, S., Sigauke, R.F., Stanley, J.T., Allen, M.A., and Dowell, R.D. (2022). Protocol variations in run-on transcription dataset preparation produce detectable signatures in sequencing libraries. BMC Genomics *23*, 187. https://doi.org/10.1186/s12864-022-08352-8.

67. Muse, G.W., Gilchrist, D.A., Nechaev, S., Shah, R., Parker, J.S., Grissom, S.F., Zeitlinger, J., and Adelman, K. (2007). RNA polymerase is poised for activation across the genome. Nat. Genet. *39*, 1507–1511. https://doi.org/10.1038/ng.2007.21.

68. Zeitlinger, J., Stark, A., Kellis, M., Hong, J.W., Nechaev, S., Adelman, K., Levine, M., and Young, R.A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat. Genet. *39*, 1512–1516. https://doi.org/10.1038/ng.2007.26.

69. Vyas, P., and Brown, D.T. (2012). N- and C-terminal domains determine differential nucleosomal binding geometry and affinity of linker histone isotypes H1(0) and H1c. J. Biol. Chem. *287*, 11778–11787. https://doi.org/10.1074/jbc.M111.312819.

70. Bowman, E.A., and Kelly, W.G. (2014). RNA polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation: A tail of two kinases. Nucleus *5*, 224–236. https://doi.org/10.4161/nucl.29347.

71. Lev Maor, G., Yearim, A., and Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. Trends Genet. *31*, 274–280. https://doi.org/10.1016/j.tig.2015.03.002.

72. Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. Science *327*, 996–1000. https://doi.org/10.1126/science.1184208.

73. de Almeida, S.F., Grosso, A.R., Koch, F., Fenouil, R., Carvalho, S., Andrade, J., Levezinho, H., Gut, M., Eick, D., Gut, I., et al. (2011). Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. Nat. Struct. Mol. Biol. *18*, 977–983. https://doi.org/10.1038/nsmb.2123.

74. Nojima, T., Rebelo, K., Gomes, T., Grosso, A.R., Proudfoot, N.J., and Carmo-Fonseca, M. (2018). RNA polymerase II phosphorylated on CTD serine 5 interacts with the spliceosome during co-transcriptional splicing. Mol. Cell *72*, 369–379.e4. https://doi.org/10.1016/j.molcel.2018.09.004.

75. Nieto Moreno, N., Giono, L.E., Cambindo Botto, A.E., Muñoz, M.J., and Kornblihtt, A.R. (2015). Chromatin, DNA structure and alternative splicing. FEBS Lett. *589*, 3370–3378. https://doi.org/10.1016/j.febslet.2015.08.002.

76. Willcockson, M.A., Healton, S.E., Weiss, C.N., Bartholdy, B.A., Botbol, Y., Mishra, L.N., Sidhwani, D.S., Wilson, T.J., Pinto, H.B., Maron, M.I., et al. (2021). H1 histones control the epigenetic landscape by local chromatin compaction. Nature *589*, 293–298. https://doi.org/10.1038/s41586-020-3032-z.

77. Eperon, L.P., Graham, I.R., Griffiths, A.D., and Eperon, I.C. (1988). Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? Cell *54*, 393–401. https://doi.org/10.1016/0092-8674(88)90202-4.

78. Kadener, S., Cramer, P., Nogués, G., Cazalla, D., de la Mata, M., Fededa, J.P., Werbajh, S.E., Srebrow, A., and Kornblihtt, A.R. (2001). Antagonistic effects of T-Ag and VP16 reveal a role for RNA Pol II elongation on alternative splicing. EMBO J. *20*, 5759–5768. https://doi.org/10.1093/emboj/20.20.5759.

79. Li, J.Y., Patterson, M., Mikkola, H.K.A., Lowry, W.E., and Kurdistani, S.K. (2012). Dynamic distribution of linker histone H1.5 in cellular differentiation. PLoS Genet. *8*, e1002879. https://doi.org/10.1371/journal.pgen.1002879.

80. Behrends, M., and Engmann, O. (2020). Linker histone H1.5 is an underestimated factor in differentiation and carcinogenesis. Environ. Epigenet. *6*, dvaa013. https://doi.org/10.1093/eep/dvaa013.

81. Kalashnikova, A.A., Rogge, R.A., and Hansen, J.C. (2016). Linker histone H1 and protein-protein interactions. Biochim. Biophys. Acta *1859*, 455–461. https://doi.org/10.1016/j.bbagrm.2015.10.004.

82. Venables, J.P. (2004). Aberrant and alternative splicing in cancer. Cancer Res. *64*, 7647–7654. https://doi.org/10.1158/0008-5472.CAN-04-1910.

83. Ghigna, C., Valacca, C., and Biamonti, G. (2008). Alternative splicing and tumor progression. Curr. Genomics *9*, 556–570. https://doi.org/10.2174/138920208786847971.

84. Murthy, T., Bluemn, T., Gupta, A.K., Reimer, M., Rao, S., Pillai, M.M., and Minella, A.C. (2018). Cyclin-dependent kinase 1 (CDK1) and CDK2 have opposing roles in regulating interactions of splicing factor 3B1 with chromatin. J. Biol. Chem. *293*, 10220–10234. https://doi.org/10.1074/jbc.RA118.001654.

85. Saha, A., Seward, C.H., Stubbs, L., and Mizzen, C.A. (2020). Site-specific phosphorylation of histone H1.4 is associated with transcription activation. Int. J. Mol. Sci. *21*, 8861. https://doi.org/10.3390/ijms21228861.

86. Zheng, Y., John, S., Pesavento, J.J., Schultz-Norton, J.R., Schiltz, R.L., Baek, S., Nardulli, A.M., Hager, G.L., Kelleher, N.L., and Mizzen, C.A. (2010). Histone H1 phosphorylation is associated with transcription by RNA polymerases I and II. J. Cell Biol. *189*, 407–415. https://doi.org/10.1083/jcb.201001148.

87. Karrer, K.M., and VanNuland, T.A. (1999). Nucleosome positioning is independent of histone H1 in vivo. J. Biol. Chem. *274*, 33020–33024. https://doi.org/10.1074/jbc.274.46.33020.

88. Kowalski, A., and Pałyga, J. (2016). Modulation of chromatin function through linker histone H1 variants. Biol. Cell *108*, 339–356. https://doi.org/10.1111/boc.201600007.

89. Wu, H., Dalal, Y., and Papoian, G.A. (2021). Binding dynamics of disordered linker histone H1 with a nucleosomal particle. J. Mol. Biol. *433*, 166881. https://doi.org/10.1016/j.jmb.2021.166881.

90. Christophorou, M.A., Castelo-Branco, G., Halley-Stott, R.P., Oliveira, C.S., Loos, R., Radzisheuskaya, A., Mowen, K.A., Bertone, P., Silva, J.C.R., Zernicka-Goetz, M., et al. (2014). Citrullination regulates pluripotency and histone H1 binding to chromatin. Nature *507*, 104–108. https://doi.org/10.1038/nature12942.

91. Kim, K., Jeong, K.W., Kim, H., Choi, J., Lu, W., Stallcup, M.R., and An, W. (2012). Functional interplay between p53 acetylation and H1.2 phosphorylation in p53-regulated transcription. Oncogene *31*, 4290–4301. https://doi.org/10.1038/onc.2011.605.

92. Talasz, H., Sarg, B., and Lindner, H.H. (2009). Site-specifically phosphorylated forms of H1.5 and H1.2 localized at distinct regions of the nucleus are related to different processes during the cell cycle. Chromosoma *118*, 693–709. https://doi.org/10.1007/s00412-009-0228-2.

93. Tapial, J., Ha, K.C.H., Sterne-Weiler, T., Gohr, A., Braunschweig, U., Hermoso-Pulido, A., Quesnel-Vallières, M., Permanyer, J., Sodaei, R., Marquez, Y., et al. (2017). An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. Genome Res. *27*, 1759–1768. https://doi.org/10.1101/gr.220962.117.

94. Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

95. Pohl, A., and Beato, M. (2014). bwtool: a tool for bigWig files. Bioinformatics *30*, 1618–1619. https://doi.org/10.1093/bioinformatics/btu056.

96. R Core Team (2022). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/.

97. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

98. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10. https://doi.org/10.14806/ej.17.1.200.

99. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

100. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

101. Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191. https://doi.org/10.1093/nar/gku365.

102. Xu, S., Grullon, S., Ge, K., and Peng, W. (2014). Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. Methods Mol. Biol. *1150*, 97–111. https://doi.org/10.1007/978-1-4939-0512-6_5.

103. Yu, G., Wang, L.-G., and He, Q.-Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics *31*, 2382–2383. https://doi.org/10.1093/bioinformatics/btv145.

104. Wang, S., Sun, H., Ma, J., Zang, C., Wang, C., Wang, J., Tang, Q., Meyer, C.A., Zhang, Y., and Liu, X.S. (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. Nat. Protoc. *8*, 2502–2515. https://doi.org/10.1038/nprot.2013.150.

105. Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. Nucleic Acids Res. *32*, D493–D496. https://doi.org/10.1093/nar/gkh103.

106. Quinlan, A.R. (2014). BEDTools: the Swiss-army tool for genome feature analysis. Curr. Protoc. Bioinformatics *47*, 11.12.1–11.1234. https://doi.org/10.1002/0471250953.bi1112s47.

107. Cunningham, F., Allen, J.E., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Austine-Orimoloye, O., Azov, A.G., Barnes, I., Bennett, R., et al. (2022). Ensembl 2022. Nucleic Acids Res. *50*, D988–D995. https://doi.org/10.1093/nar/gkab1049.

108. Gohr, A., and Irimia, M. (2019). Matt: Unix tools for alternative splicing analysis. Bioinformatics *35*, 130–132. https://doi.org/10.1093/bioinformatics/bty606.

109. Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., and Chang, W. (2022). David: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res. *50*, W216–W221. https://doi.org/10.1093/nar/gkac194.

110. Thorvaldsdóttir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief. Bioinform. *14*, 178–192. https://doi.org/10.1093/bib/bbs017.

111. Kramer, N.E., Davis, E.S., Wenger, C.D., Deoudes, E.M., Parker, S.M., Love, M.I., and Phanstiel, D.H. (2022). Plotgardener: cultivating precise multi-panel figures in R. Bioinformatics *38*, 2042–2045. https://doi.org/10.1093/bioinformatics/btac057.

112. Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallières, M., Tapial, J., Raj, B.,

O'Hanlon, D., et al. (2014). A highly conserved program of neuronal microexons is misregulated in autistic brains. Cell *159*, 1511–1523. https://doi.org/10.1016/j.cell.2014.11.035.

113. Braunschweig, U., Barbosa-Morais, N.L., Pan, Q., Nachman, E.N., Alipanahi, B., Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., and Blencowe, B.J. (2014). Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res. *24*, 1774–1786. https://doi.org/10.1101/gr.177790.114.

114. Hattori, T., Taft, J.M., Swist, K.M., Luo, H., Witt, H., Slattery, M., Koide, A., Ruthenburg, A.J., Krajewski, K., Strahl, B.D., et al. (2013). Recombinant antibodies to histone post-translational modifications. Nat. Methods *10*, 992–995. https://doi.org/10.1038/nmeth.2605.

115. Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M.R., Armean, I.M., Bennett, R., Bhai, J., Billis, K., Boddu, S., et al. (2019). Ensembl 2019. Nucleic Acids Res. *47*, D745–D751. https://doi.org/10.1093/nar/gky1113.

116. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

117. Yeo, G., and Burge, C.B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J. Comput. Biol. *11*, 377–394. https://doi.org/10.1089/1066527041410418.

118. Costantini, M., Clay, O., Auletta, F., and Bernardi, G. (2006). An isochore map of human chromosomes. Genome Res. *16*, 536–541. https://doi.org/10.1101/gr.4910606.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit Anti-H1.2	Abcam	RRID:AB_2117983
Rabbit Anti-H1.3	Abcam	RRID:AB_447908
Rabbit Anti-H1.4	Abcam	RRID:AB_10862627
Rabbit Anti-H1.5	Abcam	RRID:AB_470263
Mouse Anti-H1.0	Abcam	RRID:AB_2295032
Donkey Anti-Rabbit	Abcam	RRID:AB_2636877
Donkey Anti-Mouse	Abcam	RRID:AB_2732856
rabbit anti-RNAPII Ser-2	Abcam	RRID:AB_304749
Chemicals, peptides, and recombinant proteins		
5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole (DRB)	Sigma-Aldrich	D1916
Deposited data		
Raw and analyzed data	This paper	GEO: GSE228240
Original PCR and WB gels	This paper; Mendeley data	https://doi.org/10.17632/vfjm6k4hxg.1
Experimental models: Cell lines		
1-KO (H1.2 KO)	This paper	N/A
2-KO (H1.2 + H1.3 KO)	This paper	N/A
4-KO (H1.0+H1.2+H1.3+H1.5 KO)	This paper	N/A
5-KO (H1.0+H1.2+H1.3+H1.4+H1.5)	This paper	N/A
1-KO Rescue (H1.2)	This paper	N/A
2-KO Rescue (H1.2 +H1.3)	This paper	N/A
2-KO Rescue (H1.3 Only)	This paper	N/A
4-KO Rescue (H1.0 Only)	This paper	N/A
4-KO Rescue (H1.5 Only)	This paper	N/A
5-KO + Transfection of one H1 subtype (H1.0, H1.2, H1.3, H1.4 or H1.5). In sum five cell lines generated.	This paper	N/A
Oligonucleotides		
Primers used for RT-qPCR, see Table S4	This paper	N/A
sgRNA sequences used to target genes encoding H1 variants, see Table S1	This paper	N/A
Oligonucleotides used for PRO-seq, see Table S3	This paper	N/A
Recombinant DNA		
2A-GFP-Cas9 plasmid	Sigma	CAS9GFPP
pCDNA4-HA vector	D. Reinberg's group (NYU Medical School) and Albert Jordan's group (Institute of Molecular Biology of Barcelona)	N/A
Software and algorithms		
vast-tools v2.5.1	Tapial et al.[93]	https://github.com/vastgroup/vast-tools
FASTQC tool	Andrews[94]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
bwtool v1.0	Pohl and Beato[95]	https://github.com/CRG-Barcelona/bwtool/blob/master/INSTALL
R v4.2.1	R Foundation for Statistical Computing[96]	https://cran.r-project.org/bin/windows/base/
STAR v.2.7.10b	Dobin et al.[97]	https://github.com/alexdobin/STAR
Cutadapt v2.3 and v3.4	Martin[98]	https://cutadapt.readthedocs.io/en/stable/installation.html
Bowtie2 program v2.4.1	Langmead and Salzberg[99]	https://www.metagenomics.wiki/tools/bowtie2/install
samtools v.1.10 and v.1.12	Li et al.[100]	http://www.htslib.org/download/
deepTools v3.5.1	Ramírez et al., 2016[101]	https://deeptools.readthedocs.io/en/develop/content/installation.html
SICER v1.1	Xu et al.[102]	https://zanglab.github.io/SICER2/
ChIPseeker	Yu et al.[103]	https://bioconductor.org/packages/release/bioc/html/ChIPseeker.html
BETA	Wang et al.[104]	http://cistrome.org/BETA/
UCSC Genome Browser	Karolchik et al.[105]	https://genome.ucsc.edu/goldenpath/help/mirrorManual.html
bedtools v2.29.1	Quinlan[106]	https://bedtools.readthedocs.io/en/latest/content/installation.html
Ensembl web resource	Cunningham et al.[107]	https://www.ensembl.org/index.html
Matt v1.3.0	Gohr and Irimia[108]	https://gitlab.com/aghr/matt
DAVID v6.8	Sherman et al.[109]	https://david.ncifcrf.gov/
IGV	Thorvaldsdóttir et al.[110]	https://software.broadinstitute.org/software/igv/download
plotgardner R package v1.4.1	Kramer et al.[111]	https://bioconductor.org/packages/release/bioc/html/plotgardener.html
Other		
MNase-seq for HEK293 cell line	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM1624500	GSM1624500

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Gil Ast. Email: gilast@tauex.tau.ac.il

Materials availability

Constructed stable cell-lines generated in this study are available upon request to the lead contact, Gil Ast.

Data and code availability

- High throughput sequencing data generated in this study by ChIP-Seq, RNA-seq, PRO-seq processed data files were deposited at GEO under accession number GEO: GSE228240. Mendeley data: https://doi.org/10.17632/vfjm6k4hxg.1
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell maintenance

HEK293 cells were cultured in Dulbecco's modified Eagle's medium with 10% fetal bovine serum, 0.29 mg ml^{-1} L-glutamine, 100 U ml^{-L} penicillin, and 0.1 mg ml^{-p} streptomycin. The cells were incubated at 37 °C in a humidified atmosphere with 5% CO_2.

Cloning of sgRNAs

The single guide RNA (sgRNA) and Cas9 expression plasmid 2A-GFP-Cas9 were cut with BbsI (NEB) according to the manufacturer's protocol and purified using the Wizard Plus SV Miniprep DNA Purification System (Promega). All sgRNA sequences were designed using the CRISPR Design Tool (http://crispr.mit.edu/). Sequences with the highest scores were chosen. Each sgRNA was designed together with the complementary sequence (gRNA_comp), and five nucleotides, GAAGAC, were added to the 5′ end of each sequence to complement the BbsI restriction site. Oligonucleotides were phosphorylated at their 5′ ends using T4 Polynucleotide Kinase (NEB) in T4 ligase buffer (NEB) according to the manufacturer's protocol. In a thermocycler, the sgRNA and sgRNA_comp oligonucleotides were annealed, and the phosphorylation reaction was performed at 37 °C for 30 min. After heat inactivation at 65 °C for 20 min, samples were held at 95 °C for 5 min and then the temperature was reduced to 4 °C at 0.5 °C per second to anneal the strands. sgRNAs were diluted 1:25 and were ligated into linearized 2A-GFP-Cas9 plasmid using T4 ligase (NEB) according to manufacturer's instructions (50 ng vector with 1 μl diluted sgRNA in a 20 μl reaction volume). A 2-μl aliquot of the ligation reaction was used to transform competent 10XL Gold *Escherichia coli* after heat shock (ampicillin selection was used). Plasmids extracted from three colonies were purified using the Wizard Plus SV Miniprep Kit (Promega), and positive colonies were confirmed by Sanger sequencing at the Inter-Departmental Research Facility Unit. sgRNA sequences are listed in Table S1.

CRISPR/Cas9-mediated generation of H1-KO and rescue cell lines

HEK293 cells were transfected with 3 μg of 2A-GFP-Cas9 engineered to express an sgRNA targeting a gene encoding a particular H1 in a stepwise fashion to create cell lines with null mutations in one to five H1-encodoing genes. All transfections were performed using the TransIT®-LT1 transfection reagent (Mirus) according to the manufacturer's instructions. Expression of the sgRNA and Cas9 resulted in insertion of deletion of nucleotides in the genes encoding the specific H1 variants gene, generating frame shifts that resulted in loss of H1 protein expression. The cells were sorted based on expression of GFP on a FACS Aria and were cultured in 15-cm plates to obtain clonal growth. The clones were grown until they filled a 10-cm standard plate.

METHOD DETAILS

Protein extraction

Cells were scratch harvested from the culture plates and transferred to tubes with RIPA buffer (50 mM Tris-HCl, pH 7.5, 1% NP40, 150 mM NaCl, 0.1% SDS, 0.5% deoxycholic acid, 1 mM EDTA) with protease and phosphatase inhibitors. Mouse tissues were homogenized on ice in RIPA buffer with protease and phosphatase inhibitors. Both types of samples were incubated on ice for 30 min and then the protein-containing supernatants were collected after centrifugation for 20 min at 14,000 g at 4 °C. Protein concentrations were measured by Bradford assay (BioRad).

Western blot and antibodies

Total protein (50 μg) from cells or tissues were loaded on 10% SDS-PAGE and then electroblotted onto a Protran nitrocellulose transfer membrane. The membranes were blocked with 3% BSA in sodium azide for 1 h at room temperature and then probed for histone H1 variants with the following antibodies: rabbit anti-H1.2 (Abcam 4086, 1:1000), rabbit anti-H1.3 (Abcam 24174, 1:1500), rabbit anti-H1.4 (Abcam 105522, 1:500), rabbit anti-H1.5 (Abcam 18208, 1:2000), and mouse anti-H1.0 (Abcam 11079, 1:3000). For co-immunoprecipitation, the following antibody was used: rabbit anti-RNAPII Ser-2 (Abcam 5095, 1:400). After a 12 h incubation at 4 °C, blots were washed and incubated at room temperature for 1 h with donkey anti-rabbit (Abcam 1:30,000) or donkey anti-mouse (Abcam 1:40,000) as appropriate. Immunoblots were visualized using CYANAGEN – WESTAR NOVA 2.0, an ECL substrate with stable light output designed for low picogram detection on a VILBER Fusion FX.

Co-immunoprecipitation

Approximately 10 × 10^6 HEK293 KO H1.2 cells per sample were trypsinized, washed with PBS. Samples were centrifuged at 1500 × g, washed with cold PBS, and centrifuged again. Nuclei isolated. Cells' pellet was suspended in buffer 1 (60 mM KCl, 15 mM NaCl, 5 mM MgCl2, 0.1 mM EGTA, 15 mM Tris-HCl [pH 7.5], and 0.3 M Sucrose) supplemented with 0.5 mM DTT, 0.1 mM PMSF, 1× complete protease inhibitor (CPI), and incubated in 0.2% IGEPAL CA-630 (NP-40). After incubation for 10 min, residual NP-40 was cleared by centrifugation on a 1.2 M sucrose cushion. Nuclei were suspended in MNase digestion buffer (0.32 M sucrose, 50 mM Tris-HCl [pH 7.5], 4 mM MgCl2, 1 mM CaCl2) supplemented with 0.1 mM PMSF. MNase (10 U/10^6 nuclei, Worthington) was added, and samples were incubated at 37 °C for 10 min. The reaction was stopped by the addition of 20 mM EDTA. Nuclei were then sedimented by centrifugation. MNase-digested nuclei were suspended in immunoprecipitation (RIPA) buffer (140 mM NaCl, 1 mM EDTA [pH 8], 0.5 mM EGTA, 20 mM Tris-HCl [pH 8], 1% NP-40, 0.1% SDS, 0.4% DOC) supplemented with 0.1 mM PMSF, 1× CPI, and rotated for 1h at 4 °C. The samples were sonicated using a Bioruptor (Diagenode) at 40% amplitude in intervals of 2.2 s pulses with 9.9 s pauses for 5 min, followed by centrifugation 10,000 × g for 10 min. This supernatant is denoted as "input". Antibodies (5 μg) used for immunoprecipitation, anti HA (Abcam, ab9110) and anti IgG (Santa Cruz Biotechnology; sc2027). Mix of protein A and G beads (40 μl, Dynabeads Invitrogen) were washed and added to each sample along with the antibodies and rotated for overnight at 4 °C. Beads were washed four times with PBS (0.02% NP-40) supplemented with 0.1 mM PMSF, 1× CPI. Protein was eluted from the beads by adding 100 μl RIPA and 20 μl 6× SDS sample buffer (272 mM Tris-HCl [pH 6.8], 30% glycerol, 12% SDS, 20% β-mercaptoethanol,

0.01% bromophenol blue) and incubating at 75 °C in a thermo-shaker for 15 min with vigorous shaking at 750rpm. The supernatant was moved to a new tube and boiled for 5 min at 100 °C. 36 μl of 10% input, and 40 μl of IP and IgG were loaded. All tubes were supplemented to 40 μl with SBX1. Samples were separated by SDS-PAGE on 10% polyacrylamide gels and transferred to 0.45-μm nitrocellulose membranes (Whatman Protran). The membranes were incubated with the appropriate primary and secondary antibodies and washed with TTBS. Antibody used was anti-RNAPII p-Ser2 (Abcam; ab5095).

MNase digestion

A nucleosome ladder was generated by extracting and counting nuclei in order to determine the amount of the MNase enzyme needed: 10 units of MNase digest the chromatin in 1,000,000 cells to nucleosomes at 37 °C for 15 minutes. EDTA was used to quench the reaction. Samples were resolved on agarose gels or a TapeStation system.

Cell cycle and replication rate evaluation

Cells were incubated with DAPI, which stains the DNA inside nuclei, and then sorted on a GALLIOS flow cytometer according to staining intensity. The highest staining is obtained in cells in G2/M, there is less intense staining in S phase cells, and the least intense staining is observed in cells in G0/G1. Pulse-width/pulse-area signals were used to discriminate between cells in G2/M and cell doublets. For replication rate quantification we seeded the same number of cells 200K and allowed them to grow for 24 hours. The raw counts of cells that were counted under the microscope according to the formula. Using the following formula, we can calculate the growth rate and the growth ratio as follows:

$$\mu(\textit{growth rate}) = \frac{\ln\left(\frac{Nt(\textit{number of cells harvested})}{N0(\textit{number of cells seeded})}\right)}{\triangle t(\textit{hours of growth }(h))} * t$$

cDNA synthesis and qPCR

cDNA synthesis was performed with RT-FLEX (Quanta) according to the manufacturer's instructions. qPCR was performed using KAPA Biosystem's SYBR FAST Universal qPCR kit according to the manufacturer's instructions. A calibration curve on the control cell line was constructed by using decreasing concentrations of DNA (from 25 ng to 0.0975 ng) and specific primers for each histone H1 variant (H1.0-H1.5, H1.X, and H1t) expressed in HEK293 cells. The optimal concentration of 1.56 ng DNA was used to measure the absolute quantities of H1s in the KO cell lines and the control cells. The analysis used the calculation: Quantity = $10^{((Ct-b)/m)}$, where b is the y axis = 0, and m is the slope of the calibration curve for each of the different H1s.

For qPCR validation of genes with altered splicing level, genes were selected according to the dPSI in KO and control cells, with absolute dPSI of at least 10 for genes that passed the quality criteria.

RNA-seq analyses

RNA was extracted by adding 100 μl 1-bromo 3-chloro propane to cell lysates followed by vigorous shaking (20 s) and a 15-min centrifugation at 12,000 g at 4 °C. The top, transparent liquid phase containing the RNA was transferred to a new tube, 0.5 ml isopropanol was added, and the sample was centrifuged at 12,000 g for 8 min at 4 °C. The RNA pellet was washed with 1 ml 70% ethanol, resuspended in RNase-free water and treated with DNase (ThermoFisher) to eliminate residual DNA. A sample of 1500-2000 ng of total RNA was used to generate cDNA from oligo(dT) primers and SuperScript III reverse transcriptase. mRNA library preparation and subsequent sequencing were performed at the Weizmann Institute (Rehovot, Israel). Paired-end sequencing of mRNA samples 1-KO, 2-KO, 4-KO, 5-KO, 1-KO + rescue H1.2, 2-KO + rescue H1.2+H1.3 were performed using the Illumina HiSeq 2500 system. Paired-end sequencing of mRNA (RNA-seq) samples was performed using the Illumina NovaSeq SP system for the following samples: HEK293 derived 2-KO + H1.3 rescue, 4-KO + H1.0 rescue, 4-KO + H1.5 rescue. For the rescue experiment, cells were transfected with the cDNA of H1.

Sequenced reads were aligned to human genome assembly 19 (hg19) using the vast-tools v2.5.1 with default options.[93] The quality of sequenced reads was evaluated using the FASTQC tool.[94] Quality metrics from the alignment of samples and expression tables were obtained from vast-tools v2.5.1. Reads were normalized by quantile normalization and corrected for mappability of reads per kilobase per million mapped reads (RPKM). Between-sample comparisons were performed after quantile normalization, and log-transformed (RPKM +1) values were filtered to remove poorly expressed genes. Spearman correlation analysis and statistical P values with permutation were computed using an in-house R script.

The detection of the AS events were performed using vast-tools v2.5.1[112] with default parameters. AS events were compared between samples using vast-tools version 2.5.1.[93] The results were filtered to exclude B3- and VLOW-scored events. In addition, IR events were filtered according to the binomial test with option "–p_IR" to assess the balance of reads at the two exon-intron junctions.[113] Furthermore, compared events were classified according to the difference in percent spliced in (dPSI = PSI_B – PSI_A). Datasets of the affected events with absolute PSI difference of at least 10 used for analysis of the overlapped splicing events within analyzed samples. Venn diagrams and Upset plots built using in-house R script. Fisher test used for estimation of the significance of overlap.

To build a heatmap of the events' PSI pattern we choose instances that passed the quality criteria of the read coverage in WT, 1-KO, 2-KO, 1-Rescue (H1.2 KO and rescue) and 2-Rescue (H1.2 and H1.3 KO and rescue) samples. And then plotted the inclusion value (PSI), for each of the examined events. We then performed clustering of samples and the events; plot generated using in-house R script.

Nucleosome occupancy

The nucleosome occupancy data for hg19 was downloaded from the Hattori et al. study[114] (GEO: GSM1624500) based on MNase-seq of human HEK293 in bigWig format, and then annotated using the aggregate function from bwtool v1.0[95] and an in-house Python script. Statistical analyses were performed using R v4.2.1.[96]

ChIP-seq analyses

Cells were fixed using 1% formaldehyde, harvested, and then sonicated using a Vibra-Cell VCX600. This was followed by MNase digestion to generate chromatin fragments between 150 and 300 base pairs (bp). For ChIP, 25 ug of chromatin was immunoprecipitated overnight at 4 °C with the appropriate antibody (anti-H1.2, Abcam 4086; anti-H1.5, Abcam 18208; anti-H1.X, Abcam 31972; or anti-HA tag, Abcam 9110). Rabbit IgG (Cell Signaling Technology) was used as a control for nonspecific DNA interactions. Samples were prepared with 10% chromatin material. Immunocomplexes were recovered using 40 μl mixture of Protein-A and G magnetic beads (Invitrogen). Beads with bound antibody/protein/DNA complexes were washed, reverse cross-linked at 65 °C overnight, and immunoprecipitated. DNA was recovered using phenol:chloroform:isoamyl alcohol (Sigma) extraction. Sequences encoding H1 variants were PCR-amplified from genomic DNA and cloned into the pCDNA4-HA vector provided by D. Reinberg's group (NYU Medical School) and Albert Jordan's group (Institute of Molecular Biology of Barcelona).[12]

A library was prepared for sequencing using standard Illumina protocols. Single-end sequencing was performed with the Illumina Next-seq system for the H1.2, H1.5, and H1.X endogenous variants that were immunoprecipitated with specific antibodies from WT cells. The Illumina HiSeq system was used for H1 KO samples in which HA-tagged H1s were immunoprecipitated (Table S2, last row), with the exception of H1.X, which was precipitated with specific antibodies from both 5-KO and WT cells. Read quality was evaluated using FASTQC.[94] The adaptor sequences were trimmed using Cutadapt v2.3.[98] Processed reads were aligned to the hg19 reference genome using the Bowtie2 program v2.4.1.[99] PCR-duplicated reads and low mapping quality reads were removed using samtools v.1.10 rmdup and samtools –q 30 functions, respectively.[100] Sam-formatted files were converted to Bam format with samtools v.1.10 and sorted.

Enrichment of the H1 coverage signal ($\log_2$ ratio) of immunoprecipitate (IP) over input and the subtraction between IP and input were calculated for each sample, and samples were normalized by RPKM using deepTools v3.5.1.[101] The resulting enrichment signal files were visualized using IGV.[110] For statistical correlation analyses, we used the deepTools v3.5.1 multiBigWigsummary function. The genome was divided to bins of 80,000 bp and the $\log_2$ IP/input-normalized signal of H1 coverage calculated in these regions. Correlation data for scatter plots and the heatmap were created using the deepTools v3.5.1.[101] The skip zeros option was used, which excluded genomic regions with zero or missing values in all samples. In addition, scatter plots were generated for pairs of H1.2, H1.5, and H1.X variants from the KO cell lines (IP with antibodies to HA-tag) and WT cells (IP with specific antibodies), and a heatmap was constructed for samples H1.0, H1.2, H1.3, H1.4, and H1.5, with HA-tag, and H1.X from the KO cell lines (Table S2 last row). Hierarchical clustering was performed using deepTools v3.5.1.[101] Scatter plots and P values with permutation were generated using an in-house Python script. The peak caller program for histone, SICER v1.1[102] was used with the default parameters and FDR $\leq$ 0.01. Enriched peaks were annotated to the nearest gene features, and pie charts were generated using an R script written in-house and the Bioconductor package ChIPseeker.[103] Potential gene targets of H1s were detected using BETA.[104]

Dataset preparation and grouping

Human Ensembl genome annotations that are compatible with hg19 were downloaded from the UCSC Genome Browser.[105,115] The standard output of Ensembl annotations was processed and filtered using an in-house R script. Exons were annotated with GC content value and expression data of WT HEK293 samples using bedtools v2.29.1.[116] The first and last exon of each transcript were removed.

Following the method described by Amit et al.,[56] lengths of upstream and downstream introns of each exon were divided into five groups based on percentiles (0–20, 20–40, 40–60, 60–80, 80–100). Internal exons were classified according to the percentile group lengths of their flanking introns. Both introns flanking a given exon were conditioned to be in one percentile group. These groups represent exon-intron architectures that have different GC content patterns. An in-house R script was used to assign exons to the percentile groups (0–20, 20–40, 40–60, 60–80, 80–100) according to their corresponding gene expression values obtained from vast-tools v2.5.1[93] analysis of WT HEK293 data. Exons derived from genes for which the expression level was not defined were not included in the analysis.

Gene annotations compatible for hg19 were downloaded from the Ensembl web resource to generate a gene dataset that could be used for visualization of the H1 profiles around transcription start sites (TSSs) and transcript end sites (TESs).[107] Unique protein-coding genes were classified into TSS and TES datasets according to the Ensembl annotation of gene coordinates. Genes were also classified according to GC content using bedtools v2.29.1[106,116] for GC content extraction.

Profile plot construction

For each exon, 150 nucleotides (nt) from the upstream and downstream flanking introns as well as 75 nt from the start and end coordinate of each exon in the dataset were extracted using the bwtool v1.0.[95] For intron, 75 nucleotides (nt) from the upstream and downstream flanking exons as well as 150 nt from the start and end coordinate of each intron in the dataset were extracted. An in-house Bash script was used to analyze bigWig files of log_2-normalized IP/Input signal from ChIP-seq of H1, nucleosome occupancy in HEK293 and binary bigWig files of GC content for hg19, depending on the plot. All nucleotides participating in splice-site signals were discarded as previously described by Amit et al.: 20 nt from the 3′ SS end of the upstream introns, the first 6 nt at the 5'SS end of the downstream intron, and the first 2 and the last 3 nt of the exonic sequences.[56] The TSS and TES datasets of genes were used as regions for which the value of the log_2 normalized IP/Input signal from ChIP-seq of H1 was extracted. The region 2 kilobases (kb) upstream and 500 bp downstream were extracted for TSS, and the 2 kb upstream and downstream regions were used for TES. The average coverage value at single-nucleotide resolution was computed and plotted. Resulting profiles were smoothed using an in-house R script.

Integration between RNA-seq and ChIP-seq

The following comparisons of splicing changes were used in order to extract affected and unaffected groups of exons/introns in the given cells:

For H1.2, we generated groups as follows:

$$H1\ unaffected\ =\ dPSI(0)\ in\ (H1.2KO \sim WT)$$

$$H1\ affected\ =\ dPSI(10)\ in\ (H1.2KO \sim WT)$$

For H1.0, H1.3, and H1.5, we generated groups as follows:

$$H1\ unaffected\ =\ dPSI(0)\ in\ (KO \sim Rescue)\ and\ (WT \sim KO)$$

$$H1\ affected\ =\ dPSI(10)\ in\ (KO \sim Rescue)\ and\ (WT \sim KO)$$

For H1.3, the 2-KO sample was used, and for H1.0 and H1.5 the 4-KO sample was used.

For H1.4, we generated groups as follows:

$$H1.4\ unaffected\ =\ dPSI(0)\ in\ (5KO \sim 4KO)\ and\ dPSI(0)\ in\ (WT \sim 4KO)$$

$$H1.4\ affected\ =\ dPSI(10)\ in\ (5KO \sim 4KO)\ and\ dPSI(0)\ in\ (WT \sim 4KO)$$

Where dPSI(0) means unaffected groups with absolute dPSIs of less than 10 and dPSI(10) are groups with absolute dPSIs of at least 10. We also categorized affected groups according to the direction of PSI change: upregulated (dPSI $\geq$ 10) or downregulated (dPSI $\leq$ −10), or into groups called H1 target when one of the aforementioned groups consisted of less than 20 members.

We used the dPSI value direction obtained from the comparisons between the KO and rescue samples (for H1.0, H1.3 and H1.5), between KO and WT (for H1.2), between 5-KO and 4-KO (for H1.4) as the effect on PSI (up- or downregulated) when H1 is present in the sample. We then matched affected splicing events with potential H1 target genes. All data processing was conducted using scripts written in-house in R and Bash.

We extracted multiple features for each exon/intron from the generated datasets using the Matt program v1.3.0.[108] The bwtool v1.0[95] and an in-house R script were used to extract and compute the mean log_2 normalized IP/input signal in the windows $\pm$ 75 nucleotides from the start and end coordinate of each exon/intron. Spliced unit length and adjacent features (introns or exons) were filtered to a minimum of 75 nucleotides in length. All nucleotides participating in splice-site signals were excluded following the procedure outlined in the preceding methods section. Boxplots were generated using R v4.2.1.[96] An in-house R script was used for a Wilcoxon paired statistical test to assess the significance of H1 signal differences between exon and intron side around SSs.

Boxplots of the maximum entropy scores of SSs in generated groups were constructed according to the model presented by Yeo and Burge[117] and Wilcoxon statistical tests were used to test significance, which was done with the usage of Matt tool v1.3.0.[108]

Plotting GC levels over the chromosomes

The reference genome hg19 was divided into 100-kb windows, and the GC percent in each window was calculated using bedtools v2.29.1.[116] Assignment of the GC values to groups with different levels was as described previously,[118] then scaled to [0,1]. The input-subtracted ChIP-seq mean signal of H1 was calculated for each window, normalized and scaled to [-1:1] using an in-house R script. H1 coverage signal values for each chromosome were plotted and colored according to the normalized, scaled GC level in the same windows.

PRO-seq library preparation

For cell permeabilization, all buffers were cooled on ice, all steps were performed on ice, and all samples were spun at 300 x g at 4 °C unless otherwise noted. HEK293 control and H1 KO cell lines (1-KO, 2-KO, 4-KO, and 5-KO) cells were washed with PBS and resuspended in 1 mL Buffer W (10 mM Tris-HCl, pH 8.0, 10 mM KCl, 250 mM sucrose, 5 mM $MgCl_2$, 0.5 mM DTT, 10% glycerol), then strained through a 35-μm Falcon tube filter. Nine volumes of Buffer P (Buffer W + 0.1% IGEPAL CA-630) were immediately added to each sample, cells were rotated for 2 min at room temperature, and then were centrifuged at 300 x g at 4 °C for 4 min. Cells were washed in Buffer F (50 mM Tris-HCl, pH 8.3, 40% glycerol, 5 mM $MgCl_2$, 0.5 mM DTT, 1 μL/mL SUPERase.In [ThermoFisher]), and then resuspended in Buffer F at a final volume of 1 × 10^6 permeabilized cells per 40 μL. Samples were stored at −80 °C.

For library generation, 1 x 10^6 cells were spiked with 5% permeabilized *Drosophila* S2 cells for data normalization and used as input for PRO-seq. Nascent RNA was labeled through a biotin-NTP run-on as follows: Permeabilized cells were added to an equal volume of a 2X run-on reaction mix (10 mM Tris-HCl, pH 8.0, 300 mM KCl, 1% Sarkosyl, 5 mM $MgCl_2$, 1 mM DTT, 200 μM biotin-11-A/C/G/UTP (Perkin-Elmer), 0.8 U/μL SUPERase.In (ThermoFisher)), and incubated at 30 °C for 5 min. RNA was isolated using TRIzol Reagent (ThermoFisher). Isolated RNA was fragmented by base hydrolysis with 0.25 N NaOH for 9 min on ice, followed by neutralization with one volume of 1 M Tris-HCl, pH 6.8. To select for nascent RNA, 48 μL of washed Streptavidin M-280 magnetic beads (ThermoFisher) in binding buffer (300 mM NaCl, 10 mM Tris-HCl pH 7.4, 0.1% Triton X-100) were added to the fragmented RNA, and samples were rotated at room temperature for 20 min. The Streptavidin M-280 magnetic beads were washed twice in each of the following three buffers: high salt buffer (2 M NaCl, 50 mM Tris-HCl pH, 7.4, 0.5% Triton X-100), binding buffer, and low salt buffer (5 mM Tris-HCl, pH 7.4, 0.1% Triton X-100). Beads were resuspended in TRIzol LS Reagent (ThermoFisher) and heated at 65 °C for 5 min, cooled to room temperature, and heated at 65 °C for 5 min again to elute the RNA from the beads. The RNA was then precipitated with ethanol. Nascent RNA was resuspended in 10 μM of the VRA3 3′ end adapter in T4 Ligase buffer (Table S3) and was 3′ end ligated with T4 RNA ligase I (NEB) for 2 h at room temperature. The sample was enriched for ligated nascent RNAs by binding to Streptavidin M-280 magnetic beads as described above. The beads were washed twice in high, binding, and low salt buffers, then once in 1X ThermoPol Buffer (NEB). To prepare the nascent RNA for 5′ end adapter ligation, the 5′ ends of the RNA were decapped using RppH suspended in RppH buffer, then 5′ end was recapped with RNA 5′ pyrophosphohydrolase (NEB) at 37 °C for 1 h in T4 ligase buffer. Beads were then washed in high salt buffer, low salt buffer, and 1X T4 PNK Reaction Buffer (NEB). Samples were treated with T4 Poly-nucleotide Kinase (NEB) for 1 h at 37 °C, and the reverse 5′ RNA adapter VRA5 (Table S3) was ligated to the RNA by incubation with T4 RNA ligase I (NEB) for 2 h at room temperature. Following the 5′ end ligation, beads were washed twice in high, binding, and low salt buffers, then once in 0.25X FSS Buffer (ThermoFisher). Reverse transcription was performed using Superscript IV Reverse Transcriptase (ThermoFisher) with 25 pmol of the Illumina TRU-seq RP1 Primer (Table S3). The RT product was eluted from the beads by heating the samples twice at 95°C for 30 s. All libraries were amplified by 12 cycles of PCR with 12.5 pmol of Illumina TRU-seq RPI-index primers (Table S3), excess RP1 primer, and Q5 DNA Polymerase (NEB). The samples were sequenced using a HiSeq instrument with a mid-output 150 bp cycle run.

Pre-processing and analyses of PRO-seq experiments

PRO-seq data from H1 KOs and WT samples were preprocessed as follows: Cutadapt (v3.4, https://doi.org/10.14806/ej.17.1.200) was used to trim adapter sequences from paired-end reads and to remove low-quality bases and reads shorter than 20 nt with -m20 -q 30,30. Processed reads were first aligned to the *Drosophila* genome assembly dm3 using Bowtie2 v1.0[99] with –trim5 6 -p 20 –score-min C,-13,0. We used samtools view (v1.12)[100] with -c and -F260 for counting the primary aligned mapped read to the dm3 genome assembly to perform spike-in normalization in all samples. Paired reads that did not align to dm3, were aligned to the hg19 genome assembly using STAR v2.7.10b[97] with the following parameters: –outMultimapperOrder Random –outFilterIntronMotifs RemoveNoncanonicalUnannotated –outFilterType BySJout –outSAMtype BAM SortedByCoordinate –alignSJoverhangMin 8 –outFilterMismatchNmax 5. RNAPII elongation analysis was performed using bedtools genomecov (v 2.29.1)[116] to convert paired-end bam files to coverage files of active elongation sites with the following specifications: 5′ positions (-5), single-base resolution (-d), and strand-specific (-strand). Output files were converted to bedgraph format using an in-house Bash script, and then to bigWig format using the UCSC tool bedGraphToBigWig (http://hgdownload.soe.ucsc.edu/admin/exe/). Bwtool matrix v1.0[95] was used to cast the PRO-seq read density on gene bodies (protein-coding) including TSSs, exon-intron junctions (excluding first and last exons), and TESs. Human Ensembl transcript annotations (hg19) from the UCSC table browser (https://genome.ucsc.edu/cgi-bin/hgTables) were used for these analyses. Transcripts containing fewer than three exons were excluded. Correct orientation was achieved by casting the forward and reverse bigWig files on coordinates of the opposite strands. For the PRO-seq metagene analysis, gene coverage was aligned at TSS/TES, with mid-sections undergoing either extrapolation or interpolation as per the bwtool tripartite matrix option. To count PRO-seq reads within pausing regions, defined as -50 bp to 250 bp from the annotated transcript start site (TSS), and the elongation regions, extending 251 bp from the TSS to the transcript end site (TES), we employed Bedtools multicov (V.2.29.1) using the -S -bams option. Pausing index was calculated as follows: Pausing Index = (Read Count in Pausing Region / Length of Pausing Region) / (Read Count in Gene Body / Length of Gene Body). To evaluate the elongation pattern across individual transcripts, mean PRO-seq coverage on transcript annotations was calculated using bwtool summary v1.0.[95] Transcripts with less than three exons and mean PRO-seq coverage < 3.5 were excluded. Visualization of PRO-seq reads on selected transcripts was achieved using plotgardner R package v1.4.1[111] with the single-base resolution bedgraph files (1-KO and WT cells) after spike-in normalization. To determine how the H1s KO-mediated changes in RNAPII elongation patterns affect the splicing patterns, we used

ES and IR events with their PSI values from our RNA-seq analyses. Constitutive events with PSI < 5 for IR events and PSI > 95 for ES events as well as cryptic events with PSI > 95 for IR events or PSI < 5 for ES events were excluded. We then used bedtools multicov (v2.29.1)[116] with input bam files (-bams) and different strandedness (-S) to count PRO-seq reads within SSs (20 nt into the intron and 10 nt into the exon) for each alternative splicing event. To isolate RNAPII active elongation sites, we isolated the 2nd read mate from the paired-end bam files using samtools view -bh -f 0x80. SSs (3' and 5') with less than three PRO-seq reads were excluded from the analysis. The effect of individual histones (H1.2, H1.3, and H1.4) on ES and IR events was determined by comparing between two consecutive KO samples or the control (delta PSI and delta PRO-seq reads) as indicated in Figure 4D. Pearson's correlation tests (two-tailed) were performed between the mean delta PSI values and mean delta PRO-seq reads within the SS regions (outliers were removed in both parameters). Maximum Entropy scores of 3' and 5' splice sites flanking exons were obtained using Matt UNIX toolkit v1.3.0[108] with the *get_efeatures* and -f gene_if options. Genomic annotations (hg19) were obtained from the Ensembl Genomes project in GTF format(https://ftp.ensembl.org/pub/grch37/current/gtf/homo_sapiens/Homo_sapiens.GRCh37.87.gtf). To obtain the human genomic sequences (hg19), one FASTA file per chromosome was downloaded from the UCSC genome browser,[105] which were later merged into a single FASTA file.

RT-PCR and DRB treatment

293HEK cells were grown in 6 cm plates to a confluence of 80% and treated for 3 hr with 100 uM DRB (Sigma D1916) to stop the polymerase at the promoter so the transcription could resume at once. Cells were washed from DRB and replaced with new medium for 20 minutes followed by RNA extraction. RT-PCR was performed with RT-FLEX (Quanta), using specific primers using oligo(dT) (for splicing analyses) or Hexamers (for expression analyses) as reverse primer followed by qPCR using KAPA SYBR FAST Universal qPCR kit (KAPA Biosystems) according to the manufacturer's instructions. Primers listed in Table S4.

Gene ontology

Gene ontology (GO) enrichment for genes with altered inclusion levels of exons (from Figure 3) for each of the five H1 variants (H1.0, H1.2, H1.3, H1.4, H1.5). DAVID v6.8 was used for analysis.[109] All terms have an FDR corrected p-value < 0.05 in the two-sided Fisher's exact test. The P-value was transformed to $-\log_{10}$ (FDR p.value).

QUANTIFICATION AND STATISTICAL ANALYSIS

Details regarding the statistical tests, sample sizes, and biological replicates are indicated in the figure legends. More information on how the analyses were performed can be found in the methods section. The statistical significance threshold was set to 0.05 throughout this study.