

Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36

Sérgio Fernandes de Almeida¹, Ana Rita Grosso¹, Frederic Koch², Romain Fenouil², Sílvia Carvalho¹, Jorge Andrade¹, Helena Levezinho¹, Marta Gut³, Dirk Eick⁴, Ivo Gut³, Jean-Christophe Andrau², Pierre Ferrier² & Maria Carmo-Fonseca¹

Several lines of recent evidence support a role for chromatin in splicing regulation. Here, we show that splicing can also contribute to histone modification, which implies bidirectional communication between epigenetic mechanisms and RNA processing. Genome-wide analysis of histone methylation in human cell lines and mouse primary T cells reveals that intron-containing genes are preferentially marked with histone H3 Lys36 trimethylation (H3K36me3) relative to intronless genes. In intron-containing genes, H3K36me3 marking is proportional to transcriptional activity, whereas in intronless genes, H3K36me3 is always detected at much lower levels. Furthermore, splicing inhibition impairs recruitment of H3K36 methyltransferase HYPB (also known as Setd2) and reduces H3K36me3, whereas splicing activation has the opposite effect. Moreover, the increase of H3K36me3 correlates with the length of the first intron, consistent with the view that splicing enhances H3 methylation. We propose that splicing is mechanistically coupled to recruitment of HYPB/Setd2 to elongating RNA polymerase II.

Post-translational modifications of histones by acetylation and methylation are linked to both transcriptional activation and repression, playing a crucial role in regulation of gene expression. Although acetylation has been well known to be reversed by deacetylases, methylation was considered static and enzymatically irreversible until the recent discovery of the first histone demethylase¹. Since then, several lines of evidence indicate that histone methylation is dynamically regulated², but how differentially methylated histones contribute to chromatin function is incompletely understood. In yeast, the Set2 histone methyltransferase binds to elongating RNA polymerase II (Pol II) and mediates both di- and trimethylation of H3K36 (refs. 3–5). Methylated H3K36 subsequently recruits the Rpd3S deacetylase complex that hypoacetylates histones H3 and H4 in the chromatin reassembled behind the elongating polymerase, thereby preventing spurious intragenic transcription^{6–10}. In contrast, the mammalian ortholog HYPB/Setd2 mediates tri- but not dimethylation of H3K36 and does not affect histone acetylation across coding regions¹¹. Unexpectedly, recent studies in *Caenorhabditis elegans* and humans found that exons are enriched for H3K36me3, compared to introns within the same gene^{12,13}, thus revealing a previously unknown link between chromatin and splicing. In a series of subsequent analyses, some groups proposed that the exonic H3K36me3 enrichment reflects differences in nucleosome occupancy between exons and introns^{14,15}, whereas others reported that H3K36me3 is enriched in exons even after taking into account nucleosome occupancy^{13,16,17}. Having identified the exonic H3K36me3 signature, the major challenge remains to understand its

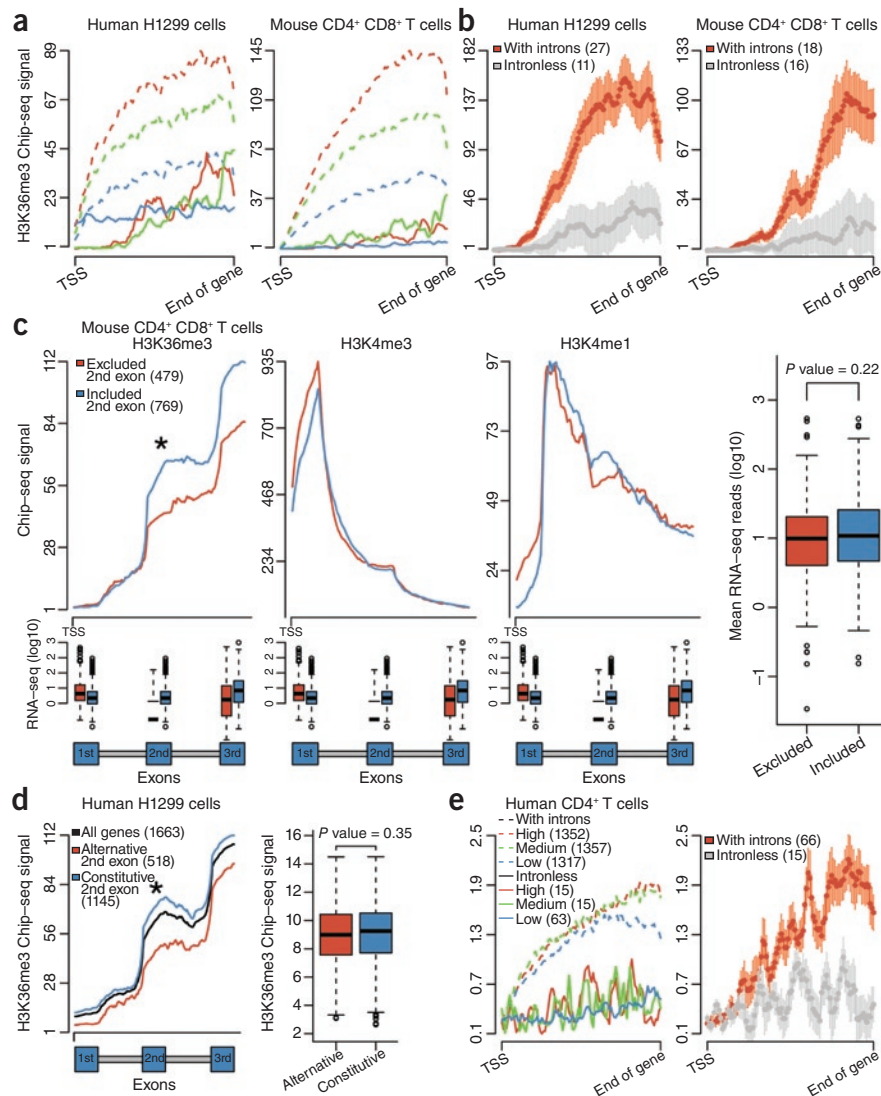
function. One possibility could be that H3K36me3 demarking of exons on actively transcribed genes provides a mechanism to direct splicing. Indeed, several recent reports suggest that histone modifications may have a role in the regulation of splicing, by modulating either the recruitment of the spliceosome or the Pol II elongation rate (recently reviewed in ref. 18). For example, it has been shown that treatment of cells with HDAC inhibitors, which favor histone acetylation, induces exon skipping¹⁹, whereas modifications associated with transcriptional repression enhance exon inclusion^{20,21}. Furthermore, trimethylation of histone H3 on Lys4 facilitates pre-mRNA splicing²², and H3K36 trimethylation affects alternative splicing decisions by modulating the recruitment of splicing regulators through a chromatin-binding protein²³.

An alternative and not mutually exclusive possibility is that splicing influences histone modification, namely H3K36 trimethylation. If so, then the chromatin of exons that are constitutively spliced should have higher H3K36me3 enrichment than exons excluded by alternative splicing. This has been observed in *C. elegans* and mouse cells¹² but not in human cells^{13,17}. If splicing influences H3K36 trimethylation, another prediction is that intron-containing and intronless genes should differ in the marking. Here, we show that in intronless genes, H3K36me3 is much reduced compared to intron-containing genes, irrespective of expression levels. In addition, we observe that splicing inhibition in human cells impairs recruitment of H3K36 methyltransferase HYPB/Setd2 and reduces H3K36me3. Inclusion of alternative exons has the opposite effect, increasing HYPB/Setd2 recruitment and H3K36me3.

¹Instituto de Medicina Molecular, Faculdade de Medicina, Universidade de Lisboa, Lisboa, Portugal. ²Centre d'Immunologie de Marseille-Luminy, Centre National de la Recherche Scientifique (CNRS) UMR6102, Institut National de la Santé et de la Recherche Médicale (INSERM), U631, Université de la Méditerranée, Marseille, France. ³Centre Nacional D'Anàlisi Genòmica, Parc Científic de Barcelona, Barcelona, Spain. ⁴Department of Molecular Epigenetics, Helmholtz Center Munich, Center of Integrated Protein Science, Munich, Germany. Correspondence should be addressed to M.C.-F. (carmo.fonseca@fm.ul.pt), J.-C.A. (andrau@ciml.univ-mrs.fr) or P.F. (ferrier@ciml.univ-mrs.fr).

Received 11 April; accepted 13 July; published online 26 July 2011; doi:10.1038/nsmb.2123

Figure 1 Patterns of H3K36 trimethylation in intron-containing and intronless genes. **(a)** H3K36me3 ChIP-seq average profiles across intron-containing (dashed lines) and intronless genes (solid lines), grouped according to expression level (red, high; green, medium; blue, low). Human H1299 cells: for intron-containing genes, the profiles correspond to 634 genes with high, 593 genes with medium and 436 genes with low level of expression; for intronless genes, the profiles correspond to 11 genes with high, 16 genes with medium and 62 genes with low level of expression. Mouse CD4⁺ CD8⁺ T cells: for intron-containing genes, the profiles correspond to 917 genes with high, 920 genes with medium and 872 genes with low level of expression; for intronless genes, the profiles correspond to 16 genes with high, 13 genes with medium and 63 genes with low expression level. The x axis extends from the TSS to the 3' end of the genes and is scaled along the gene bodies in 100 equally sized bins (introns not depicted). **(b)** H3K36me3 ChIP-seq average profiles for intron-containing and intronless genes with equal average lengths and equal high-expression levels. The x axis is scaled as in **a**. **(c)** ChIP-seq average profiles for indicated chromatin marks on genes containing a second exon that is either excluded or included in the mRNA, as assessed from RNA-seq reads of the same cells. Exons and introns represented in 20 and 40 equally sized bins, respectively. Box plots represent distribution of expression levels for each gene set. **(d)** Same as in **c**, but alternatively or constitutively spliced exons were defined according to the University of California, Santa Cruz (UCSC) Genome Browser database. **(e)** H3K36me3 ChIP-seq average profiles at single nucleosome resolution across intron-containing and intronless genes grouped according to expression level. The graph on the left depicts all genes. The graph on the right depicts genes with equal average lengths and equal high-expression levels. Graphs show mean and s.e.m. **(b,e)**. Asterisks denote $P < 0.005$. Student's *t*-test.



We therefore demonstrate that H3K36me3 marking is directly influenced by splicing, and we propose a model in which co-transcriptional pre-mRNA splicing enhances H3K36 trimethylation by promoting the recruitment of HYPB/Setd2 to elongating Pol II.

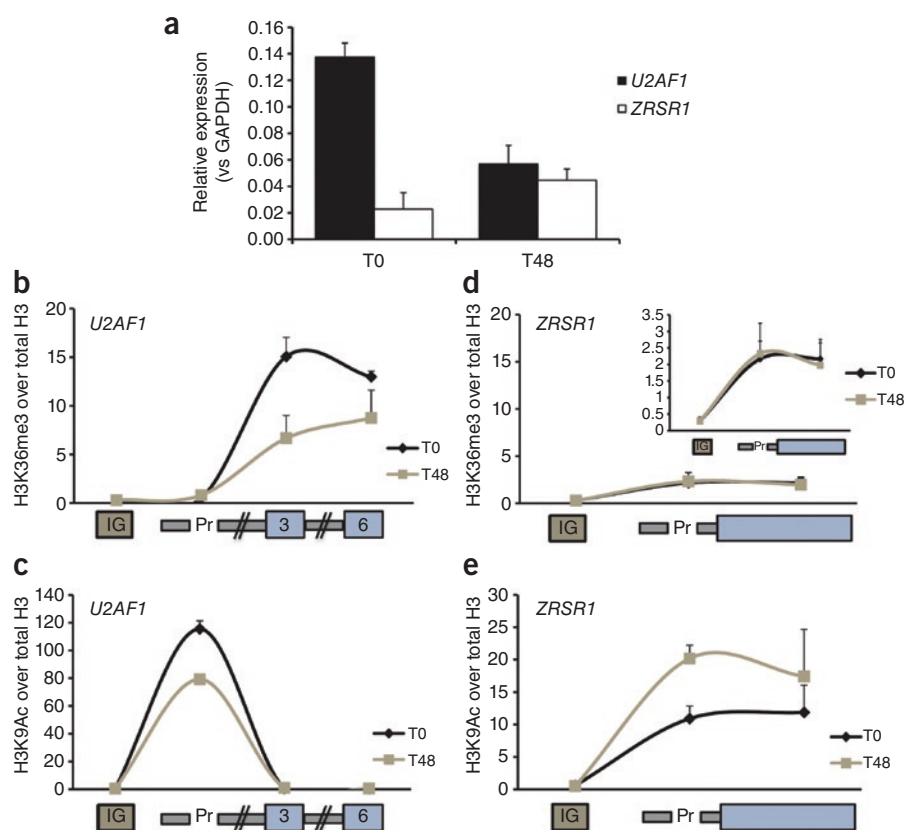
RESULTS

H3K36me3 is enriched in intron-containing genes

We used chromatin immunoprecipitation coupled to high-throughput sequencing (ChIP-seq) to compare the distribution of histone H3 methylation in intron-containing and intronless genes across the mouse and human genomes. We did immunoprecipitations on mouse thymus-derived CD4⁺ CD8⁺ double-positive T cells²⁴ and human H1299 lung carcinoma cells with antibodies directed against mono- or trimethylation of Lys4 and Lys36 of histone H3 (H3K4me1, H3K4me3 and H3K36me3). We also did ChIP-seq on these cells with antibodies that specifically recognize TATA-binding protein (TBP), initiating (Ser5P) and elongating (Ser2P) Pol II. As a correlation was previously reported between transcription and H3K36me3 histone modification^{12,25}, we grouped the intron-containing and intronless genes into three sets according to expression levels (high, medium and low) determined by either poly(A) RNA sequencing (RNA-seq) or microarray analysis of the same cells (**Supplementary Fig. 1**). Genes were aligned at the first and last nucleotides of the annotated

transcripts. Each gene was divided in 100 equally sized bins, and ChIP-seq tags in each bin were counted and averaged. The results show that the level of H3K36me3 remains relatively low along the body of intronless genes, irrespective of their transcriptional activity (**Fig. 1a**), in marked contrast to intron-containing genes that characteristically show low levels of H3K36me3 near the transcript start site (TSS, defined as the first base of the annotated transcript) and a progressive increase in H3K36me3 over the gene body, in proportion to transcriptional activity^{26–28}. A very distinct profile of distribution is observed for H3K4me3 and H3K4me1 (**Supplementary Fig. 2**). In intron-containing genes, H3K4me3 and H3K4me1 levels peak near the TSS and decrease progressively in the rest of the gene body, as previously reported²⁹, whereas intronless genes have high levels of these marks throughout their gene bodies (**Supplementary Fig. 2**). To exclude a putative influence of the different genome-wide average length of intron-containing and intronless genes, we mapped the distribution of H3K36me3 along two equally sized groups of intron-containing and intronless genes (**Fig. 1b** and **Supplementary Fig. 3**). We selected intron-containing genes shorter than 2,880 base pairs (bp) (75th percentile of intronless gene length) and compared groups of intronless and intron-containing genes with an average length of 2,000 bp. This analysis confirmed that H3K36me3 marking on intron-containing genes is considerably higher than on intronless genes in both human

Figure 2 In an intronless gene, H3K36me3 remains low irrespective of transcriptional activation. **(a)** Results of experiments with RNA from undifferentiated (T0) or differentiated (T48) MEL cells that was reverse transcribed with random primers and PCR amplified using primer pairs that detect *U2AF1* pre-mRNA and *ZRSR1* RNA. The amount of PCR product was estimated by qRT-PCR and normalized to the levels of GAPDH. **(b–e)** Results of ChIP assays to compare H3K36me3 in the intronless *ZRSR1* gene and the intron-containing *U2AF1* gene using MEL cells that were either uninduced (T0) or induced to undergo erythroid differentiation for 48 h (T48). ChIP signals for H3K36me3 (**b,d**) and H3K9Ac (**c,e**) were normalized to total H3. Histograms and graphs depict mean and s.d. from at least three independent experiments. The schematic diagrams of the regions amplified by each primer set are represented below each graph. Blue boxes represent exons; IG, intergenic region; Pr, promoter region.



and mouse cells (Fig. 1b and Supplementary Fig. 3). We used TBP and Ser5P as controls and analyzed their distributions (Supplementary Fig. 4). Consistent with previous reports^{30,31}, we observed peaks in TBP and Ser5P Pol II distribution at the TSS, which correlated with mRNA levels in both intron-containing and intronless genes (Supplementary Fig. 4a,b).

Taken together, our results show that irrespective of gene size and expression level, H3K36me3 marking is considerably higher on intron-containing genes than in intronless genes. This difference had not been observed before and suggests that the chromatin landscape can be influenced not only by transcription but also by splicing.

To further investigate a potential relationship between H3K36me3, H3K4me3 and H3K4me1 modifications and splicing, we used poly(A) RNA-seq data from mouse CD4⁺ CD8⁺ T cells to identify genes with similar expression levels that contain a second exon that is either constitutively included in the mature transcript or is excluded by alternative splicing. Plotting H3K36me3 signals across aligned first, second and third exons (Fig. 1c) revealed that the levels of H3K36me3 remain relatively low from the TSS to the first intron-exon boundary, rise sharply over the second exon when this is included, persist constantly over the second intron and rise sharply again over the third exon. In contrast, genes that contain an excluded second exon have much weaker enrichment of H3K36me3 levels at the first intron-exon boundary. No significant ($P > 0.05$) differences are detected in levels of H3K4me1 and H3K4me3 associated with included or excluded second exons (Fig. 1c). Similar results were obtained with human cells (Fig. 1d). These observations suggest that splicing preferentially affects specific chromatin modifications, namely H3K36me3.

Reduced H3K36me3 is independent of total H3 occupancy

Because exons have been shown to contain increased nucleosome occupancy relative to introns^{14,15}, we wondered whether the ChIP-seq results could reflect differences in nucleosome occupancy of intronless genes. To address this possibility, we first took advantage of a previously published dataset describing the genome-wide distribution of H3K36me3 at single-nucleosome resolution in human CD4⁺ T cells³². In the latter study, histone modification signals were resolved to individual nucleosomes using mononucleosome templates generated by micrococ-

cal nuclease (MNase) digestion of native chromatin for ChIP³². After we separated genes into groups according to their expression levels, our results confirmed that H3K36me3 levels are systematically higher in intron-containing genes relative to intronless genes (Fig. 1e). This strongly suggests that the observed differences in H3K36me3 levels are independent of nucleosome occupancy.

Next, we analyzed the intronless *Zrsr1* gene (previously named *U2AF1L1*), which originated recently in evolution as a result of retrotransposition events that occurred independently in rodents and primates³³. This gene encodes a protein with high homology to the small subunit of the heterodimeric splicing factor U2AF, and its expression is specifically activated during erythroid differentiation (Fig. 2a)³⁴. We prepared chromatin extracts from murine erythroleukemia (MEL) cells induced to differentiate *in vitro*, and immunoprecipitations were conducted with antibodies directed against total histone H3, H3K36me3 and H3K9Ac. The total histone H3 ChIP signal was used to normalize H3K36me3 and H3K9Ac values. In the canonical, intron-containing *U2AF1* gene, internal exons from undifferentiated MEL cells were highly enriched in H3K36me3 (by approximately 25- and 50-fold compared to the promoter region or the control intergenic region, respectively; Fig. 2b). After differentiation, the H3K36me3 levels decreased (Fig. 2b). This correlates with lower transcriptional activity of the *U2AF1* gene, confirmed by lower levels of total *U2AF1* transcripts (Fig. 2a) and reduced H3K9Ac at the promoter region (Fig. 2c). Indeed, acetylation of histone H3 Lys9 is well known to occur in nucleosomes associated with promoters of actively transcribed genes^{28,35–37}. In contrast, we detected no H3K36me3 enrichment in the body of the intronless *ZRSR1* gene relative to the promoter region, with levels approximately seven-fold higher compared to the control intergenic region (Fig. 2d). Moreover, no significant ($P > 0.05$) change in H3K36me3 level was detected between undifferentiated and differentiated MEL cells, despite the fact that tran-

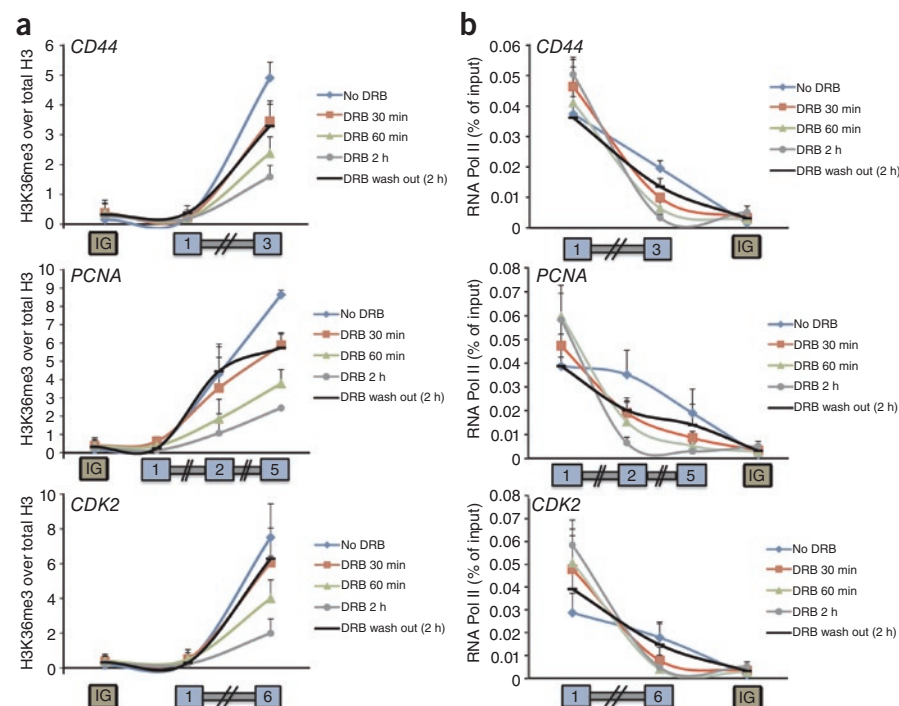


Figure 3 H3K36me3 is highly dynamic in intron-containing genes. Results from experiments with HeLa cells after transcription inhibition with DRB for the indicated times. ChIP signals obtained with H3K36me3 antibody are normalized to total histone H3 (**a**), and ChIP signals obtained with antibody N20 against Pol II (**b**) are represented as percentage of the input. Schematic diagrams represent the regions amplified by each primer set for the indicated genes. An intergenic region (IG) was amplified as a control. Exons are indicated as blue boxes. All graphs depict mean and s.d. from at least three independent experiments.

of total protein extracts revealed that in both RNAi- and meayamycin-treated cells, the level of H3K36me3 was markedly reduced, whereas there was no change in total histone H3 (Fig. 4a). Accumulation of unspliced transcripts from *CD44*, *PCNA* and *CDK2* genes measured by quantitative PCR with reverse-transcriptase (qRT-PCR) confirmed the occurrence of splicing inhibition (Fig. 4b). To determine whether inhibition of splicing had an effect on transcription, we measured for each gene the total amount of RNA present in the cell (total RNA), as well as the number of nascent

transcripts associated with chromatin^{40–43}. Similar levels of RNA were detected (Fig. 4b), arguing that neither treatment with meayamycin nor downregulation of SAP130 reduced transcription of the analyzed genes. ChIP experiments also showed similar levels of exonic Pol II occupancy, in further agreement with the view that splicing inhibition did not affect transcription of *CD44*, *PCNA* and *CDK2* genes (Supplementary Fig. 5). A similar observation was very recently reported⁴⁴ in a study that found that treatment with meayamycin or its analog spliceostatin A does not alter transcription elongation rates or levels of Pol II occupancy along intron-containing model genes.

Splicing enhances H3K36me3 and recruitment of HYPB/Setd2

Together, our results argue that chromatin marking by H3K36me3 is a dynamic process that is dependent on both transcription and splicing. A prediction from this model is that the level of H3K36me3 along the body of actively transcribed intron-containing genes should decrease following inhibition of either transcription or splicing. To address this prediction, we selected three genes that are expressed in HeLa cells (*CD44*, *PCNA* and *CDK2*) and conducted chromatin immunoprecipitations with antibodies directed against Pol II, histone H3 and H3K36me3. Differences in nucleosome occupancy were corrected by normalizing H3K36me3 values to the histone H3 ChIP signal. Consistent with the genome-wide data observed in mouse and human cells (Fig. 1), the level of H3K36me3 in the first exon of each of these three genes was very low and did not differ from that detected in a non-transcribed intergenic region³⁸ (Fig. 3a). However, in downstream internal exons there was a 20- to 50-fold enrichment in H3K36me3 (Fig. 3a). After treating cells with the reversible inhibitor of transcription 5,6-dichloro-1- β -D-ribofuranosylimidazole (DRB) for 30 min, 60 min and 2 h, we observed a progressive decrease in H3K36me3 levels over internal exons (Fig. 3a). By 2 h of treatment, the H3K36me3 levels in internal exons were drastically reduced but remained higher than in the control intergenic region. In contrast, the levels of Pol II in these exons were reduced to levels similar to those in the control intergenic region (Fig. 3b), suggesting that removal of the H3K36me3 mark lags behind clearance of Pol II. Upon removal of DRB, both Pol II and H3K36me3 levels increased again to near normal values, within a 2-h period (Fig. 3a,b).

We next inhibited splicing by either treating cells with meayamycin³⁹ for 4 h or by downregulating expression of the essential splicing factor SAP130 by RNA interference (RNAi) (Fig. 4a). Western blot analysis

transcripts associated with chromatin^{40–43}. Similar levels of RNA were detected (Fig. 4b), arguing that neither treatment with meayamycin nor downregulation of SAP130 reduced transcription of the analyzed genes. ChIP experiments also showed similar levels of exonic Pol II occupancy, in further agreement with the view that splicing inhibition did not affect transcription of *CD44*, *PCNA* and *CDK2* genes (Supplementary Fig. 5). A similar observation was very recently reported⁴⁴ in a study that found that treatment with meayamycin or its analog spliceostatin A does not alter transcription elongation rates or levels of Pol II occupancy along intron-containing model genes.

Our ChIP experiments showed reduced H3K36me3 enrichment in internal exons of the three genes analyzed upon splicing inhibition (Fig. 4c). Because H3K36 trimethylation in mammalian cells is catalyzed by the Huntingtin-interacting protein HYPB/Setd2 histone methyltransferase^{11,45}, we wondered whether inhibition of splicing affects recruitment of this enzyme to the chromatin. Our ChIP experiments using anti-HYPB/Setd2 antibodies revealed a distribution pattern very similar to that of H3K36me3. HYPB/Setd2 enrichment was detected in internal exons relative to the first exon, and a reduction of HYPB/Setd2 levels in internal exons occurred following treatment with meayamycin or RNAi of Sap130 (Fig. 4d). This strongly suggests that recruitment of the histone methyltransferase and subsequent H3K36me3 modification are splicing-dependent. According to this view, the levels of HYPB/Setd2 and H3K36me3 associated with alternatively spliced exons should increase when the splicing pattern is induced to shift from exon exclusion to exon inclusion. To address this possibility, we took advantage of the *CD44* gene, which contains multiple clustered variant exons (Fig. 5a) that can be induced to splice in response to the activation of signaling pathways⁴⁶. As previously reported⁴⁷, treatment of HeLa cells with phorbol myristate acetate (PMA) increases inclusion of variant exons 5 and 10 (exons v5 and v10) in the *CD44* mRNA (Fig. 5b). ChIP experiments further revealed that PMA treatment increased the levels of Pol II (Fig. 5c), H3K36me3 and HYPB/Setd2 (Fig. 5d) in exons v5 and v10 but not in the constitutive exons located upstream and downstream of the alternative exons. A local accumulation of Pol II in the regions encoding the variant exons of the

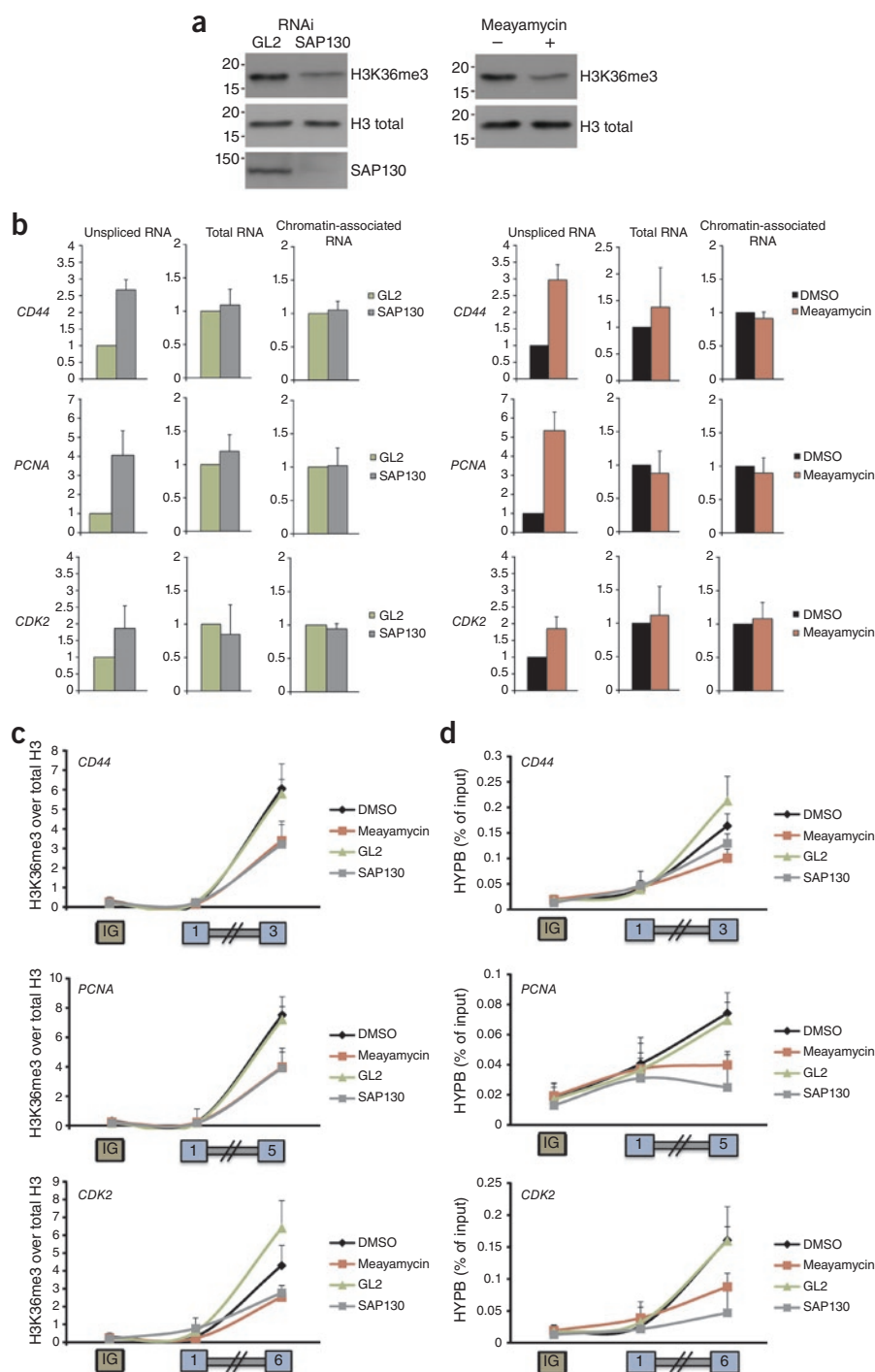


Figure 4 Splicing inhibition reduces H3K36me3 and HYPB/Setd2 recruitment in intron-containing genes. **(a)** Western blot analysis of HeLa cell lysates prepared 48 h after transfection with siRNAs against luciferase (GL2) and SAP130. The blot was probed with the indicated antibodies. Molecular weight markers are shown on the left. **(b)** Results from experiments with RNA from HeLa cells that were either transfected with the indicated siRNAs or treated with meayamycin. qRT-PCR was performed to measure the amount of total RNA present in whole-cell lysates or chromatin fractions and the amount of unspliced RNA from the indicated genes. The amount of PCR product estimated by qRT-PCR was normalized to the levels of U6 small nucleolar RNA (snRNA). Results are shown as the fold change over GL2- or DMSO-treated cells. **(c, d)** ChIP assays were carried out for the indicated genes. Signals obtained with H3K36me3 antibody are normalized to total histone H3 (c), and signals obtained with anti-HYPB/Setd2 antibody are represented as percentage of the input (d). Schematic diagrams represent the regions amplified by each primer set. An intergenic region (IG) was amplified as a control. Exons are indicated as blue boxes. Histograms and graphs depict mean and s.d. from at least three independent experiments.

(ref. 45). We therefore asked whether H3K36me3 mirrors the genome-wide distribution of Ser2P Pol II in mouse T cells. We conducted ChIP-seq experiments with antibodies that specifically recognize Ser5P and Ser2P forms of the CTD, and these showed that Ser5P Pol II peaks in the TSS region, whereas Ser2P Pol II is detected predominantly downstream of the TSS region (Fig. 6a), as recently reported in mouse embryonic stem cells³¹ and through our own work with mouse CD4⁺ CD8⁺ T cells²⁴. The density of Ser2P Pol II increases ~200 to 600 bp downstream of the TSS (Fig. 6a), consistent with the idea that Pol II progresses from initiation to processive elongation in the promoter-proximal region⁵¹. As reported in mouse embryonic stem cells³¹, the density of Ser2P Pol II remains relatively constant throughout the gene body. A very different profile is observed for H3K36me3, which increases progressively over the gene body (Fig. 1a,b,e). To analyze H3K36me3 distribution in relation to exon-intron architecture, we

CD44 gene was previously reported to correlate with increased inclusion of these exons⁴⁷. The PMA treatment effects that we observed were completely blocked when cells were simultaneously treated with PMA and DRB (Fig. 5d). We therefore conclude that recruitment of HYPB/Setd2 to chromatin requires transcriptional activity. Whether Pol II pausing that is associated with the inclusion of alternative exons or splicing itself contributes to further enhance the recruitment of HYPB/Setd2 remains to be determined.

In yeast, the H3K36 methyltransferase Set2 binds directly to Pol II CTD phosphorylated on Ser2 (Ser2P)^{48–50}, and the human homolog HYPB also interacts with hyperphosphorylated Pol II

grouped genes into distinct sets according to the distance from the TSS to the second exon, and in each group, we observed that H3K36me3 enrichment peaks near the 3' end of the first intron (Fig. 6b). Thus, the presence of elongating Ser2P Pol II is not sufficient to account for the pattern of H3K36me3 chromatin marking, which appears to be enhanced by spliceosome assembly as Pol II transcribes past the 3' splice site (Fig. 6b and Supplementary Fig. 6).

DISCUSSION

Histone modifications are being discovered to be regulators of pre-mRNA splicing¹⁸, and our study highlights that the communication between

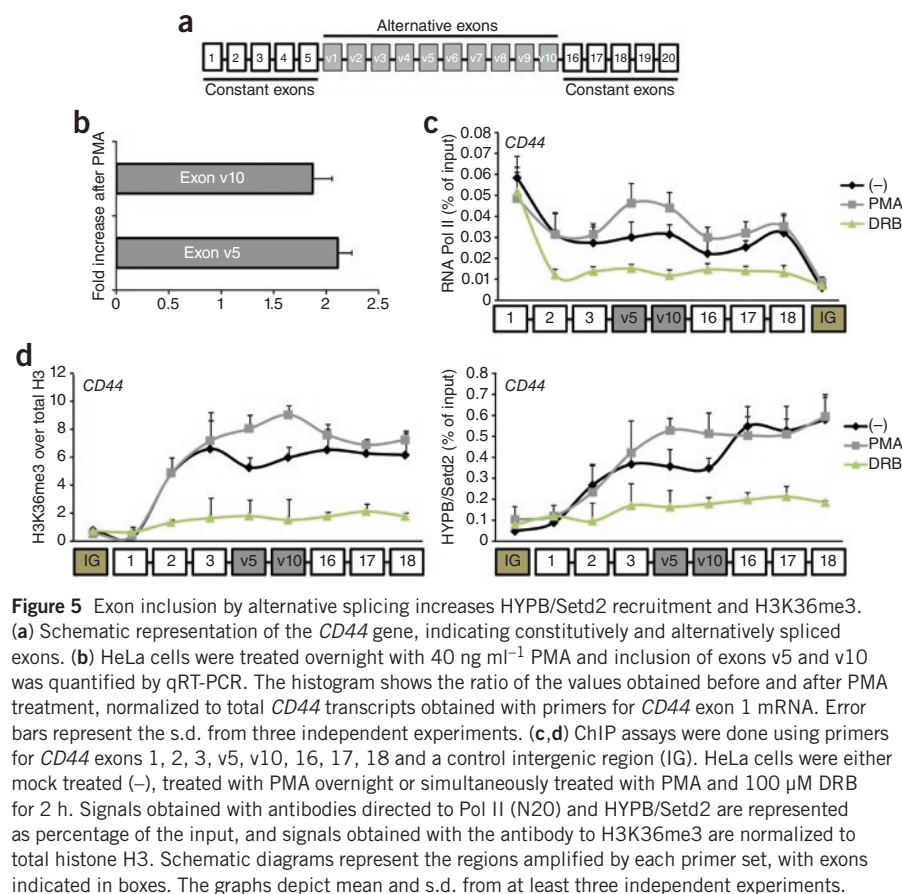


Figure 5 Exon inclusion by alternative splicing increases HYPB/Setd2 recruitment and H3K36me3. (a) Schematic representation of the *CD44* gene, indicating constitutively and alternatively spliced exons. (b) HeLa cells were treated overnight with 40 ng ml⁻¹ PMA and inclusion of exons v5 and v10 was quantified by qRT-PCR. The histogram shows the ratio of the values obtained before and after PMA treatment, normalized to total *CD44* transcripts obtained with primers for *CD44* exon 1 mRNA. Error bars represent the s.d. from three independent experiments. (c, d) ChIP assays were done using primers for *CD44* exons 1, 2, 3, v5, v10, 16, 17, 18 and a control intergenic region (IG). HeLa cells were either mock treated (-), treated with PMA overnight or simultaneously treated with PMA and 100 μ M DRB for 2 h. Signals obtained with antibodies directed to Pol II (N20) and HYPB/Setd2 are represented as percentage of the input, and signals obtained with the antibody to H3K36me3 are normalized to total histone H3. Schematic diagrams represent the regions amplified by each primer set, with exons indicated in boxes. The graphs depict mean and s.d. from at least three independent experiments.

epigenetic mechanisms and RNA processing is probably bidirectional. The first hint that a histone mark could be influenced by splicing was the observation that in *C. elegans* and mouse genomes, the chromatin of exons that are constitutively spliced have higher levels of H3K36me3

than exons excluded by alternative splicing¹². In contrast to earlier reports^{13,17}, we show that higher H3K36me3 marking relative to exons that are excluded is also a characteristic of human exons that are included in mRNA. Although it is considerably reduced, there is still substantial H3K36 trimethylation at excluded alternative exons (Figs. 1c,d and 5d), arguing that splicing is not essential for histone modification. Rather, we consider it more likely that H3K36me3 marking is controlled by the interplay between several cellular processes, including splicing. It is therefore not unexpected that differences in H3K36me3 levels may not be detected in particular genes or splicing events, as recently reported for the alternative exons of *CD45* (*PTPRC*) and *YPEL5* genes¹⁷.

Our study demonstrates for the first time that intronless genes contain much lower H3K36me3 marking compared to intron-containing genes, irrespective of expression levels. Most importantly, we show that when splicing is experimentally inhibited, recruitment of H3K36 methyltransferase HYPB/Setd2 is reduced, whereas forcing the inclusion of alternative exons has the opposite effect of increasing HYPB/Setd2 recruitment and H3K36 trimethylation. Taken together, our data suggest a model for H3K36me3 marking in which the H3K36me3-specific methyltransferase HYPB/Setd2 binds to Ser2P Pol II and deposits the mark on nucleosomes as the

polymerase elongates^{45,48–50}; co-transcriptional spliceosome assembly and splicing additionally enhance recruitment of HYPB/Setd2, leading to enrichment of H3K36me3 associated with splicing (Fig. 6c). This could explain why H3K36me3 levels increase progressively along intron-containing genes but remain relatively constant on intronless genes (Fig. 1b,e). Although further studies are needed to characterize in detail the molecular interactions between chromatin and the splicing machinery, splicing factors have been reported to associate with both chromatin and Pol II (refs. 22,52), making such a model plausible.

METHODS

Methods and any associated references are available in the online version of the paper at <http://www.nature.com/nsmb/>.

Accession codes. ChIP-seq data for human H1299 cells can be accessed in the Gene Expression Omnibus (GEO) database (code: GSE30902).

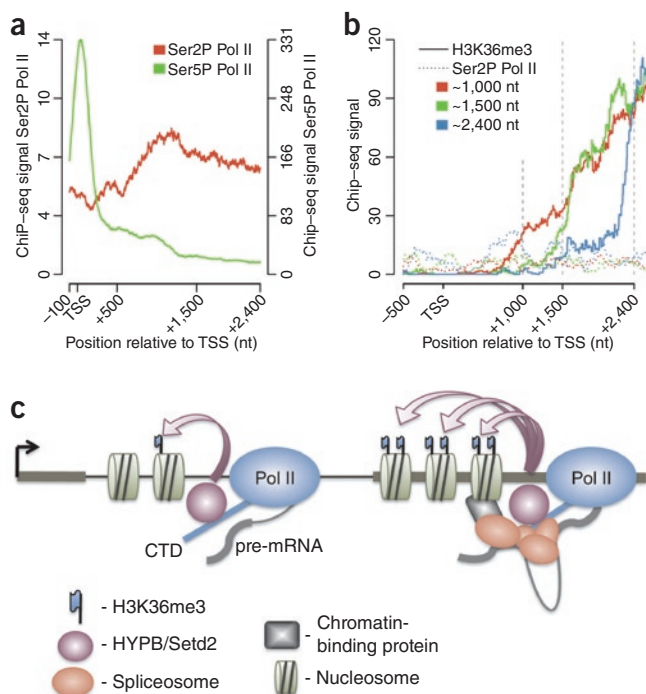


Figure 6 H3K36me3 does not mirror Ser2P Pol II occupancy. (a, b) ChIP-seq average profiles for initiating Pol II (Ser5P), elongating Pol II (Ser2P) and H3K36me3 along highly and medium expressed intron-containing genes. Panel (b) depicts average profiles for H3K36me3 and Ser2P Pol II along genes that contain the second exon located within 1,000 $\pm$ 30 (red), 1,500 $\pm$ 30 (green) or 2,400 $\pm$ 30 (blue) nucleotides from the TSS. (c) Splicing enhances H3K36me3 through HYPB/Setd2 recruitment. HYPB/Setd2 binds to the hyperphosphorylated CTD of Pol II and methylates the nucleosomes reassembled behind the elongating polymerase. As Pol II transcribes through the 3' splice site, allowing for spliceosome assembly on the pre-mRNA, recruitment of HYPB/Setd2 is enhanced, resulting in higher levels of H3K36me3.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We are grateful to K. Koide (University of Pittsburgh), for kindly providing the meayamycin used in this study. We are also thankful to S. Marinho (Faculdade de Medicina, Universidade de Lisboa) for technical assistance. This work was supported by grants from Fundação para a Ciência e Tecnologia (PTDC-BIA-BCM-101575-2008 to M.C.-F. and PTDC-BIA-BCM-111451-2009 to S.F.deA.) and the European Commission (MRTN-CT-2006-035733 to M.C.-F. and P.F.). S.F.deA. and A.R.G. are supported by fellowships from Fundação para a Ciência e Tecnologia (SFRH-BPD-34679-2007 and SFRH-BPD-62911-2009). Work in the P.F. laboratory is supported by institutional grants from Institut National de la Santé et de la Recherche Médicale (INSERM) and Centre National de la Recherche Scientifique (CNRS), and by specific grants from Fondation Princesse Grace de Monaco, the Agence Nationale de la Recherche (ANR), the Institut National du Cancer (INCa) and the Commission of the European Communities. F.K. was supported by a Marie Curie research training fellowship (MRTN-CT-2006-035733); and is now supported by Association pour la Recherche sur le Cancer (ARC). R.F. was supported by Marseille-Nice Genopole; and is now supported by a grant from CNRS.

AUTHOR CONTRIBUTIONS

S.F.deA. and M.C.-F. conceived the project and designed the experiments. S.F.deA., S.C., J.A., H.L. conducted and analyzed the wet-lab experiments. F.K., I.G., J.-C.A. and P.F. conceived the framework of the ChIP-seq studies. F.K. and J.-C.A. designed the ChIP-seq experiments. D.E. produced and provided the Ser2P and Ser5P Pol II antibodies. R.F. and F.K. carried out the bioinformatics preprocessing of ChIP-seq data. ChIP-seq and RNA-seq preprocessing materials were prepared by F.K., M.G. and I.G. conducted all ChIP-seq- and RNA-sequencing experiments. A.R.G. carried out the bioinformatic analysis of microarray, ChIP-seq and RNA-seq data. S.F.deA., A.R.G. and M.C.-F. wrote the manuscript. All authors reviewed the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/nsmb/>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Shi, Y. *et al.* Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**, 941–953 (2004).
- Shi, Y. & Whetstone, J.R. Dynamic regulation of histone lysine methylation by demethylases. *Mol. Cell* **25**, 1–14 (2007).
- Li, J., Moazed, D. & Gygi, S.P. Association of the histone methyltransferase Set2 with RNA polymerase II plays a role in transcription elongation. *J. Biol. Chem.* **277**, 49383–49388 (2002).
- Li, B., Howe, L., Anderson, S., Yates, J.R. III & Workman, J.L. The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **278**, 8897–8903 (2003).
- Xiao, T. *et al.* Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev.* **17**, 654–663 (2003).
- Carrozza, M.J. *et al.* Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**, 581–592 (2005).
- Joshi, A.A. & Struhl, K. Eaf3 chromodomain interaction with methylated H3–K36 links histone deacetylation to Pol II elongation. *Mol. Cell* **20**, 971–978 (2005).
- Keogh, M.C. *et al.* Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**, 593–605 (2005).
- Li, B. *et al.* Combined action of PHD and chromo domains directs the Rpd3S HDAC to transcribed chromatin. *Science* **316**, 1050–1054 (2007).
- Li, B. *et al.* Infrequently transcribed long genes depend on the Set2/Rpd3S pathway for accurate transcription. *Genes Dev.* **21**, 1422–1430 (2007).
- Edmunds, J.W., Mahadevan, L.C. & Clayton, A.L. Dynamic histone H3 methylation during gene induction: HYPB/Set2 mediates all H3K36 trimethylation. *EMBO J.* **27**, 406–420 (2008).
- Kolasinska-Zwier, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.* **41**, 376–381 (2009).
- Spies, N., Nielsen, C.B., Padgett, R.A. & Burge, C.B. Biased chromatin signatures around polyadenylation sites and exons. *Mol. Cell* **36**, 245–254 (2009).
- Tilgner, H. *et al.* Nucleosome positioning as a determinant of exon recognition. *Nat. Struct. Mol. Biol.* **16**, 996–1001 (2009).
- Schwartz, S., Meshorer, E. & Ast, G. Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* **16**, 990–995 (2009).
- Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C. & Komorowski, J. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.* **19**, 1732–1741 (2009).
- Huff, J.T., Plocik, A.M., Guthrie, C. & Yamamoto, K.R. Reciprocal intronic and exonic histone modification regions in humans. *Nat. Struct. Mol. Biol.* **17**, 1495–1499 (2010).
- Luco, R.F., Allo, M., Schor, I.E., Kornblitt, A.R. & Misteli, T. Epigenetics in alternative pre-mRNA splicing. *Cell* **144**, 16–26 (2011).
- Schor, I.E., Rascovan, N., Pelisch, F., Allo, M. & Kornblitt, A.R. Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc. Natl. Acad. Sci. USA* **106**, 4325–4330 (2009).
- Alló, M. *et al.* Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* **16**, 717–724 (2009).
- Saint-André, V., Batsche, E., Rachez, C. & Muchardt, C. Histone H3 lysine 9 trimethylation and HP1gamma favor inclusion of alternative exons. *Nat. Struct. Mol. Biol.* **18**, 337–344 (2011).
- Sims, R.J. III *et al.* Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell* **28**, 665–676 (2007).
- Luco, R.F. *et al.* Regulation of alternative splicing by histone modifications. *Science* **327**, 996–1000 (2010).
- Koch, F. *et al.* Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nat. Struct. Mol. Biol.* published online, doi:10.1038/nsmb.2085 (17 July 2011).
- Li, B., Carey, M. & Workman, J.L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
- Strahl, B.D. *et al.* Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol. Cell* **22**, 1298–1306 (2002).
- Bannister, A.J. *et al.* Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J. Biol. Chem.* **280**, 17732–17736 (2005).
- Pokholok, D.K. *et al.* Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**, 517–527 (2005).
- Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R. & Young, R.A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88 (2007).
- van Werven, F.J., van Teeffelen, H.A., Holstege, F.C. & Timmers, H.T. Distinct promoter dynamics of the basal transcription factor TBP across the yeast genome. *Nat. Struct. Mol. Biol.* **16**, 1043–1048 (2009).
- Rahl, P.B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
- Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
- Mollet, I., Barbosa-Morais, N.L., Andrade, J. & Carmo-Fonseca, M. Diversity of human U2AF splicing factors. *FEBS J.* **273**, 4807–4816 (2006).
- Grosso, A.R. *et al.* Tissue-specific splicing factor gene expression signatures. *Nucleic Acids Res.* **36**, 4823–4832 (2008).
- Liang, G. *et al.* Distinct localization of histone H3 acetylation and H3–K4 methylation to the transcription start sites in the human genome. *Proc. Natl. Acad. Sci. USA* **101**, 7357–7362 (2004).
- Schubeler, D. *et al.* The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **18**, 1263–1271 (2004).
- Bernstein, B.E. *et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169–181 (2005).
- Listerman, I., Sapra, A.K. & Neugebauer, K.M. Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat. Struct. Mol. Biol.* **13**, 815–822 (2006).
- Albert, B.J. *et al.* Meayamycin inhibits pre-messenger RNA splicing and exhibits picomolar activity against multidrug-resistant cells. *Mol. Cancer Ther.* **8**, 2308–2318 (2009).
- Wuarin, J. & Schibler, U. Physical isolation of nascent RNA chains transcribed by RNA polymerase II: evidence for cotranscriptional splicing. *Mol. Cell Biol.* **14**, 7219–7225 (1994).
- West, S., Proudfoot, N.J. & Dye, M.J. Molecular dissection of mammalian RNA polymerase II transcriptional termination. *Mol. Cell* **29**, 600–610 (2008).
- Dye, M.J., Gromak, N. & Proudfoot, N.J. Exon tethering in transcription by RNA polymerase II. *Mol. Cell* **21**, 849–859 (2006).
- Pandya-Jones, A. & Black, D.L. Co-transcriptional splicing of constitutive and alternative exons. *RNA* **15**, 1896–1908 (2009).
- Brody, Y. *et al.* The *in vivo* kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol.* **9**, e1000573 (2011).
- Sun, X.J. *et al.* Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J. Biol. Chem.* **280**, 35261–35271 (2005).
- König, H., Ponta, H. & Herrlich, P. Coupling of signal transduction to alternative pre-mRNA splicing by a composite splice regulator. *EMBO J.* **17**, 2904–2913 (1998).
- Batsche, E., Yaniv, M. & Muchardt, C. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat. Struct. Mol. Biol.* **13**, 22–29 (2006).
- Kizer, K.O. *et al.* A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol. Cell Biol.* **25**, 3305–3316 (2005).
- Li, M. *et al.* Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc. Natl. Acad. Sci. USA* **102**, 17636–17641 (2005).
- Vojnic, E., Simon, B., Strahl, B.D., Sattler, M. & Cramer, P. Structure and carboxyl-terminal domain (CTD) binding of the Set2 SRI domain that couples histone H3 Lys36 methylation to transcription. *J. Biol. Chem.* **281**, 13–15 (2006).
- Fuda, N.J., Ardehali, M.B. & Lis, J.T. Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**, 186–192 (2009).
- Allemand, E., Batsche, E. & Muchardt, C. Splicing, transcription, and chromatin: a ménage à trois. *Curr. Opin. Genet. Dev.* **18**, 145–151 (2008).

ONLINE METHODS

Cells. T cells were isolated from mouse thymuses as previously described²⁴. Briefly, cell suspensions were obtained from thymuses extracted from wild-type mice, and CD4⁺ CD8⁺ double-positive cells were subsequently sorted. For biological replicates, equally aged littermates were used. H1299 cells were grown in Dulbecco's modified Eagle medium (DMEM), supplemented with 10% (v/v) FCS, 100 U ml⁻¹ penicillin-streptomycin and 2 mM L-glutamine. HeLa cells were grown as monolayers in DMEM, supplemented with 10% (v/v) FBS, 2 mM L-glutamine and 100 U ml⁻¹ penicillin-streptomycin. The murine erythroleukemia (MEL) cell line C88 was grown in DMEM-GlutaMAX supplemented with 10% (v/v) FCS. Erythroid differentiation was induced by addition of 2% (v/v) DMSO (Sigma), 5% (w/v) BSA (Sigma) and 1.8 × 10⁻³ mM iron dextran (Sigma), as described previously⁵³. All cell culture reagents were from Invitrogen. To inhibit transcription, HeLa cells were treated with 100 μM DRB (Sigma) for 30 min, 60 min or 2 h. To resume transcription, DRB was washed out after 2 h and the cells incubated for an additional 2 h in drug-free medium. To inhibit splicing, HeLa cells were treated with 50 nM mecamycin for 4 h.

ChIP sequencing. ChIP-seq analysis was carried out as previously described²⁴. Briefly, H1299 and purified murine CD4⁺ CD8⁺ cells were incubated with 1% (v/v) formaldehyde for cross-linking and the reaction was quenched with 250 mM glycine. Cells were lysed, and chromatin was collected by centrifugation and sonicated. After the addition of Triton X-100 to a final concentration of 1% (v/v) and subsequent centrifugation, supernatants were combined for batch processing. ChIP experiments were carried out using Dynabeads (Invitrogen) coated with protein G. For biological replicates, all steps were repeated using independent samples. Immunoprecipitated chromatin and input extracts were incubated with SDS, digested with RNase A and treated with proteinase K. DNA was subsequently purified and quantified. At least 1 ng of ChIP or input DNA was used for library preparation according to the Illumina protocol. The following antibodies were used: rabbit polyclonal antibody directed against total Pol II (N-20, Santa Cruz Biotechnology); rat monoclonal antibodies against Ser2P (3E10) and Ser5P (3E8)⁵⁴; rabbit polyclonal antibodies against H3K4me1, H3K4me3 and H3K36me3 (Abcam references ab8895, ab8580 and ab9050, respectively); and mouse monoclonal antibody against TBP (Diagenode, reference Mab-TBPCSH-100).

RNA sequencing. Strand-specific sequencing of mouse T-cell RNA was done as previously described²⁴. Total RNA was extracted using TRIzol (Invitrogen), and DNA was digested. Poly(A) RNA was purified from total RNA using the Illumina purification kit. Poly(A) RNA samples were fragmented to ~150 bp using RNase H (Ambion). Library preparation using the Illumina Small RNA Sample Prep Kit was carried out according to the manufacturer's instructions. The library was reverse transcribed, PCR amplified, size selected and sequenced according to the kit instructions, except that 36-bp tags were sequenced instead of 75-bp tags.

Bioinformatics analysis of ChIP- and RNA-sequencing data. ChIP-seq data for H3K36me3 in human CD4⁺ T cells was previously published³². Bioinformatics pre-processing of sequenced tags was done as previously described²⁴ (see **Supplementary Methods** for details). To generate average ChIP-seq profiles, we used the mm9 and hg18 UCSC KnownGene and Refseq gene annotations⁵⁵ to extract values from wiggle files associated with expressed genes. Grouping of genes as intron-containing or intronless, and alternative or constitutive, second exon was done according to UCSC database⁵⁵ guidelines. Murine and human genes were divided according to expression levels (high, medium and low), by splitting them in three equally sized groups according to RNA-seq levels (mean reads number for all exons) or microarray expression data, respectively. Microarray data for human cells were obtained from GSE10437 for CD4⁺ T cells (GEO, <http://www.ncbi.nlm.nih.gov/geo/>) and E-TABM-767 for lung carcinoma cell line (H1299) (ArrayExpress, <http://www.ebi.ac.uk/arrayexpress>). Microarray data were analyzed using the R packages affy⁵⁶ and annaffy (<http://www.bioconductor.org/packages/2.2/bioc/html/annaffy.html>). The identification of genes with excluded or included second exon

was based on the ratio of the average number of poly(A) RNA-seq reads between the first and second exons for each gene: genes with a two-fold or greater increase were considered as part of the 'excluded second exon' group.

RNA interference. To reduce the levels of SAP130 by RNAi, we used siRNA duplexes. As an unspecific siRNA control, GL2—a sequence targeting the firefly luciferase gene—was used⁵⁷. HeLa cells were plated 1 d before transfection so that they were 50–60% confluent at the time of transfection. The siRNA duplexes were transfected using Lipofectamine 2000 transfection reagent (Invitrogen) according to the manufacturer's protocol. At 24 h after the first transfection, cells were re-transfected with the same siRNA duplex and harvested on the following day. The siRNA sequences are listed in **Supplementary Table 1**.

RNA isolation and quantitative reverse-transcriptase PCR. Total RNA was extracted from HeLa and MEL cells using TRIzol (Invitrogen). Chromatin fractions were prepared from HeLa cells as described previously⁵⁸, and chromatin-associated nascent RNA was extracted with TRIzol. cDNA was made using Superscript II Reverse Transcriptase (Invitrogen) according to the manufacturer's instructions. qRT-PCR was carried out in the 7000 Real-Time PCR System (Applied Biosystems), using SYBR Green PCR Master Mix (Applied Biosystems). The relative expression was estimated as follows: $2^{C_t(\text{reference}) - C_t(\text{sample})}$, where C_t (reference) and C_t (sample) are mean threshold cycles of qRT-PCR done in duplicate on cDNA samples from GAPDH or U6 snRNA (reference) and the cDNA from our genes of interest (sample), respectively. All primer sequences are presented in **Supplementary Table 2**.

Western blot. Whole-cell lysates were prepared, resolved and transferred as described⁵⁸. Incubations with primary antibodies were followed by incubations with the appropriate secondary antibodies (Bio-Rad) and by detection using enhanced luminescence substrate (Amersham/GE Healthcare).

Chromatin immunoprecipitation. ChIP was carried out on HeLa and MEL cells as previously described⁵⁸, with some modifications. After incubation with the specific antibodies and beads, DNA from immunoprecipitated samples was extracted with Chelex 100 (Bio-Rad) as previously described⁵⁹. The total number of cells was kept constant between experimental batches in order to yield similar amounts of input chromatin and thus avoid variability due to changes in precipitation efficiency. Quantitative PCR was carried out as described for qRT-PCR above. The relative occupancy of the immunoprecipitated protein at each DNA site was estimated as follows: $2^{C_t(\text{input}) - C_t(\text{IP})}$, where C_t (input) and C_t (IP) are mean threshold cycles of qRT-PCR done in duplicate on DNA samples from input and specific immunoprecipitations, respectively. Gene-specific and intergenic-region primer pairs are presented in **Supplementary Table 2**. The following antibodies were used: rabbit polyclonal antibodies anti-Pol II (N20; Santa Cruz Biotechnology); anti-histone H3 (ab1791; Abcam); anti-H3K36me3 (ab9050; Abcam); anti-H3K9Ac (ab10812; Abcam) and anti-HYPB (ab69836; Abcam); anti-SAP130 (SF3B3, ab86992; Abcam).

53. Volloch, V. & Housman, D. Terminal differentiation of murine erythroleukemia cells: physical stabilization of end-stage cells. *J. Cell Biol.* **93**, 390–394 (1982).
54. Chapman, R.D. *et al.* Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. *Science* **318**, 1780–1782 (2007).
55. Rhead, B. *et al.* The UCSC Genome Browser database: update 2010. *Nucleic Acids Res.* **38**, D613–D619 (2010).
56. Gautier, L., Cope, L., Bolstad, B.M. & Irizarry, R.A. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
57. Elbashir, S.M. *et al.* Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498 (2001).
58. de Almeida, S.F., Garcia-Sacristan, A., Custodio, N. & Carmo-Fonseca, M. A link between nuclear RNA surveillance, the human exosome and RNA polymerase II transcriptional termination. *Nucleic Acids Res.* **38**, 8015–8026 (2010).
59. Nelson, J.D., Denisenko, O., Sova, P. & Bomsztyk, K. Fast chromatin immunoprecipitation assay. *Nucleic Acids Res.* **34**, e2 (2006).