Transcriptional repression by FACT is linked to regulation of chromatin accessibility at the promoter of ES cells

Constantine Mylonas[1] and Peter Tessarz[1,2,3]

1 - Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Str. 9b, 50931 Cologne, Germany

2 - Cologne Excellence Cluster on Cellular Stress Responses in Ageing Associated Diseases (CECAD), University of Cologne, Joseph-Stelzmann-Strasse 26, 50931 Cologne, Germany

3 - corresponding author: ptessarz@age.mpg.de

Word count: 4,210

The conserved and essential histone chaperone FACT (Facilitates Chromatin Transcription) reorganizes nucleosomes during DNA transcription, replication and repair and ensures both, efficient elongation of RNA Pol II and nucleosome integrity. In mammalian cells, FACT is a heterodimer, consisting of SSRP1 and SUPT16. Here, we show that in contrast to yeast, FACT accumulates at the transcription start site of genes reminiscent of RNA Polymerase II profile. Depletion of FACT in mouse embryonic stem cells leads to up-regulation of pro-proliferative genes and key pluripotency factors concomitant with hyper-proliferation of mES cells. Using MNase-, ATAC-, and Nascent Elongating Transcript Sequencing (NET-seq) we show that up-regulation of genes coincides with loss of nucleosomes upstream of the TSS and concomitant increase in antisense transcription, indicating that FACT impacts the promoter architecture to regulate expression of these genes. Finally, we demonstrate a role for FACT in cell fate determination and show that FACT depletion primes ES cells for the neuronal lineage.

Introduction

The basic functional unit of chromatin is the nucleosome consisting of around 147 bp of DNA wrapped around an octamer of histone proteins – two copies each of histones H2A, H2B, H3 and H4. *In vitro*, chromatinized DNA templates are refractory to transcription suggesting that the nucleosome might provide a barrier for the elongating RNA polymerase. Using elegant biochemical fractionation assays coupled to *in vitro* transcription assays, FACT was initially characterised as a factor that alleviated the repressive nature of chromatin *in vitro* (Orphanides, Wu, Lane, Hampsey, & Reinberg, 1999). Meanwhile, it has been demonstrated that FACT can cooperate with all RNA polymerases in the cell and ensure both, efficient transcription elongation and nucleosome integrity. Both FACT subunits are highly conserved across all eukaryotes with the exception of an HMG-like domain present in SSRP1 but absent in the yeast homolog Pob3. In yeast, an HMG domain protein named Nhp6 has been proposed to provide the DNA binding capacity of FACT (Formosa et al., 2001).

The molecular basis for FACT activity has long remained elusive. However, recent biochemical and structural studies are starting to elucidate how FACT engages nucleosomes (Hondele et al., 2013; Hsieh et al., 2013; Kemble, McCullough, Whitby, Formosa, & Hill, 2015; Winkler & Luger, 2011). Via its several domains, FACT binds to multiple surfaces on the nucleosome octamer and acts by shielding histone-DNA interactions. Initially, it was proposed that FACT would evict an H2A/B dimer from the nucleosome in front of the polymerase and then reinstate nucleosome integrity in its wake. However, other data suggests that this dimer replacement is not part of FACT function as it leaves the histone composition of the nucleosome intact (Formosa, 2012). Based on recent biochemical data (Hsieh et al., 2013), a model emerges in which RNA Pol II enters the nucleosome and partially uncoils the nucleosomal DNA. At the same time, FACT binds to the proximal and distal H2A/H2B dimer and these FACT–dimer interactions facilitate nucleosome survival.

Although the genetics and biochemistry of FACT are relatively well understood, it is not known whether cell-type dedicated functions are conferred by this histone chaperone. Interestingly, genome-wide expression analyses across cell and tissue types implicate a role of FACT in maintaining an undifferentiated state. Depletion of FACT subunits leads to growth reduction in transformed, but not in immortalized cells (Garcia et al., 2013), indicating that FACT is essential for tumour growth, but not for proliferation of untransformed cells. Finally, FACT regulates the expression of Wnt-target genes during osteoblast differentiation in mesenchymal stem cells and its deletion leads to a differentiation skew (Hossan et al., 2016). Taken together, these data suggested a more specific role for the FACT complex in undifferentiated cells as previously assumed.

Recent studies have demonstrated that RNA Pol II can transcribe in both sense and anti-sense directions near many mRNA genes (Kwak, Fuda, Core, & Lis, 2013; Mayer et al., 2015). At these so-called bidirectional promoters, RNA Pol II initiates transcription and undergoes promoter-proximal pausing in both the sense (at the protein-coding TSS) and anti-sense orientation (Kwak et al., 2013; Mayer et al., 2015). Divergent transcription is often found at mammalian promoters that are rich in CpG content, but lack key core promoter elements such as the TATA motif (Scruggs et al., 2015). A broader nucleosome free region (NFR) in the promoter region is often accompanied by divergent transcription, and can lead to binding of more transcription factors resulting in higher gene activity (Scruggs et al., 2015).

Here, we have confirmed an indispensable role of FACT in undifferentiated cells based on the expression levels of both FACT subunits and, thus chose mouse embryonic stem cells as a model to investigate how FACT might shape the transcriptome and maintain an undifferentiated state. To achieve this, we performed ChIP- and RNA-seq to identify genes bound and regulated by FACT. To address at a mechanistic level how FACT might regulate transcription in ES cells, we combined this analysis with MNase digestion of chromatin coupled to deep sequencing (MNase-seq), Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq), and

Nascent Elongating Transcript Sequencing (NET-seq). Using these approaches, we have identified a specific gene cluster comprising of genes involved in embryogenesis/ neuronal development that are up-regulated upon FACT depletion. In addition, we observed a concomitant increase in chromatin accessibility around the transcription start site, suggesting that maintenance of nucleosomes at this position by FACT is part of the mechanism how FACT impacts on the regulation of these genes. Finally, our data support a role of FACT in the maintenance of a pluripotent state by showing that its depletion leads to faster differentiation into the neuronal lineage.

Results

Occupancy of FACT correlates with marks of active gene expression

High expression of FACT has been associated with stem or less-differentiated cells (Garcia et al., 2011). Indeed, we were able to confirm that low FACT levels correlate with highly differentiated cell lines as opposed to stem and cancer cells (**Supplementary Fig. 1A**). In addition, differentiation of murine ES cells into terminally differentiated cardiomyocytes (Wamstad et al., 2012) reveals that FACT levels diminish throughout the course of differentiation (**Supplementary Fig. 1B**). Thus, we chose to explore how FACT contributes to the transcriptome of undifferentiated cells using mouse ES cells. Initially, we applied to mESCs a chromatin immunoprecipitation and sequencing (ChIP-seq) assay to identify potential DNA binding regions for both FACT subunits. Subsequently, we examined FACT co-enrichment with several other transcription factors, histone marks, and chromatin remodellers over the gene body area of all uniquely annotated protein-coding genes (n = 11,305). High correlation scores were observed between SSRP1, SUPT16, H3K4me3, H3K27ac, and Pol II variants (Pol II S5ph, Pol II S2ph) confirming the role of FACT in active gene expression (**Fig. 1A & Supplementary**

Fig. 1E). A good correlation was also observed between both FACT subunits and Chd1, in line with data demonstrating physical interaction and co-localization in mammalian cells (Kelley, Stokes, & Perry, 1999). However, only a moderate correlation was observed between FACT and H3K36me3 on a genome wide level despite the fact that H3K36me3 directly recruits FACT to actively transcribed genes (Carvalho et al., 2013). We suspect that the strong enrichment of FACT subunits around the TSS might mask this potential correlation. Nevertheless, FACT subunits also co-localize to the gene body of actively transcribed genes and enrich towards the TES, similarly to H3K36me3 (**Supplementary Fig. 1C,D**). Pearson's correlation among FACT and active marks remained elevated when we focused on promoter and enhancer regions (*n* = 19,461) (**Fig. 1B**). Both subunits displayed very similar binding pattern to each other over the transcription start site (TSS) of all the annotated genes and were tightly linked to H3K4me3 levels (**Fig. 1C**).

Regulation of gene expression by FACT

To investigate how FACT orchestrates transcriptional regulation in mESCs, we depleted SSRP1 levels using short hairpin RNAs (shRNA – **Supplementary Fig. 2A**). Importantly, this also led to a simultaneous depletion of SUPT16 levels as assessed by Mass spectrometry (Supplementary Table 7). This interdependence of the two FACT subunits has been observed before (Garcia et al., 2013). Surprisingly, we observed an increase in mESC proliferation following *Ssrp1* knock-down (KD) as measured by proliferation rate via MTT (metabolic activity measurement) cell proliferation assays using independent shRNAs (**Fig. 2A & Supplementary Fig. 2B**). This is in contrast to previously published data from tumour cell lines, in which proliferation rates decrease and also from terminally differentiated cells, where FACT depletion has no effect on proliferation (Garcia et al., 2013). Subsequently, we sequenced the whole transcriptome (RNA-

seq). In total, we characterised 3,003 differentially expressed genes; 1,655 down-regulated and 1,348 up-regulated (**Fig. 2B**). Down-regulated genes were over-represented for pathways involved in development, while up-regulated genes were involved in metabolic processes and positive regulation of proliferation (**Fig. 2C**), indicating that the change in the transcriptome accounts for the faster proliferation rates. These results suggest that FACT impacts developmental processes and negatively controls cell proliferation in mES cells by controlling gene expression patterns. A low correlation (Pearson's R = 0.11) was observed between the coverage of SSRP1 (ChIP-seq) and the gene fold change (RNA-seq) of those genes in the *Ssrp1* KD (**Fig. 2D**), indicating that FACT binding alone is not a predictor for gene expression changes. Taking these findings together, FACT can work directly as an enhancer or repressor of transcription in mES cells.

Given the high correlation of FACT with H3K4me3 (Fig. 1A) and to understand how the transcriptional changes might be linked to differences in recruitment of transcriptional regulators we performed an IP for H3K4me3 followed by Mass spectrometry both in Control and SSRP1-depleted ES cells (**Supplementary Fig. 3A, Supplementary Table 5**). We observed an increased binding of Oct4 and Sox2 to H3K4me3 in the *Ssrp1* KD state, in line with the observation that FACT depletion impacts developmental processes. Interestingly, we observed reduced binding of many splicing factors on H3K4me3 in the absence of FACT (**Supplementary Fig. 3A**). Differential splicing analysis between Control and *Ssrp1* KD conditions confirmed in total 356 Exon skipping/inclusion and 97 Intronic retention events following FACT depletion (**Supplementary Fig. 3B,C**). Interestingly, a fraction of the differential gene isoforms generated in the *Ssrp1* KD group is over-represented in limbic system and dendrite development pathways (**Supplementary Fig. 3D**), suggesting that genes involved in neuronal development might be influenced by FACT.

Depletion of FACT induces very specific changes in chromatin accessibility

Since FACT is responsible for the remodelling of nucleosomes in front of RNA polymerase and the re-establishment of nucleosome integrity in its wake (Formosa, 2012), we speculated whether some of the observed transcriptional alterations could be connected to changes in nucleosome occupancy upon depletion of FACT. Mononucleosome-sized DNA fragments upon treatment with MNase (135-170 bp) were purified from Control and *Ssrp1*-depleted conditions, and sequenced (**Supplementary Fig. 4A,B**). Nucleosome occupancy was plotted for four different gene classes according to the presence of SSRP1 in the control group (ChIP-Seq) and their relative gene Fold Change (RNA-seq) in the *Ssrp1* KD state. Overall, we observed little changes in nucleosome occupancy genome-wide (**Fig. 3A**). Genes that are down-regulated in the *Ssrp1* KD ("Down-regulated" class) and bound by FACT exhibit a global mononucleosomal shift by a few nucleotides right after the +1 nucleosome. Up-regulated genes showed a loss of nucleosome occupancy in the gene body area regardless of FACT-bound status (Non-SSRP1 and SSRP1 targets) (**Fig. 3A,B**), potentially reflecting the higher transcription rate through these genes. However, specifically in up-regulated genes bound by FACT ("Up-regulated" class), we observed a significant loss of nucleosomes upstream of the TSS (**Fig. 3A,B**). This difference in nucleosome occupancy at the promoter region is highly reproducible among the different replicates (**Supplementary Fig. 4C**). Splitting the up-regulated genes by the amount of H3K4me3 levels (k-means clustering) as a proxy for gene-expression levels, also revealed that the loss of nucleosomes at the promoter is more profound over the promoters of lowly expressed/ repressed genes (Control state) (**Supplementary Fig. 4C**). The observed nucleosome depleted regions (NDRs) were different between up- and down-regulated genes. Such architectural differences have been previously attributed to different levels of GC frequency. Indeed, GC frequency over SSRP1 targets was higher and broader in the "Down-

regulated" class corroborating a more open chromatin state (Fenouil et al., 2012) (**Fig. 3A & Supplementary Fig. 4D).**

To confirm this difference in chromatin accessibility using an additional approach, we performed ATAC-seq in Control and *Ssrp1*-depleted ES cells (**Fig. 3C**). In line with the observations of the MNase-seq experiments, we observed a statistically significant increase ($P < 10^{-10}$) in chromatin accessibility in the absence of FACT upstream of the promoter region of FACT-bound, up-regulated genes (**Fig. 3C,D,E**). In combination with the RNA-seq data, this reduction in nucleosome occupancy (and subsequently increase in chromatin accessibility) at the TSS suggests that FACT might act as a repressor by enabling a more closed chromatin conformation state at promoter regions.

Gain in chromatin accessibility upon FACT depletion upstream of the TSS coincides with antisense transcription

Over the last decade, it has become apparent that promoters can drive expression of sense and antisense RNAs, with proximally paused RNA Pol II on both strands (Jonkers, Kwak, & Lis, 2014; Seila et al., 2008). *In vitro*, FACT has been demonstrated to facilitate transcription through chromatinized templates (Orphanides et al., 1999) and reduces pausing of the elongating polymerase when it encounters nucleosomes (Hsieh et al., 2013). In yeast, depletion of Spt16 leads to upregulation of antisense transcription from gene-internal cryptic promoters (Feng et al., 2016). Thus, in order to understand how the observed changes in chromatin accessibility would impact transcription initiation and to get more mechanistic insight into how FACT might dampen expression of genes in mES cells, we performed NET-seq (Mayer & Churchman, 2016) (**Supplementary Fig. 5A**), a method that allows quantitative, strand-specific and nucleotide resolution mapping of RNA Pol II.

Initially, we sought to determine whether nascent transcription positively correlates with mRNA levels. A higher correlation of nascent RNA – mRNA expression and a significantly higher slope ($P < 10^{-5}$) was observed over SSRP1-target regions in the Control state suggesting higher levels of Pol II pausing and mRNA levels in the presence of FACT (**Fig. 4A**). Nevertheless, in the *Ssrp1* KD state the SSRP1-bound regions maintained a higher slope, suggesting that pausing and elongation speed of RNA Pol II are not controlled entirely by FACT (**Supplementary Fig. 5B**). To confirm this, we measured the travelling ratio of RNA Pol II over down-regulated and up-regulated genes. Indeed, "up-regulated" SSRP1-bound genes show a lower travelling ratio overall, but no significant difference was observed among this group of genes following FACT depletion (Control to *Ssrp1* KD comparison), indicating that FACT is not involved into the release of Pol II towards successive elongation in mammalian cells (**Fig. 4B**).

Next, we assessed RNA Pol II pausing and directionality over up-regulated genes. NET-seq density plots identified that FACT targets displayed higher levels of promoter-proximal RNA Pol II in the sense strand, but not in the antisense strand, compared to SSRP1-unbound promoters (**Fig. 4C,D**). Upon knock-down of FACT, SSRP1 targets displayed an increase ($P < 10^{-6}$) in divergent transcription compared to the non-SSRP1 targets (**Fig. 4E**). This occurred precisely at locations where nucleosomes were depleted upon knock-down of FACT (**Fig. 4F**). No change in antisense transcription was observed for Down-regulated (**Supplementary Fig. 6A,B**) or Unchanged (**Supplementary Fig. 6C,D**) genes suggesting that the presence of FACT over a specific gene class (Up-regulated genes) decreases the rate of antisense transcription by maintaining higher nucleosome density upstream of the TSS.

A correlation between loss of nucleosomes upstream of the TSS, increase in antisense and sense transcription has recently been reported to occur in mammalian cells (Scruggs et al., 2015). Furthermore, this study showed that antisense transcription can lead to a more open chromatin structure enabling increased binding of transcription factors, which is favourable for

sense transcription. Taken together, our data suggest that the repressive function of FACT is linked to nucleosome deposition at the promoter and obstruction of anti-sense transcription.

ES cells differentiate more efficiently into the neuronal lineage upon FACT depletion

Finally, we wanted to investigate, whether the transcriptional changes induced by depletion of FACT have physiological consequences. We tested this by differentiating mES cells into the neuronal lineage. The rationale for this approach stems from previous studies that pinpoint a specific role for FACT in neurons (Neumüller et al., 2011; Vied et al., 2014) as well as from the gene ontology enrichment for neuronal terms that we obtained from mRNA isoform analysis (**Supplementary Fig. 3E**). We induced differentiation of ES cells towards a neuronal lineage via embryoid body formation and treatment with retinoic acid (Bibel, Richter, Lacroix, & Barde, 2007). We created early stage Neural Precursor Cells (NPCs – 3 days into the differentiation process) and interrogated the whole transcriptome via RNA-seq. We identified that in early stage NPCs, expression of key neurogenesis markers (*Pax6*, *Nes*, *Tubb3*) increases whereas FACT mRNA levels and key pluripotency factors are unchanged but still maintained at a high level (**Fig. 5A**). A quarter of the up-regulated genes in ES cells after *Ssrp1* KD overlaps with the up-regulated genes instigated by neuronal differentiation ($P < 10^{-13}$, Fisher's exact test; **Fig. 5B**) and are over-represented in pathways involved in neuronal development. Similarly to our previous observations, β3-Tubulin (*Tubb3*) (SSRP1-bound gene), as an example for neurogenesis genes up-regulated upon FACT depletion, shows higher chromatin accessibility levels at the promoter region upon knock-down of FACT. This opening of the promoter is accompanied by an increase in anti-sense transcription (**Fig. 5C**).

We then depleted *Ssrp1* levels at the onset of neuronal differentiation and performed immunofluorescence for neurogenesis (β3-Tubulin) and dendritic (MAP2) markers at the same

time point as the RNA-seq experiment. *SSRP1 KD* caused a substantial increase in the expression of those markers, indicating that loss of FACT function primes ES cells for the neuronal lineage and enhances early neuronal differentiation (**Fig. 5D**).

Discussion

In this study we have addressed the role of the histone chaperone FACT in mouse ES cells. In contrast to the genomic profile identified for *S. cerevisiae* FACT, where the protein occupancy is depleted at the transcription start site and accumulates in the gene body (True et al., 2016), the genomic profile of mammalian FACT over active genes is reminiscent of a profile of the Ser5 phosphorylated form of RNA Pol II. This recruitment to the transcription start site might reflect binding of FACT to RNA Pol II. A similar profile for SSRP1 has been reported recently in HT1080 cells (Garcia et al., 2013).

In general, FACT depletion does not lead to gross alterations of the nucleosomal landscape as measured by MNase- and ATAC-seq. In particular, genes down-regulated upon FACT depletion only show a slight shift of nucleosomes, similar to what has been observed in yeast upon FACT inactivation (Feng et al., 2016). It is tempting to speculate that the reason for down-regulation lies in the originally described function of FACT to help passage of RNA Pol II through chromatin (Orphanides et al., 1999) and its depletion makes this process less efficient. FACT-bound genes that are up-regulated upon *Ssrp1* depletion show a significant alteration in nucleosomal occupancy around the transcriptional start site (TSS). FACT depletion leads to loss of nucleosomes and increased rates of bi-directional nascent transcription suggesting that these genes are usually dampened or repressed (in case of silent genes) by the maintenance of nucleosomes at these sites. The loss of nucleosomal occupancy upon depletion of FACT goes hand-in-hand with an increase in antisense transcription. Based on the data presented here, we

cannot determine if the loss of nucleosomes precedes upregulation of antisense transcription or vice versa. Also, it is not clear, whether this is driven by FACT alone, or in combination with RNA polymerase and/or chromatin remodellers. However, it is clear that this observed effect is very specific to SSRP1-bound genes, in which the histone chaperone operates as a repressor, suggesting that FACT is required to maintain the observed high level of nucleosome occupancy and to inhibit antisense transcription. One should note however, that this gene class shows low levels of antisense transcription (**Fig. 4C,D**). Therefore, one plausible model would be that FACT is required on these promoters to reinstate nucleosomes after initiation of antisense transcription. Depletion of FACT would lead to loss of this function and loss of nucleosomes, which in turn would drive higher levels of antisense transcription. It is of interest to note that FACT depletion in *S. cerevisiae* by using a thermosensitive allele of *spt16* also leads to upregulation of sense/antisense transcription. However, this occurs at cryptic promoters within the coding region of the gene due to a defect in re-establishing chromatin structure after passage of the elongating polymerase (Feng et al., 2016). Given the differences of FACT occupancy between mammals (this study and Garcia et al., 2013) and yeast (True et al., 2016), this might reflect evolutionary differences between mammalian and yeast FACT.

This scenario described for mammalian FACT would lead to a wider NFR and allow more efficient recruitment of TFs and RNA polymerase. In addition, the torque generated by two divergently elongating RNAPII molecules can create sufficient negative supercoiling density in the DNA between the two promoters, which is known to increase RNAPII transcription efficiency (Seila, Core, Lis, & Sharp, 2009). Taken together, we have shown that FACT can function both as an enhancer and a repressor of transcription. The repressive function of FACT correlates well with nucleosomal occupancy at the TSS and suppression of antisense transcription. In ES cells, genes repressed by FACT in this way encode for proteins involved in embryogenesis,

particularly in early neuronal differentiation, which is accelerated when FACT is depleted (**Fig. 6**).

FACT expression correlates with the differentiation state of the cell, being highest in undifferentiated and lowest in terminally differentiated cells. This cannot be simply explained by differences in proliferation rates as e.g. NIH-3T3 also exhibit low levels of FACT expression, but proliferate comparably to mouse ES cells. These observations suggest that FACT assists to maintain a chromatin/ transcription state that allows self-renewal. Indeed, depletion of FACT leads to an imbalance of the ES cell transcriptome. On the one hand, pro-proliferative genes are up-regulated and lowly expressed developmental factors are further down-regulated leading in the hyper-proliferation of ES cells. Moreover, the FACT-depleted gene signature has a large overlap with gene expression changes observed upon differentiation into the neuronal lineage. Interestingly, a comparison of expression patterns in the early developing mouse brain identified a set of only 13 genes, including *Ssrp1* with high correlation of expression in the proliferating cells of the VZ (Ventricular Zone) of the neocortex at early stages of development (Vied et al., 2014). This is a transient embryonic layer of tissue containing neural stem cells (Rakic, 2009) and a place for neurogenesis during development dependent on the Notch pathway (Rash, Lim, Breunig, & Vaccarino, 2011). Similarly to our study, hyperproliferation in a stem cell compartment upon FACT depletion has been observed before. Drosophila neuroblasts hyperproliferate upon deletion of SSRP1 suggesting that it is involved in the regulation of balancing neuroblast self-renewal and differentiation (Neumüller et al., 2011). A very recent report also highlights the role of FACT in assisting cell fate maintenance. Using a genetic screen in *C. elegans*, all FACT subunits were identified as barriers for cellular reprogramming of germ cells into the neuronal lineage (Kolundzic et al., 2017). Comparable to our results, the authors did not observe major chromatin architecture alterations, but observed larger colonies during reprogramming assays in the absence of FACT, indicative of higher proliferation rates. In

agreement with these reports, our data demonstrate that FACT-depleted ES cells differentiate much more efficiently into early neuronal precursors. Taken together, the data suggest a role for FACT activity during neuronal differentiation and the proper levels of FACT might assist in balancing proliferation speed and timing of differentiation processes.

Acknowledgments

We would like to thank Ilian Attanassov of the Max Planck Institute for Biology of Ageing Proteomics Core Facility for Mass Spectrometry Analysis and the FACS and Imaging Facility for help with microscopy. We are particularly grateful to Franziska Metge and Sven Templer (MPI-AGE Bioinformatics Core) for their assistance with coding script formatting. Sequencing was performed at the Max Planck Genome core centres in Berlin and Cologne and data analysis was done on servers of the GWDG, Göttingen and the MPI-AGE cluster. We thank Andy Bannister, Antonis Kirmizis and members of the Tessarz laboratory for discussion and comments on the manuscript. This work was funded by the Max Planck Society.

Author Contributions

C.M. and P.T. designed the study, C.M. performed all experiments and analysed data, C.M. and P.T. interpreted results and wrote the manuscript.

358 ***Conflict of Interest***

359 The authors declare no conflict of interest.

360

361 **Deposition of sequencing data**

362 Data have been deposited in Gene Expression Omnibus (GEO) under accession numbers GSE

363 90906 (ChIP-seq, RNA-seq, chrRNA-seq, MNase-seq, ATAC-seq, and NET-seq).

364

365

366

367

368

369

370

371

372

373

374

375

376

Materials and Methods

Cell culture. The E14 cell line (mESCs) was cultured at 37 °C, 7.5% CO_2, on 0.1% gelatin coated plates, in DMEM + GlutaMax™ (Gibco) with 15% fetal bovine serum (Gibco), MEM non-essential amino acids (Gibco), penicillin/streptomycin (Gibco), 550 µM 2-mercaptoethanol (Gibco), and 10 ng/ml of leukaemia inhibitory factor (LIF) (eBioscience). HEK293T, N2a, MEFs, NIH3T3, and B16 cell lines were cultured at 37 °C, 5% CO_2 in DMEM + GlutaMax™ (Gibco) with 10% fetal bovine serum (Gibco), and penicillin/streptomycin (Gibco). Early Neuronal Precursor Cells (NPCs) were generated as previously described (Bibel et al., 2007). Briefly, embryoid bodies were created with the hanging drop technique and were further treated with 1 µM retinoic acid (RA) for 4 days. RA-treated embryoid bodies were trypsinised and cultured in DMEM + GlutaMax™ (Gibco) with 15% fetal bovine serum without LIF for 3 days.

Depletion of SSRP1 from mESCs via shRNA and RNA preparation. E14 were transfected with lentiviral vectors containing either a scramble Control or *Ssrp1* shRNAs (MISSION® shRNA, Sigma) with the following sequences:

Scramble Control	CCGGGCGCGATAGCGCTAATAATTTCTCGAGAAATTATTAGCGCTATCGCGCTTTTT
shRNA 1 (*Ssrp1*)	CCGGCCTACCTTTCTACACCTGCATCTCGAGATGCAGGTGTAGAAAGGTAGGTTTTTG
shRNA 2 (*Ssrp1*)	CCGGGCGTACATGCTGTGGCTTAATCTCGAGATTAAGCCACAGCATGTACGCTTTTTG
shRNA 3 (*Ssrp1*)	CCGGGCAGAGGAGTTTGACAGCAATCTCGAGATTGCTGTCAAACTCCTCTGCTTTTTG
shRNA 4 (*Ssrp1*)	CCGGCCGTCAGGGTATCATCTTTAACTCGAGTTAAAGATGATACCCTGACGGTTTTTG

A combination of two different *Ssrp1* shRNAs was used (1&2, 3&4) at a time and depletion was quantified via western blotting using a monoclonal anti-Ssrp1 antibody (Biolegends). Anti-alpha

Tubulin was used as a reference control. The 1&2 combination was used for subsequent experiments as it yielded higher depletion of SSRP1 levels (Supplementary Figure 2a,b). Fourty-eight hours (48h) after transfection, puromycin (2 µg /ml) selection was applied for an additional 24h period, before cell collection and RNA preparation. Total RNA was obtained via phenol-chloroform extraction (QIAzol Lysis Reagent – QIAGEN) followed by purification via Quick-RNA™ MicroPrep (Zymo Research). Library preparation and ribosomal depletion were performed via the NEBNext Directional RNA Ultra Kit (NEB) and the RiboZero Kit (Illumina) according to the manufacturer's instructions, respectively. Four different biological replicates (Control or SSRP1-depleted mESCs) were prepared and processed for transcriptome analysis.

MTT proliferation assay. 48h after transfection, different cell densities ($3x10^4$, $2x10^4$, $1x10^4$) were seeded on 96-well plates (Sarstedt) along with puromycin (2 µg /ml). Twenty-four hours later, the CellTiter 96® Non-Radioactive Cell Proliferation Assay kit (Promega) was used according to the manufacturer's instructions in order to assess the rate of cell proliferation between the two conditions (Control, *Ssrp1* KD). Statistical analysis was performed using a two-tailed *t-test*.

Transcriptome analysis in SSRP1-depleted mESCs. Sequenced reads were aligned to the mm10 genome via STAR (v 2.4.1b) (Dobin et al., 2013). Gene and exon counts were obtained from featureCounts of the Rsubread package (R/Bioconductor). Only reads with CPM (counts per million) > 1 were kept for subsequent analysis. Counts were normalised using the internal TMM normalisation in edgeR (Robinson, McCarthy, & Smyth, 2009) and differential expression was performed using the limma (Ritchie et al., 2015) package. All of the RNA-seq data presented in this manuscript have been normalised to the total library size. Significant genes

418 with an absolute logFC > 1 and Adjusted P.Value < 0.01 were considered as differentially
419 expressed (Supplementary Table 1). The “Unchanged” gene class (n=2,179) was obtained from
420 genes with an Adjusted P.Value > 0.05. The diffSplice function implemented in limma was used
421 to identify differentially spliced exons between the two conditions (Supplementary Table 2).
422 Significant exons with an FDR < 0.001 were considered as differentially spliced. Retention
423 Introns were identified using the MISO (Katz, Wang, Airoldi, & Burge, 2010) (Mixture of
424 Isoforms) probabilistic framework (Supplementary Table 3).

425

426 **Retention intron events.** We verified the presence of retained introns in the *Ssrp1* KD by
427 randomly selecting ten intron retention events. The FastStart SYBR Green Master (Roche) was
428 used along with the following primers to amplify via PCR the retained intragenic regions;

Gene name	Forward primer	Reverse primer
Men1	ATTTCCCAGCAGGCTTCAGG	GGGATGACACGGTTGACAGC
Dvl1	CCTGGGACTACCTCCAGACA	CCTTCATGATGGATCCAATGTA
Map4k2	GCTGCAGTCAGTCCAGGAGG	TCCTGTTGCTTCAGAGTAGCC
Ctsa	GCAATACTCCGGCTACCTCA	TGGGGACTCGATATACAGCA
Pol2ri	CGAAATCGGGAGTGAGTAGC	GGTGGAAGAAGGAACGATCA
Wipf2	TAGAGATGAGCAGCGGAATC	TCGAGAGCTGGGGACTTGCA
Fuz	GACCCAGTGTGTGGACTGTG	GACAAAGGCTGTGCCAGTGG
Rfx5	CACCAGTTGCCCTCTCTGAA	CAATTCTCTTCCTCCCATGC
Fhod1	CACCAGGGAGCAGAGATGAT	CCATCAACATTGGCCTAACC
Tcirg1	AGCGACAGCACTCACTCCTT	CAACACCCCTGCTTCCAGGC

429

Amplified products were run on a 1.5% Agarose gel and visualized under UV. Band quantification was performed with ImageJ.

Chromatin Immunoprecipitation (ChIP) of FACT subunits. ChIP was performed in ~20 million ES cells, per assay, as described previously (Tessarz et al., 2014) with a few modifications. Briefly, cells were crosslinked with 1% formaldehyde for 20 min followed by quenching for 5 min with the addition of glycine to a final concentration of 0.125 M. After washing with PBS buffer, cells were collected and lysed in Cell Lysis buffer (5 mM Tris pH8.0, 85 mM KCl, 0.5% NP40) with proteinase inhibitors (10 µl/mL Phenylmethylsulfonyl fluoride (PMSF), 1 µl/mL Leupeptin and 1 µl/mL Pepstatin). Pellets were spun for 5 min at 5000 rpm at 4°C. Nuclei were lysed in Nuclei Lysis Buffer (1% SDS, 10 mM EDTA, 50 mM Tris HCl) and samples were sonicated for 12 min. Samples were centrifuged for 20 min at 13,000 rpm at 4°C and the supernatant was diluted in IP buffer (0.01% SDS, 1.1% Triton-X-100, 1.2 mM EDTA, 16.7 mM Tris HCl, 167 mM NaCl) and the appropriate antibody was added and left overnight with rotation at 4°C. Anti-Ssrp1 and anti-Supt16 antibodies were purchased from Biolegends (#609702) and Cell Signalling (#12191), respectively. Anti-AP-2γ (Tfap2c) antibody was purchased from Santa Cruz (#sc-12762). Two biological replicates were prepared for each FACT subunit, using independent cell cultures and chromatin precipitations. Protein A/G Dynabeads (Invitrogen) were added for 1h and after extensive washes, samples were eluted in Elution Buffer (1% SDS, 0.1 M $NaHCO_3$). 20 µL of 5 M NaCl were added and samples were reverse-crosslinked at 65°C for 4h. Following phenol-chloroform extraction and ethanol precipitation, DNA was incubated at 37°C for 4h with RNAse (Sigma).

ChIP-seq library preparation, sequencing, and peak-calling. Approximately 10-20 ng of ChIP material was used for library preparation. End-repair and adaptor ligation was prepared as described previously with a few modifications (Tessarz et al., 2014). Double sided size selections (~200 – 650bp) were performed using the MagSI-NGS Dynabeads (MagnaMedics,

#MD61021) according to the manufacturer's instructions. Purified adapter-ligated ChIP material was run on a high sensitivity DNA chip on a 2200 TapeStation (Agilent) to assess size distribution and adaptor contamination.

Samples were single-end deep-sequenced and reads were aligned to the mm10 genome using Bowtie2 (v 2.2.6)(Langmead & Salzberg, 2012). Peak-calling was performed using PePr (v 1.1) (Zhang, Lin, Johnson, Rozek, & Sartor, 2014) with peaks displaying an FDR < 10^{-5} considered as statistically significant (Supplementary Table 4). Peak annotation was performed via the chipenrich (Welch et al., 2014) R package with the following parameters (locusdef = "nearest_gene", method = "broadenrich").

ChIP-seq normalisation and metagene analysis. All the ChIP-seq BAM files were converted to bigwig (10 bp bin) and normalised to x1 sequencing depth using Deeptools (v 2.4) (Ramirez et al., 2016). Blacklisted mm9 co-ordinates were converted to mm10 using the LiftOver tool from UCSC and were further removed from the analysis. Average binding profiles were visualised using R (v 3.3.0). Heatmaps were generated via Deeptools. For the average profiles in **Supplementary Fig. 1C,D**, RPKM values from Control ES RNA-seq data were divided into four different quantiles and the average profile for each FACT subunit was generated for each quantile. The Pearson's correlation plot in **Figure 1A** was generated using all unique annotated mm10 RefSeq genes (*n* = 11,305) from UCSC (blacklisted regions were removed).

MNase-seq following SSRP1 depletion in mESCs. ES cells were cultured and transfected with shRNA vectors as described above. Biological replicates were obtained from two independent transfection experiments for each shRNA vector. Briefly, ~5 million cells were crosslinked with 1% formaldehyde for 20 min followed by quenching for 5 min with the addition

of glycine to a final concentration of 0.125 M. After washing with PBS buffer, cells were collected and lysed in Cell Lysis buffer (5 mM Tris pH8.0, 85 mM KCl, 0.5% NP40) with proteinase inhibitors (10 µl/mL Phenylmethylsulfonyl fluoride (PMSF), 1 µl/mL Leupeptin and 1 µl/mL Pepstatin). Nuclei were gathered by centrifugation (5000 rpm for 2 min) and were treated with 10 Kunitz Units/10^6 cells of micrococcal nuclease (NEB, #M0247S) for 5 min at 37°C in 40 µl of Micrococcal Nuclease Buffer (NEB, #M0247S). The reaction was stopped with the addition of 60 µl 50 mM EDTA, 25 µl 5 M NaCl, 15 µl 20% NP-40 and incubated on a rotator for 1h at room temperature to release soluble nucleosomes. Samples were centrifuged for 5 min at 10,000 g and supernatant was transferred to a new tube. This centrifugation step is important to obtain highly soluble nucleosomes and remove nucleosome-protein complexes, which can raise bias in subsequent data interpretation (Carone et al., 2014) (**Supplementary Fig. 7**). Samples were reverse-crosslinked by incubating overnight at 65°C with 0.5% SDS and proteinase K. Following phenol-chloroform extraction and ethanol precipitation, DNA was incubated at 37°C for 4h with RNAse (Sigma). All samples were run in a 2% agarose gel and fragments <200 bp were extracted and purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel) according to the manufacturer's instructions.

Purified DNA (500 ng) was used for library preparation as described above. The only difference was the PCR amplification step where we used the same conditions as mentioned in (Henikoff et al., 2011) but with only three amplification cycles. Libraries were verified using a 2200 TapeStation and were paired-end deep-sequenced (~250 million reads per sample). For quality checks and reproducibility, please refer to **Supplementary Fig. 7**.

MNase-seq normalisation and metagene analysis. All the MNase-seq BAM files were converted to bigwig, binned (1 bp), smoothed (20-bp window), and normalised to x1 sequencing depth using Deeptools (v 2.4). Moreover, they were split into two different categories according

to fragment length; <80 bp Transcription factor (TF)-sized fragments and 135-170 bp mononucleosome fragments). Average nucleosome occupancy profiles were visualised using R (v 3.3.0). For the **Supplementary Fig. 7D,E** the mm10 annotated exon list for mononucleosomal profiling was obtained from UCSC.

ATAC-seq following SSRP1 depletion in mESCs. ES cells were cultured and transfected with shRNA vectors as described above. Biological replicates were obtained from two independent transfection experiments for each shRNA vector. ATAC-seq was performed on 50,000 cells as previously described (Buenrostro, Giresi, Zaba, Chang, & Greenleaf, 2013). All samples were PCR amplified for 9 cycles were paired-end sequenced on an Illumina Hi-Seq 2500 platform.

ATAC-seq normalisation and metagene analysis. Sequenced paired mates were mapped on mm10 genome build using Bowtie2 with the following parameters: –X 2000. Reads corresponding to nucleosome free regions were selected via a random forest approach using the "ATACseqQC" R package. All the ATAC-seq BAM files were converted to bigwig, binned (1 bp), and normalised to x1 sequencing depth using Deeptools (v 2.4). Duplicated reads were removed. Chromatin accessibility profiles were visualised using R (v 3.3.0).

Mass spectrometry sample preparation and analysis. Nuclei were isolated from ~5 million ES cells under hypotonic conditions and samples were incubated overnight at 4°C with an anti-H3K4me3 antibody (Active Motif, #39159) in the presence of low-salt Binding buffer (150 mM NaCl, 50 mM Tris-HCl pH 8.0, 1% NP-40), protease inhibitors, and Protein G Dynabeads (Invitrogen). The following day, after several rounds of bead washing with Binding Buffer, samples were incubated overnight at 37°C with Tris pH 8.8 and 300 ng Trypsin Gold (Promega). In total, four samples were prepared for each condition (Control, *Ssrp1* KD). For the full protein interactome of both FACT subunits, nuclei were extracted as descripted above, and anti-Ssrp1 and anti-Supt16 antibodies were used. Peptides were desalted using StageTips(Rappsilber,

Ishihama, & Mann, 2003) and dried. The peptides were resuspended in 0.1% formic acid and analyzed using liquid chromatography - mass spectrometry (LC-MS/MS).

LC-MS/MS analysis. For mass spectrometric analysis, peptides were separated online on a 25 cm 75 µm ID PicoFrit analytical column (New Objective) packed with 1.9 µm ReproSil-Pur media (Dr. Maisch) using an EASY-nLC 1000 (Thermo Fisher Scientific). The column was maintained at 50°C. Buffer A and B were 0.1% formic acid in water and 0.1% formic acid in acetonitrile respectively. Peptides were separated on a segmented gradient from 5% to 25% buffer B for 45 min, from 25% to 35% buffer B for 8 min, and from 35% to 45% buffer B for 4 min, at 200nl / min. Eluting peptides were analyzed on a QExactive HF mass spectrometer (Thermo Fisher Scientific). Peptide precursor mass to charge ratio (m/z) measurements (MS1) were carried out at 60000 resolution in the 300 to 1500 m/z range. The top ten most intense precursors with charge state from 2 to 7 only, were selected for HCD fragmentation using 27% collision energy. The m/z of the peptide fragments (MS2) were measured at 15000 resolution, using an AGC target of 1e6 and 80 ms maximum injection time. Upon fragmentation, precursors were put on an exclusion list for 45 seconds.

LC-MS/MS data analysis. The raw data were analysed with MaxQuant (Jurgen Cox & Mann, 2008) (v 1.5.2.8) using the integrated Andromeda search engine (Jürgen Cox et al., 2011). Fragmentation spectra were searched against the canonical and isoform sequences of the mouse reference proteome (proteome ID UP000000589, downloaded August 2015) from UniProt. The database was automatically complemented with sequences of contaminating proteins by MaxQuant. For the data analysis, methionine oxidation and protein N-terminal acetylation were set as variable modifications. The digestion parameters were set to “specific”

552 and “Trypsin/P,” allowing for cleavage after lysine and arginine, also when followed by proline.
553 The minimum number of peptides and razor peptides for protein identification was 1; the
554 minimum number of unique peptides was 0. Protein identification was performed at a peptide
555 spectrum matches and protein false discovery rate of 0.01. The “second peptide” option was on
556 in order to identify co-fragmented peptides. Successful identifications were transferred between
557 the different raw files using the “Match between runs” option, using a match time window of 0.7
558 min. Label-free quantification (LFQ) (Jurgen Cox, Hein, Luber, & Paron, 2014) was performed
559 using an LFQ minimum ratio count of 2.

560

561 **Identification of co-enriched proteins.** Analysis of the label-free quantification results was
562 done using the Perseus computation platform (Tyanova et al., 2016) (v 1.5.0.0) and R. For the
563 analysis, LFQ intensity values were loaded in Perseus and all identified proteins marked as
564 “Reverse”, “Only identified by site”, and “Potential contaminant” were removed. Upon log2
565 transformation of the LFQ intensity values, all proteins that contained less than four missing
566 values in one of the groups (control or *Ssrp1* KD) were removed. Missing values in the resulting
567 subset of proteins were imputed with a width of 0.3 and down shift of 1.8. Next, the imputed
568 LFQ intensities were loaded into R where a two side testing for enrichment was performed using
569 limma (Kammers, Cole, Tiengwe, & Ruczinski, 2015; Ritchie et al., 2015). Proteins with an
570 adjusted p-value of less than 0.05 were designated as significantly enriched in the control or
571 knockdown (H3K4me3 IP) (Supplementary Table 5). The complete list of differential protein
572 expression between Control and *Ssrp1* KD can be found in Supplementary Table 7.

573 **NET-seq library preparation.** ES cells were cultured and transfected with shRNA vectors as
574 described above. Biological replicates were obtained from two independent transfection
575 experiments for each shRNA vector. NET-seq libraries were prepared as previously described

(Mayer & Churchman, 2016) with a few modifications. Briefly, chromatin associated nascent RNA was extracted through cell fractionation in the presence of α-amanitin, protease and RNAase inhibitors. > 90% recovery of ligated RNA and cDNA was achieved from 15 % TBE-Urea (Invitrogen) and 10% TBE-Urea (Invitrogen), respectively, by adding RNA recovery buffer (Zymo Research, R1070-1-10) to the excised gel slices and further incubating at 70°C (1500 rpm) for 15 min. Gel slurry was transferred through a Zymo-Spin IV Column (Zymo Research, C1007-50) and further precipitated for subsequent library preparation steps. cDNA containing the 3' end sequences of a subset of mature and heavily sequenced snRNAs, snoRNAs, and rRNAs, were specifically depleted using biotinylated DNA oligos (Supplementary Table 6). Oligo-depleted circularised cDNA was amplified via PCR (5 cycles) and double stranded DNA was run on a 4% low melt agarose gel. The final NET-seq library running at ~150 bp was extracted and further purified using the ZymoClean Gel DNA recovery kit (Zymo Research). Sample purity and concentration was assessed in a 2200 TapeStation and further deep sequenced in a HiSeq 2500 Illumina Platform (~400 million reads per replicate).

NET-seq analysis. All the NET-seq fastq files were processed using custom Python scripts (https://github.com/BradnerLab/netseq) to remove PCR duplicates and reads arising from RT bias. Reads mapping exactly to the last nucleotide of each intron and exon (Splicing intermediates) were further removed from the analysis. The final NET-seq BAM files were converted to bigwig (1 bp bin), separated by strand, and normalized to x1 sequencing depth using Deeptools (v 2.4) with an "–Offset 1" in order to record the position of the 5' end of the sequencing read. NET-seq tags sharing the same or opposite orientation with the TSS were assigned as 'sense' and 'anti-sense' tags, respectively. Promoter-proximal regions were carefully selected for analysis to ensure that there is minimal contamination from transcription arising from other transcription units. Genes overlapping within a region of 2.5 kb upstream of the TSS were removed from the analysis. For the NET-seq metaplots, genes underwent several

rounds of k-means clustering in order to filter regions; in a 2kb window around the TSS, rows displaying very high Pol II occupancy within a <100 bp region were removed from the analysis as they represent non-annotated short non-coding RNAs. Average Pol II occupancy profiles were visualised using R (v 3.3.0). In **Figure 4B** the Proximal Promoter region was defined as -30 bp and +250 bp around the TSS. For **Figure 4A,B**, gene body coverage was retrieved by averaging all regions (FACT-bound and non-FACT-bound) +300 bp downstream of TSS (Transcription Start site) and -200 bp upstream of TES (Transcription End Site). Comparison of the two linear regressions was performed by calculating the z-score via

$$z = \frac{\beta 1 - \beta 2}{\sqrt{s_{\beta 1}^2 + s_{\beta 2}^2}}$$

where β and s_β represent the 'slope' and the 'standard error of the slope', respectively. *P value* was calculated from the respective confidence level yielded by the z score.

Immunofluorescence and confocal microscopy. Early Neuronal Precursor Cells (NPCs) were generated and *Ssrp1* levels were knocked-down as described above. Cells were fixed with 100% Ethanol for 10 min and processed for immunofluorescence. Permeabilisation and blocking was performed for 1 h at room temperature with 1% BSA and 0.1% NP-40 in PBS. Incubation with primary antibodies was carried at room temperature for 2 hours by using rabbit anti-β3-Tubulin (1:300; Cell Signaling) and mouse anti-MAP2 (1:300; Millipore.). After washing in blocking buffer, the secondary antibodies anti-rabbit and anti-mouse Alexa Fluor 568 (1:1,000; Life Technologies.) were applied for 2 h at room temperature. Slides were extensively washed in PBS and nuclei were counterstained with DAPI before mounting. Fluorescence

images were acquired using a laser-scanning confocal microscope (TCS SP5-X; Leica), equipped with a white light laser, a 405-diode UV laser, and a 40× objective lens.

Gene Ontology Analysis. All GO terms were retrieved from the metascape online platform (http://metascape.org/).

Accession numbers and references of publicly available datasets. H3K4me3, H3K27me3, Pol II S5ph, H3K4me1, H3K27Ac, CTCF (ENCODE Consortium – E14 cell line); Chd1, Chd2, (de Dieuleveult et al., 2016) : GSE64825; p53 (Li et al., 2012): GSE26360; Pol II S2ph (Brookes et al., 2016): GSM850470.

References

Bibel, M., Richter, J., Lacroix, E., & Barde, Y. (2007). Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nature Protocols*, *2*(5), 1034–1043. http://doi.org/10.1038/nprot.2007.147

Brookes, E., de Santiago, I., Hebenstreit, D., Morris, K. J., Carroll, T., Xie, S. Q., … Pombo, A. (2016). Polycomb Associates Genome-wide with a Specific RNA Polymerase II Variant, and Regulates Metabolic Genes in ESCs. *Cell Stem Cell*, *10*(2), 157–170. http://doi.org/10.1016/j.stem.2011.12.017

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., & Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*, *10*, 1213. Retrieved from http://dx.doi.org/10.1038/nmeth.2688

Carone, B. R., Hung, J. H., Hainer, S. J., Chou, M. Te, Carone, D. M., Weng, Z., … Rando, O. J. (2014). High-resolution mapping of chromatin packaging in mouse embryonic stem cells and sperm. *Developmental Cell*, *30*(1), 11–22. http://doi.org/10.1016/j.devcel.2014.05.024

Carvalho, S., Raposo, A. C., Martins, F. B., Grosso, A. R., Sridhara, S. C., Rino, J., … De Almeida, S. F. (2013). Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription. *Nucleic Acids Research*, *41*(5), 2881–2893. http://doi.org/10.1093/nar/gks1472

Cox, J., Hein, M. Y., Luber, C. a, & Paron, I. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & Cellular …*, *13*(9), 2513–2526. http://doi.org/10.1074/mcp.M113.031591

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotech*, *26*(12), 1367–1372. Retrieved from http://dx.doi.org/10.1038/nbt.1511

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V, & Mann, M. (2011). Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *Journal of Proteome Research*, *10*(4), 1794–1805. http://doi.org/10.1021/pr101065j

de Dieuleveult, M., Yen, K., Hmitou, I., Depaux, A., Boussouar, F., Dargham, D. B., … Gérard, M. (2016). Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. *Nature*. http://doi.org/10.1038/nature16505

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., … Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. http://doi.org/10.1093/bioinformatics/bts635

Feng, J., Gan, H., Eaton, M. L., Zhou, H., Li, S., Belsky, J. A., … Li, Q. (2016). Non-coding transcription is a driving force for nucleosome instability in spt16 mutant cells. *Molecular and Cellular Biology*, *36*(13), MCB.00152-16. http://doi.org/10.1128/MCB.00152-16

Fenouil, R., Cauchy, P., Koch, F., Descostes, N., Cabeza, J. Z., Innocenti, C., … Andrau, J. C. (2012). CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. *Genome Research*, *22*(12), 2399–2408. http://doi.org/10.1101/gr.138776.112

Formosa, T. (2012). The role of FACT in making and breaking nucleosomes. *Biochimica et Biophysica Acta*, *1819*(3–4), 247–55. http://doi.org/10.1016/j.bbagrm.2011.07.009

Formosa, T., Eriksson, P., Wittmeyer, J., Ginn, J., Yu, Y., & Stillman, D. J. (2001). Spt16–Pob3 and the HMG protein Nhp6 combine to form the nucleosome- binding factor SPN. *The*

688 *EMBO Journal*, *20*(13), 3506 LP-3517. Retrieved from
689 http://emboj.embopress.org/content/20/13/3506.abstract

690 Garcia, H., Fleyshman, D., Kolesnikova, K., Safina, A., Commane, M., Paszkiewicz, G., …
691 Gurova, K. (2011). Expression of FACT in mammalian tissues suggests its role in
692 maintaining of undifferentiated state of cells. *Oncotarget*, *2*(10), 783–96. http://doi.org/340
693 [pii]

694 Garcia, H., Miecznikowski, J. C., Safina, A., Commane, M., Ruusulehto, A., Kilpinen, S., …
695 Gurova, K. V. (2013). Facilitates chromatin transcription complex is an “accelerator” of
696 tumor transformation and potential marker and target of aggressive cancers. *Cell Reports*,
697 *4*(1), 159–73. http://doi.org/10.1016/j.celrep.2013.06.013

698 Hondele, M., Stuwe, T., Hassler, M., Halbach, F., Bowman, A., Zhang, E. T., … Ladurner, A. G.
699 (2013). Structural basis of histone H2A-H2B recognition by the essential chaperone FACT.
700 *Nature*, *499*(7456), 111–4. http://doi.org/10.1038/nature12242

701 Hossan, T., Nagarajan, S., Baumgart, S. J., Xie, W., Magallanes, R. T., Hernandez, C., …
702 Johnsen, S. A. (2016). Histone Chaperone SSRP1 is Essential for Wnt Signaling Pathway
703 Activity During Osteoblast Differentiation. *Stem Cells (Dayton, Ohio)*, *34*(5), 1369–76.
704 http://doi.org/10.1002/stem.2287

705 Hsieh, F.-K., Kulaeva, O. I., Patel, S. S., Dyer, P. N., Luger, K., Reinberg, D., & Studitsky, V. M.
706 (2013). Histone chaperone FACT action during transcription through chromatin by RNA
707 polymerase II. *Proceedings of the National Academy of Sciences*, *110*(19), 7654–7659.
708 http://doi.org/10.1073/pnas.1222198110

709 Jonkers, I., Kwak, H., & Lis, J. T. (2014). Promoter-proximal pausing of RNA polymerase II:
710 emerging roles in metazoans. *eLife*, *3*(e02407), 1–25.

711 http://doi.org/http://dx.doi.org/10.7554/eLife.02407

712 Kammers, K., Cole, R. N., Tiengwe, C., & Ruczinski, I. (2015). Detecting significant changes in
713 protein abundance. *EuPA Open Proteomics*, *7*, 11–19.
714 http://doi.org/10.1016/j.euprot.2015.02.002

715 Kaplan, C. D., Laprade, L., & Winston, F. (2003). Transcription Elongation Factors Repress
716 Transcription Initiation from Cryptic Sites. *Science*, *301*(5636), 1096 LP-1099. Retrieved
717 from http://science.sciencemag.org/content/301/5636/1096.abstract

718 Katz, Y., Wang, E. T., Airoldi, E. M., & Burge, C. B. (2010). Analysis and design of RNA
719 sequencing experiments for identifying isoform regulation. *Nature Methods*, *7*(12), 1009–
720 15. http://doi.org/10.1038/nmeth.1528

721 Kelley, D. E., Stokes, D. G., & Perry, R. P. (1999). CHD1 interacts with SSRP1 and depends on
722 both its chromodomain and its ATPase/helicase-like domain for proper association with
723 chromatin. *Chromosoma*, *108*(1), 10–25. http://doi.org/10.1007/s004120050347

724 Kemble, D. J., McCullough, L. L., Whitby, F. G., Formosa, T., & Hill, C. P. (2015). FACT
725 Disrupts Nucleosome Structure by Binding H2A-H2B with Conserved Peptide Motifs.
726 *Molecular Cell*, *60*(2), 294–306. http://doi.org/10.1016/j.molcel.2015.09.008

727 Kolundzic, E., Ofenbauer, A., Uyar, B., Sommermeier, A., Seelk, S., He, M., … Tursun, B.
728 (2017). FACT sets a barrier for cell fate reprogramming in C. elegans
729 and Human. *bioRxiv*. Retrieved from
730 http://biorxiv.org/content/early/2017/09/06/185116.abstract

731 Kwak, H., Fuda, N. J., Core, L. J., & Lis, J. T. (2013). Precise Maps of RNA Polymerase Reveal
732 How Promoters Direct Initiation and Pausing. *Science*, *339*(6122), 950 LP-953. Retrieved
733 from http://science.sciencemag.org/content/339/6122/950.abstract

734 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat*
735 *Methods*, *9*(4), 357–359. http://doi.org/10.1038/nmeth.1923

736 Li, M., He, Y., Dubois, W., Wu, X., Shi, J., & Huang, J. (2012). Distinct Regulatory Mechanisms
737 and Functions for p53-Activated and p53-Repressed DNA Damage Response Genes in
738 Embryonic Stem Cells. *Molecular Cell*, *46*(1), 30–42.
739 http://doi.org/10.1016/j.molcel.2012.01.020

740 Mayer, A., & Churchman, L. S. (2016). Genome-wide profiling of RNA polymerase transcription
741 at nucleotide resolution in human cells with native elongating transcript sequencing. *Nat*
742 *Protocols*, *11*(4), 813–833. http://doi.org/10.1038/nprot.2016.047

743 Mayer, A., Di Iulio, J., Maleri, S., Eser, U., Vierstra, J., Reynolds, A., … Churchman, L. S.
744 (2015). Native elongating transcript sequencing reveals human transcriptional activity at
745 nucleotide resolution. *Cell*, *161*(3), 541–544. http://doi.org/10.1016/j.cell.2015.03.010

746 Neumüller, R. A., Richter, C., Fischer, A., Novatchkova, M., Neumüller, K. G., & Knoblich, J. A.
747 (2011). Genome-wide analysis of self-renewal in Drosophila neural stem cells by
748 transgenic RNAi. *Cell Stem Cell*, *8*(5), 580–593. http://doi.org/10.1016/j.stem.2011.02.022

749 Orphanides, G., Wu, W. H., Lane, W. S., Hampsey, M., & Reinberg, D. (1999). The chromatin-
750 specific transcription elongation factor FACT comprises human SPT16 and SSRP1
751 proteins. *Nature*, *400*(6741), 284–8. http://doi.org/10.1038/22350

752 Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nat Rev*
753 *Neurosci*, *10*(10), 724–735. Retrieved from http://dx.doi.org/10.1038/nrn2719

754 Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., … Manke, T.
755 (2016). deepTools2: a next generation web server for deep-sequencing data analysis.
756 *Nucleic Acids Research*, *44*(April), 160–165. http://doi.org/10.1093/nar/gkw257

Rappsilber, J., Ishihama, Y., & Mann, M. (2003). Stop and Go Extraction Tips for Matrix-Assisted Laser Desorption/Ionization, Nanoelectrospray, and LC/MS Sample Pretreatment in Proteomics. *Analytical Chemistry*, *75*(3), 663–670. http://doi.org/10.1021/ac026117i

Rash, B. G., Lim, H. D., Breunig, J. J., & Vaccarino, F. M. (2011). FGF Signaling Expands Embryonic Cortical Surface Area by Regulating Notch-Dependent Neurogenesis. *The Journal of Neuroscience*, *31*(43), 15604 LP-15617. Retrieved from http://www.jneurosci.org/content/31/43/15604.abstract

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research*, *43*(7), e47. http://doi.org/10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2009). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, *26*(1), 139–140. http://doi.org/10.1093/bioinformatics/btp616

Scruggs, B. S., Gilchrist, D. A., Nechaev, S., Muse, G. W., Burkholder, A., Fargo, D. C., & Adelman, K. (2015). Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. *Molecular Cell*, *58*(6), 1101–1112. http://doi.org/10.1016/j.molcel.2015.04.006

Seila, A. C., Calabrese, J. M., Levine, S. S., Yeo, G. W., Rahl, P. B., Flynn, R. A., … Sharp, P. A. (2008). Divergent Transcription from Active Promoters. *Science*, *322*(5909), 1849 LP-1851. Retrieved from http://science.sciencemag.org/content/322/5909/1849.abstract

Seila, A. C., Core, L. J., Lis, J. T., & Sharp, P. A. (2009). Divergent transcription: A new feature of active promoters. *Cell Cycle*, *8*(16), 2557–2564. http://doi.org/10.4161/cc.8.16.9305

Tessarz, P., Santos-Rosa, H., Robson, S. C., Sylvestersen, K. B., Nelson, C. J., Nielsen, M. L.,

& Kouzarides, T. (2014). Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature*, *505*(7484), 564–8. http://doi.org/10.1038/nature12819

True, J. D., Muldoon, J. J., Carver, M. N., Poorey, K., Shetty, S. J., Bekiranov, S., & Auble, D. T. (2016). The modifier of transcription 1 (Mot1) ATPase and Spt16 histone chaperone co-regulate transcription through preinitiation complex assembly and nucleosome organization. *Journal of Biological Chemistry*, *291*(29), 15307–15319. http://doi.org/10.1074/jbc.M116.735134

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., … Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Meth*, *13*(9), 731–740. Retrieved from http://dx.doi.org/10.1038/nmeth.3901

Vied, C. M., Freudenberg, F., Wang, Y., Raposo, A. a S. F., Feng, D., & Nowakowski, R. S. (2014). A multi-resource data integration approach: identification of candidate genes regulating cell proliferation during neocortical development. *Frontiers in Neuroscience*, *8*(August), 257. http://doi.org/10.3389/fnins.2014.00257

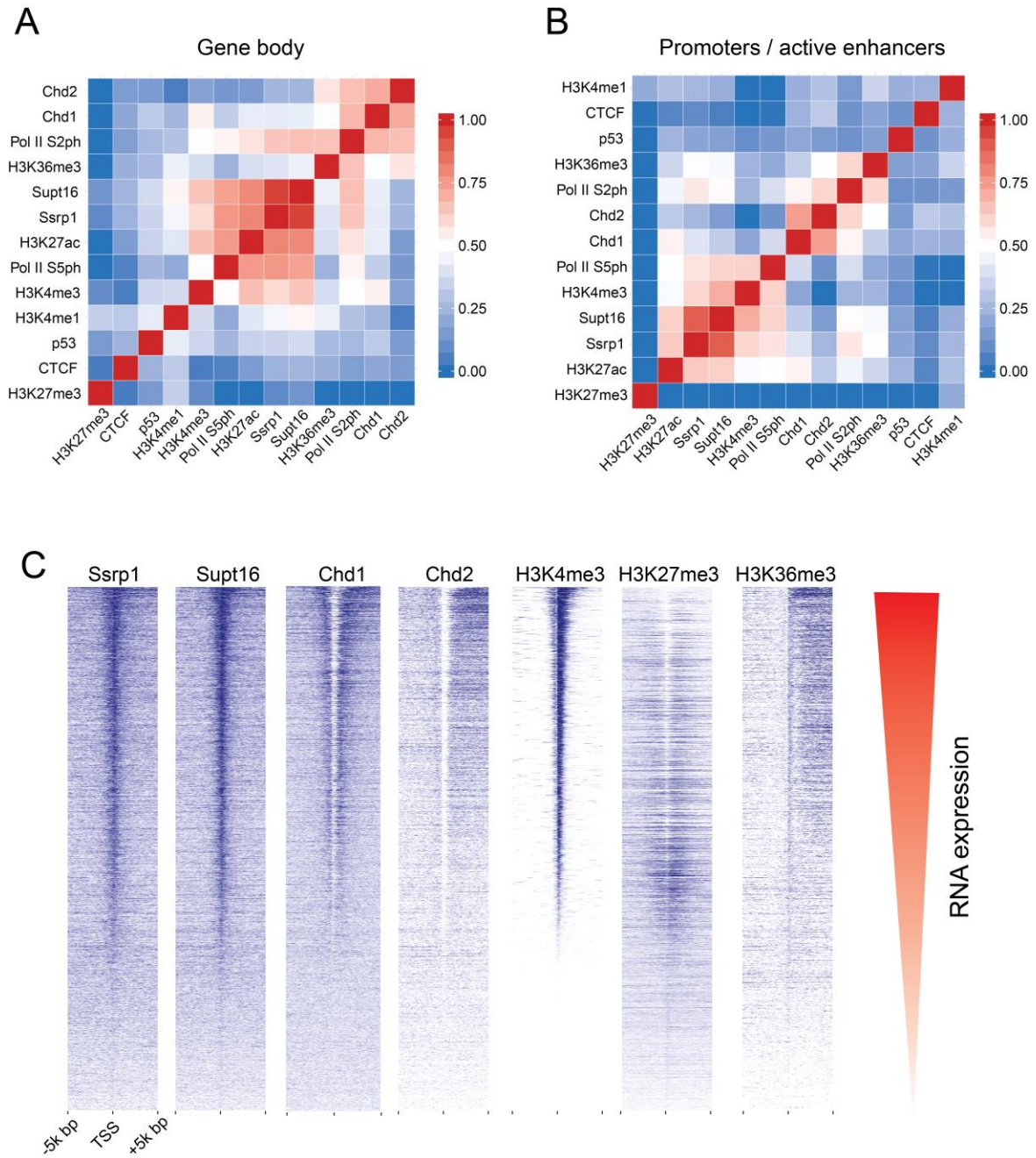
Wamstad, J. A., Alexander, J. M., Truty, R. M., Shrikumar, A., Li, F., Eilertson, K. E., … Bruneau, B. G. (2012). Dynamic and Coordinated Epigenetic Regulation of Developmental Transitions in the Cardiac Lineage. *Cell*, *151*(1), 206–220. http://doi.org/10.1016/j.cell.2012.07.035

Welch, R. P., Lee, C., Imbriano, P. M., Patil, S., Weymouth, T. E., Smith, R. A., … Sartor, M. A. (2014). ChIP-Enrich: Gene set enrichment testing for ChIP-seq data. *Nucleic Acids Research*, *42*(13), 1–13. http://doi.org/10.1093/nar/gku463

Winkler, D. D., & Luger, K. (2011). The Histone Chaperone FACT: Structural Insights and Mechanisms for Nucleosome Reorganization. *Journal of Biological Chemistry*, *286*(21),

18369–18374. http://doi.org/10.1074/jbc.R110.180778

Zhang, Y., Lin, Y. H., Johnson, T. D., Rozek, L. S., & Sartor, M. A. (2014). PePr: A peak-calling prioritization pipeline to identify consistent or differential peaks from replicated ChIP-Seq data. *Bioinformatics*, *30*(18), 2568–2575. http://doi.org/10.1093/bioinformatics/btu372



809 Figure 1: **Correlated occupancies across FACT-bound regions**. **A**, Heatmap representing
810 Pearson's correlation between FACT subunits (SSRP1, SUPT16), and other factors over the
811 gene body area of all uniquely annotated protein-coding genes (*n* = 11,305). **B**, Same as (**A)** but
812 for promoter/active enhancer regions (*n* = 19,461) characterised by high H3K27ac and/or Pol II
813 density. **c,** Distribution of FACT and other factors (ChIP-seq tags indicated in blue) over the TSS
814 of 11,305 unique RefSeq genes, sorted by H3K4me3 levels. Coinciding RNA expression levels
815 are shown in red.

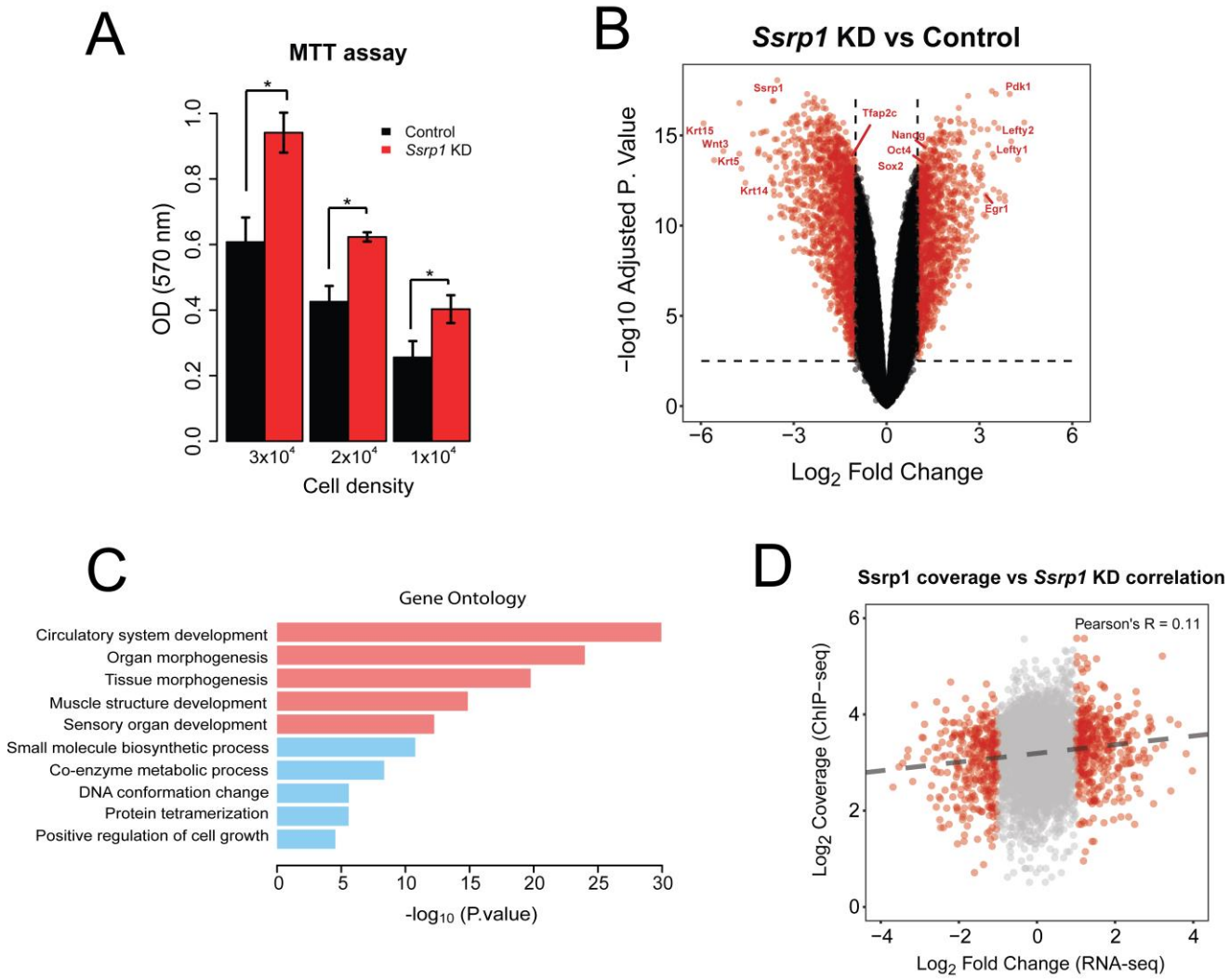


Figure 2: **Regulation of gene expression by FACT.** **A**, MTT assay assessing cell metabolic activity in mESCs at different cell densities following depletion of FACT levels. Values are mean and SE of three independent transfection experiments are displayed. Significance was calculated via a two-tailed *t-test* ($P < 0.05$). **B**, Volcano plot of differentially expressed genes between the Control and KD group. Values with logFC > 1 or logFC <-1 and Adjusted P.value < 0.01 are highlighted in red. **C**, Gene ontology analysis of all differentially expressed genes (Red: pathways for down-regulated genes, Blue: pathways for up-regulated genes). **D**, Scatterplot of log (SSRP1 coverage) (ChIP-seq) over logFC (RNA-seq).

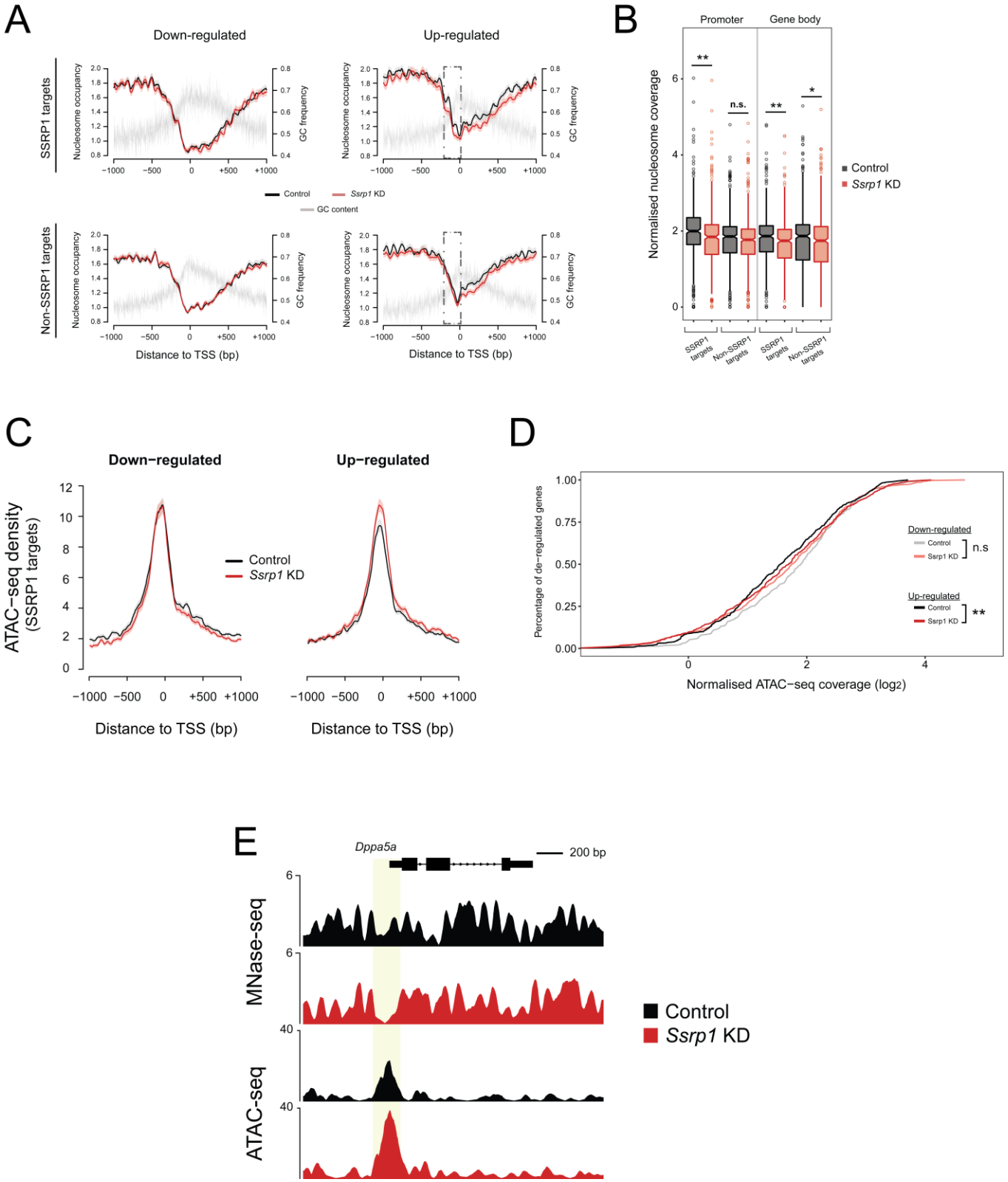


Figure 3: **Regulation of gene expression by FACT through chromatin accessibility. A**, Nucleosome occupancy of all de-regulated genes. Datasets are split by their FACT occupancy status (SSRP1 and Non-SSRP1 targets) and their relative transcriptional direction ("Down-regulated", "Up-regulated") following SSRP1 depletion. Solid lines indicate the mean values, whereas the shading represents the SE of the mean. **B.** Boxplots measuring the nucleosome occupancy (log_2) over promoters and gene body area of Up-regulated genes (** $p < 0.001$, * $p < 0.05$, n.s. =not significant). The assessed promoter region is shown in dashed boxes indicated in (**A**). Significance was calculated using the Welch Two-Sample t-test. **C.** Metaplot of open chromatin assessed by ATAC-seq among Down-regulated and Up-regulated genes both in Control and *Ssrp1* KD conditions. **D,** Cumulative distribution of ATAC-seq density for genes and conditions displayed in (**C**). Significance was calculated using the Welch Two-Sample t-test (** $p < 10^{-9}$, n.s. =not significant. **E,** Interrogation of nucleosome occupancy (MNase-seq) and chromatin accessibility (ATAC-seq) over the *Dppa5a* gene promoter for Control and *Ssrp1* KD conditions. Changes in nucleosome occupancy and chromatin accessibility are highlighted in yellow.

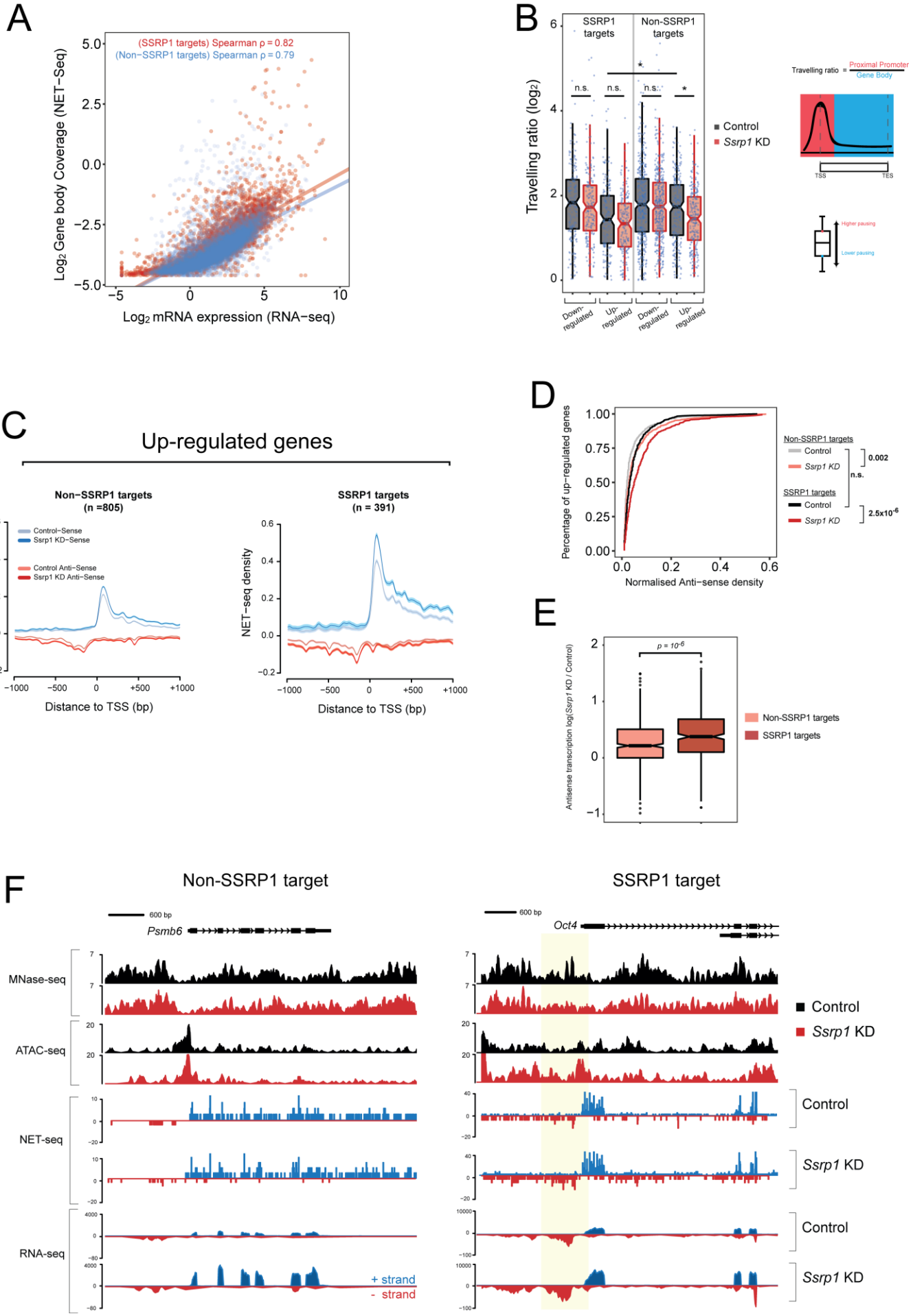
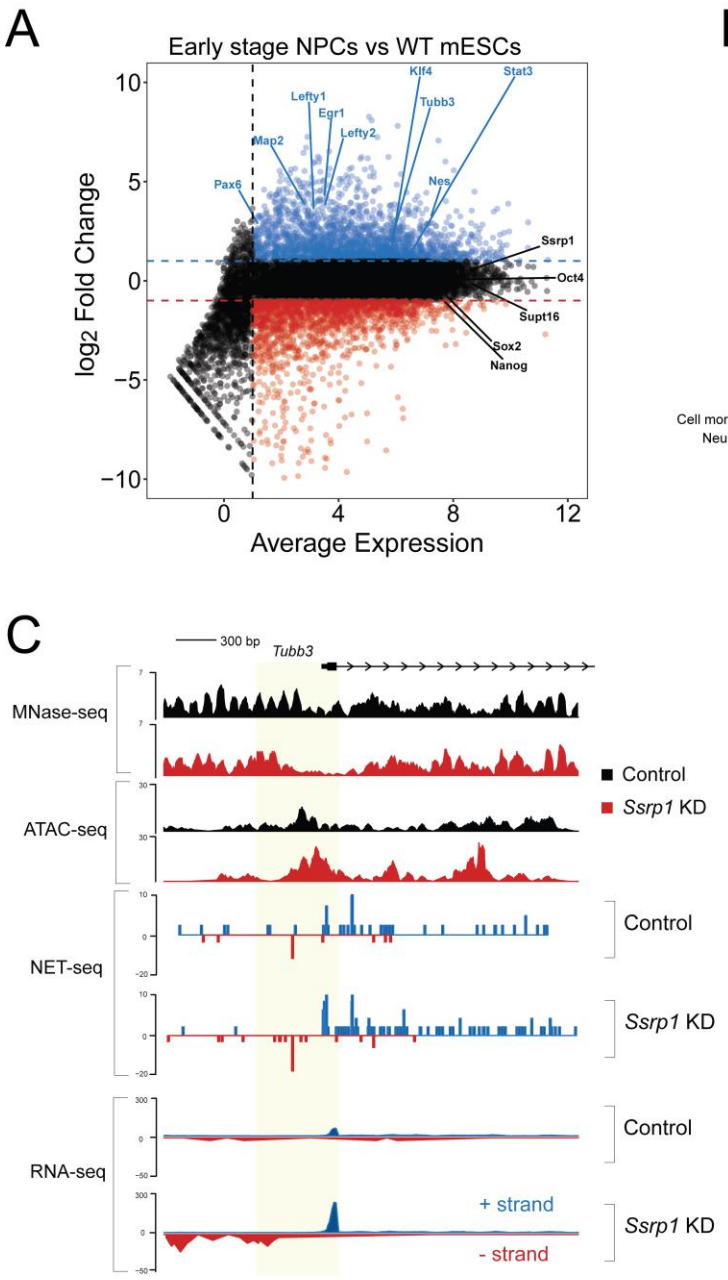


Figure 4: **Regulation of RNA Pol II directionality by FACT. A,** Scatterplots of log gene body coverage (NET-seq) versus log mRNA expression (RNA-seq) for SSRP1 (n=4,576) and Non-SSRP1 (n=8,844) target regions in the Control state (Z-score = 5.3, $P < 10^{-5}$). **B,** Measure of Pol II pause/release. Travelling ratio is defined as NET-seq density of proximal promoter versus gene body area. The log transformed travelling ratio for each gene class is displayed with boxplots. The Welch Two-sided t-test was used to calculate significance between Control and *Ssrp1* KD (* $p < 0.05$, n.s. =not significant). **C,** NET-seq density plots (Control and *Ssrp1* KD group) of Up-regulated genes split by FACT-bound status (Non-SSRP1 and SSRP1 targets). Solid lines indicate mean values, whereas the shading represents the 95% confidence interval. **D,** Cumulative distribution of anti-sense transcription (NET-seq) in a window 1000 bp upstream of the TSS. The Welch Two-sided t-test was used to calculate significance between Control and *Ssrp1* KD among Non-SSRP1 and SSRP1 targets. **E,** Boxplots assessing fold change (*Ssrp1* KD vs Control) in anti-sense transcription (NET-seq) in a window 1000 bp upstream of the TSS. The Welch Two-sided t-test was used to calculate significance between Non-SSRP1 and SSRP1 targets. **F**, Nucleosome occupancy (MNase-seq), open chromatin (ATAC-seq), and transcriptional activity (NET-seq/ RNA-seq) over an SSRP1 (*Oct4*) and non-SSRP1 (*Psmb6*) target gene between Control and *Ssrp1* KD conditions. Nucleosomal loss and increase in antisense transcription at the *Oct4* promoter is highlighted in yellow.



870

Figure 5: **FACT regulates neurogenesis through Pol II / nucleosome dynamics. A**, MA plot depicting differential expression in NPCs versus WT ES cells. Up-regulated genes are highlighted in blue whereas down-regulated genes are highlighted in red. **B**, Venn diagram showing the overlap of up-regulated genes between NPC vs mESCs and Control vs *Ssrp1* KD mESCs. **C**, Interrogation of nucleosome occupancy (MNase-seq), chromatin accessibility (ATAC-seq), and transcriptional activity (NET-seq/ RNA-seq) over the *Tubb3* gene promoter for Control and *Ssrp1* KD conditions. Changes in nucleosome occupancy, chromatin accessibility, and Pol II occupancy are highlighted in yellow. **D**, Immunofluorescence (IF) analysis of early stage NPCs following *Ssrp1* depletion. (Blue) DAPI, nuclei; (Green) β3-Tubulin (Tubb3), neurons; (Red) MAP2, dendrites.

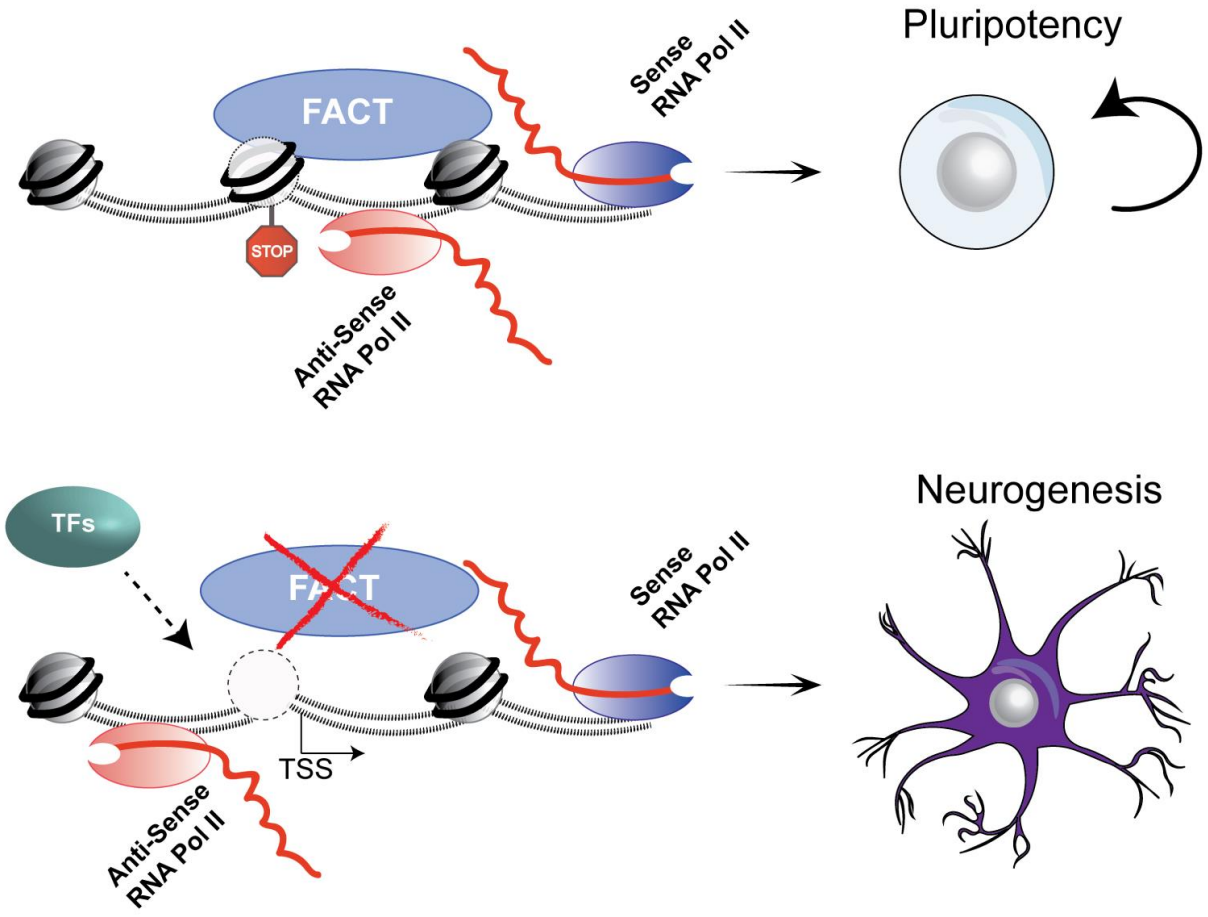
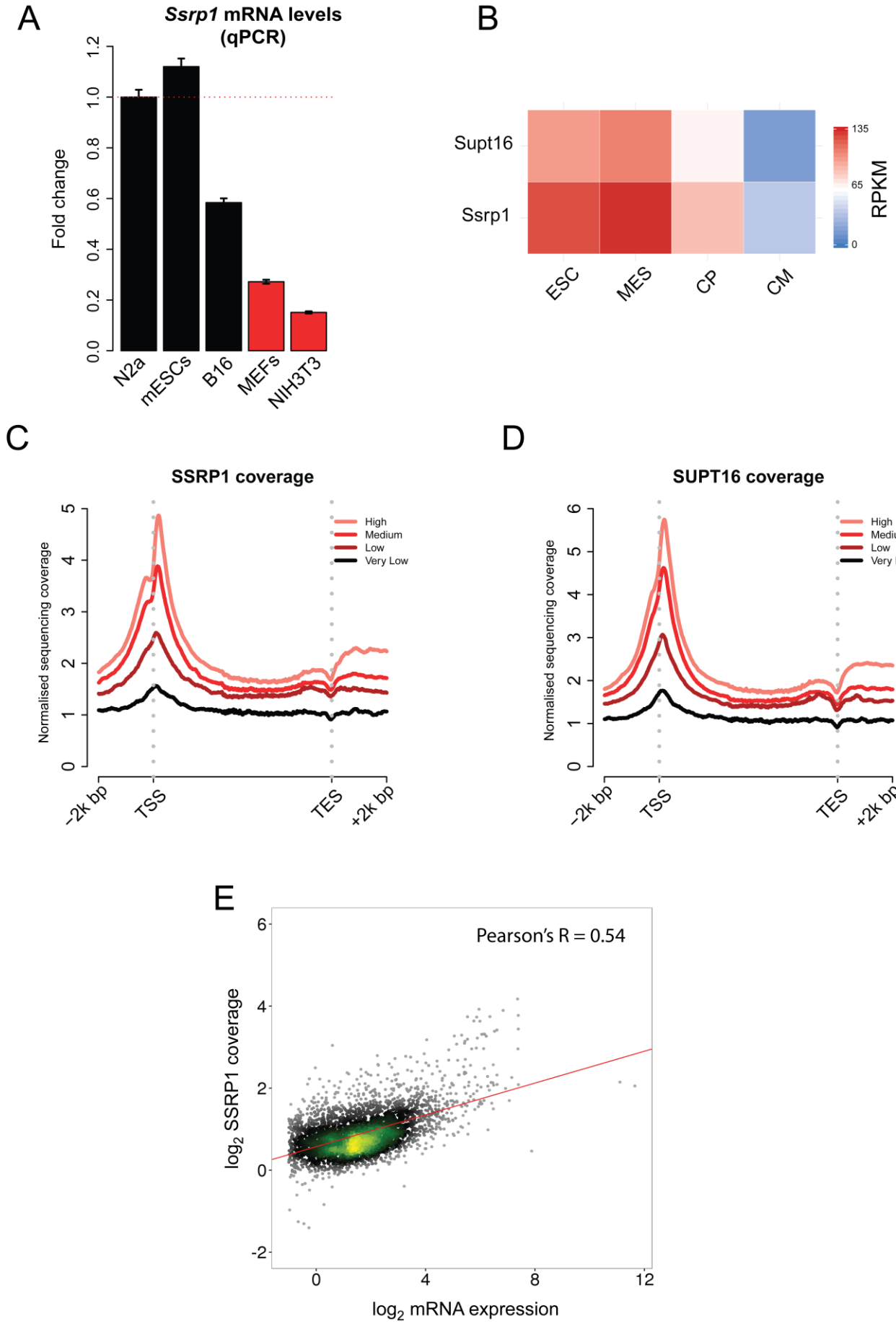


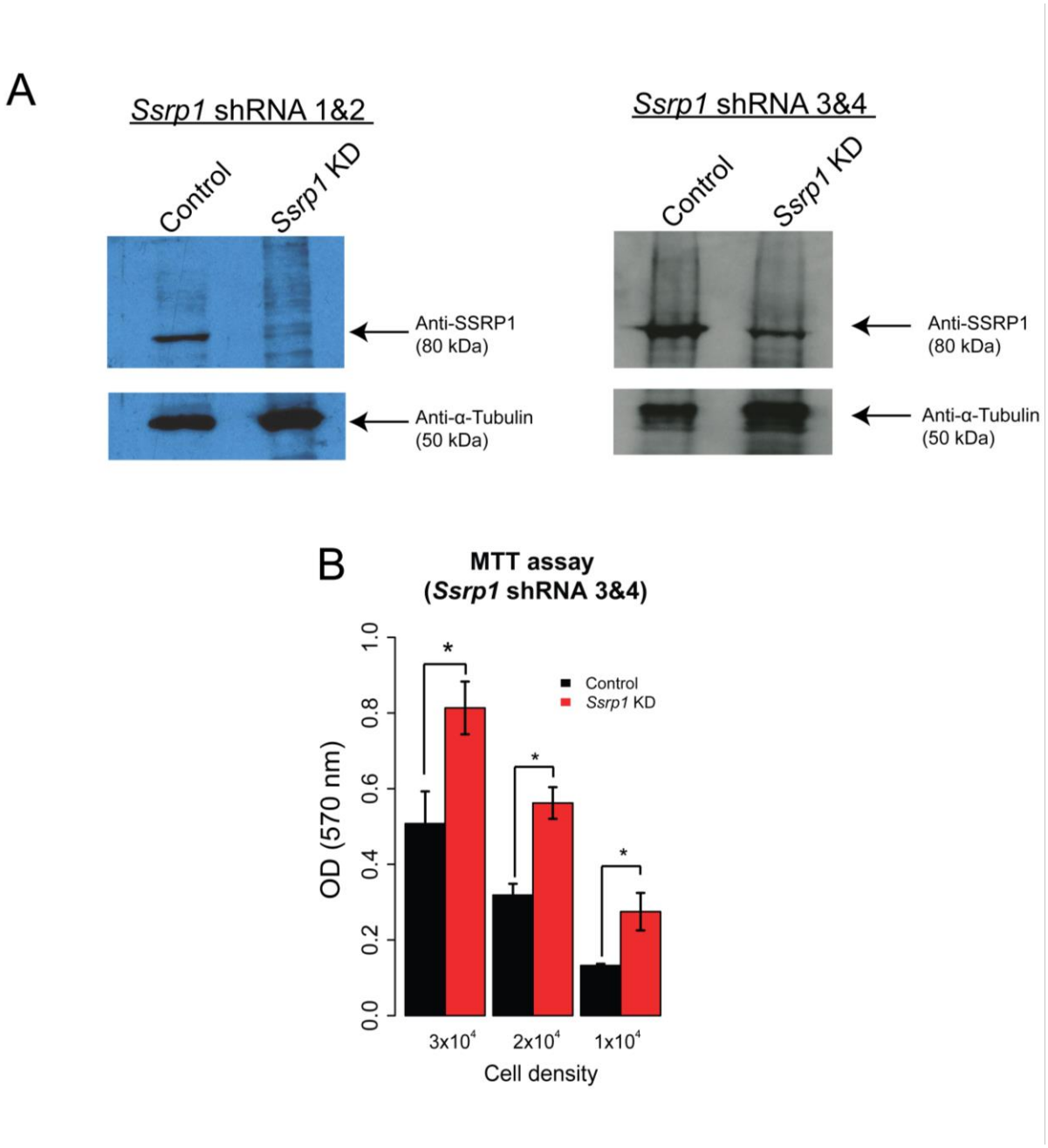
Figure 6: Model of the suppressive role of FACT in gene expression and maintenance of pluripotency. Upper; FACT places a nucleosomal barrier at the promoter region of genes involved in embryogenesis/ neurogenesis that hinders divergent travelling of Pol II and enables a closed chromatin conformation state. Lower; In the absence of FACT, the nucleosomal barrier is alleviated, thus allowing bi-directional travelling of Pol II, recruitment of TFs, increased gene expression, and ultimately, activation of neurogenesis cues.



892 **Supplementary Figure 1: FACT correlates with active gene expression. A**, qPCR of *Ssrp1*
893 levels among stem (mESCs), cancer (N2a, B16), and differentiated cell lines (MEFs, NIH3T3).
894 Mean and SE of three biological replicates are shown. Data were normalised to *Gapdh* mRNA
895 levels. **B**, Heatmap assessing the mRNA levels (RPKM) of FACT (*Ssrp1, Supt16*) at different
896 timepoints of differentiation of ES cells to cardiomyocytes (Wamstad et al., 2012) (ESC =
897 Embryonic Stem Cells, MES = Mesodermal cells, CP = Cardiac Precursors, CM =
898 cardiomyocytes). **C**, Distribution of SSRP1 relative to the TSS (± 2000 bp) and the TES (± 2000
899 bp) for four different gene classes ranked by level of RNA abundance (High, Medium, Low, Very
900 Low). **D**, Same as **(C)** but for SUPT16. **E**, Scatterplot of the $\log_2$ SSRP1 coverage versus $\log_2$
901 mRNA expression.

902

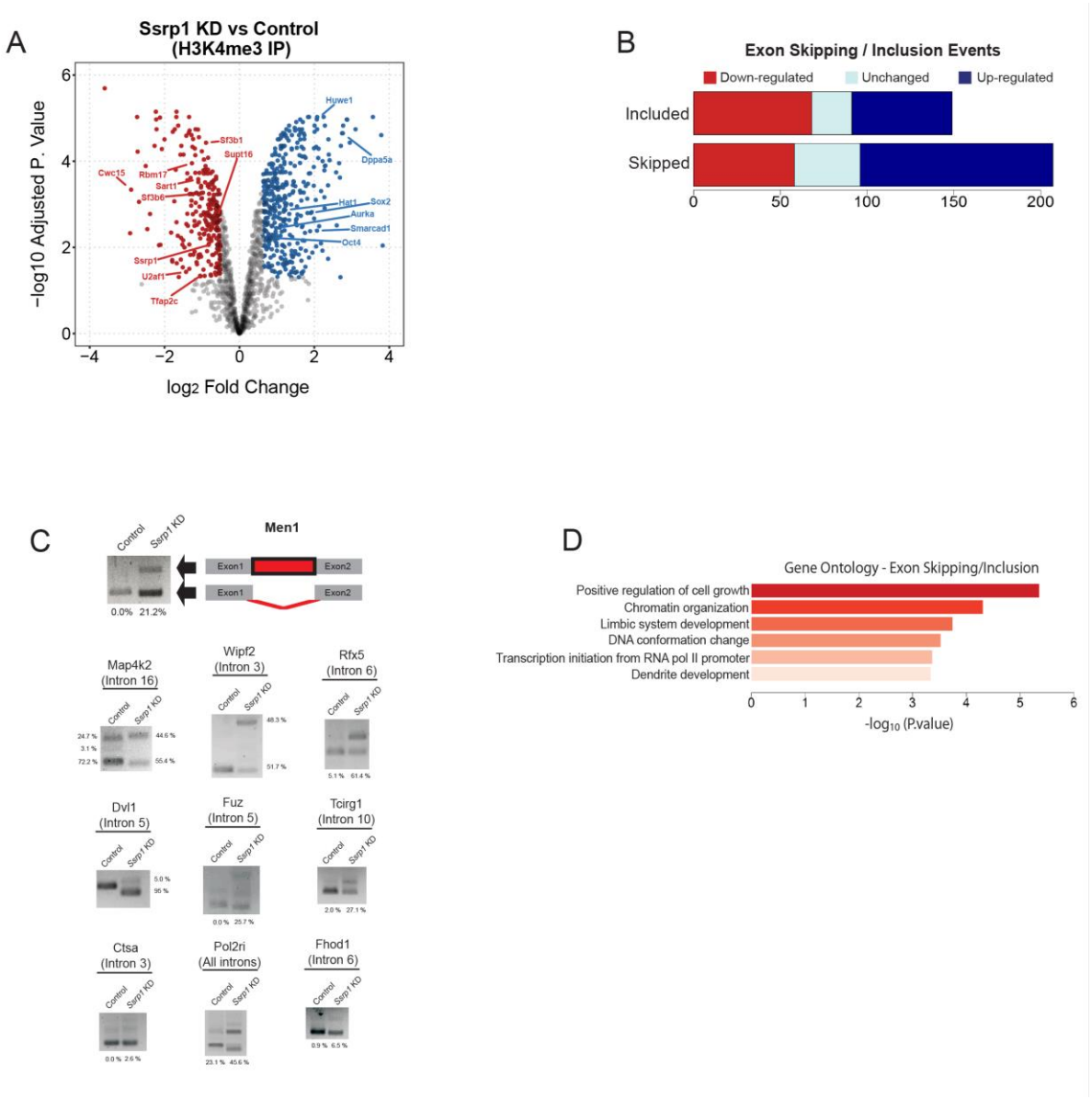
903



904

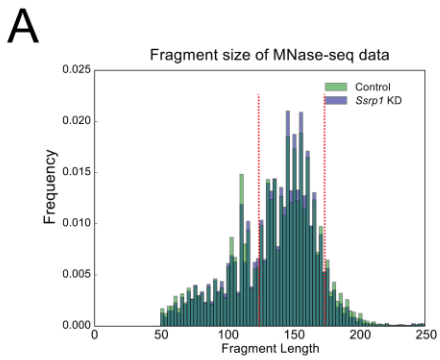
905 **Supplementary Figure 2: FACT depletion promotes increased proliferation rate in**
906 **mESCs. a**, Western blots after transfection of mESCs with different combinations of *Ssrp1*
907 shRNA vectors (shRNA 1&2 or shRNA 3&4). Anti-α-Tubulin was used as a reference. **b,** MTT
908 assay following transfection with *Ssrp1* shRNA 3&4 vectors. Values are mean and SE of three
909 independent transfection experiments are displayed. Significance was calculated via a two-
910 tailed *t-test* (*$P < 0.05$).

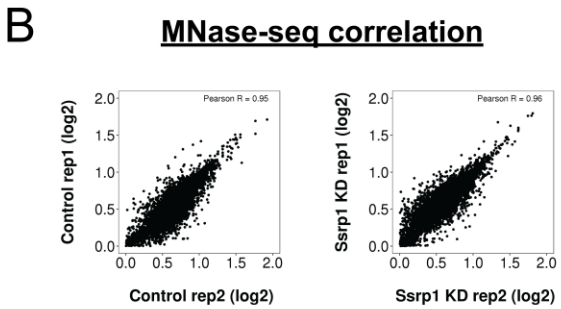
911

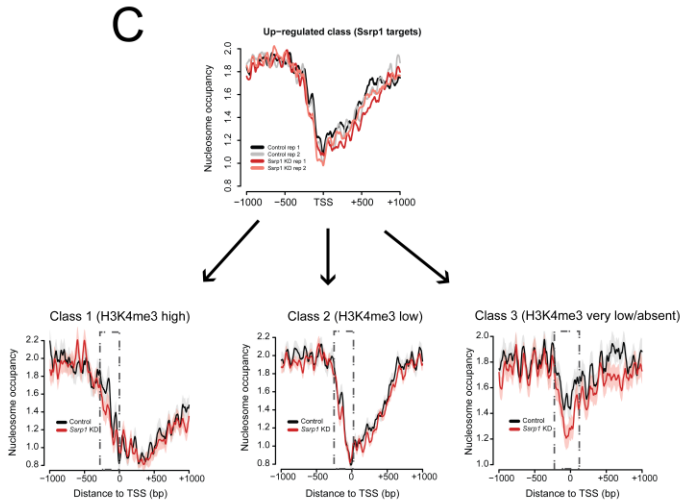


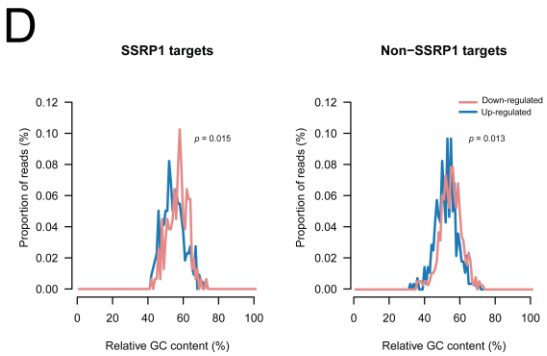
912

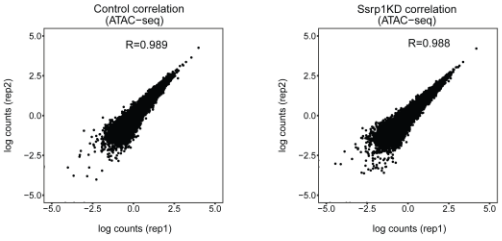
913 **Supplementary Figure 3: FACT facilitates alternative splicing of RNA transcripts. A,**
914 Volcano plot of depleted/enriched proteins at H3K4me3 following *Ssrp1* depletion. **B,** Average
915 distribution of SSRP1, SUPT16, and H3K4me3 aligned to the 5' Splice Site (5'SS) of all genes
916 grouped by first exon length. **C**, Barplots representing the number of included/skipped exons
917 categorized by their gene expression status (red: “Down-regulated”, cyan: “Unchanged”, blue:
918 “Up-regulated”). In total, we have identified 149 included and 207 skipped exon events in the
919 *Ssrp1* KD group. **D**, Graphical representation of an intronic retention event (*Men1*) in the KD
920 group. Also, analysis of intron inclusion events or isoform switches after FACT depletion.
921 Unspliced transcript percentage was measured according to band intensity. **E,** Gene ontology
922 analysis of transcripts (FACT-bound) that display alternative exon usage between the two
923 conditions











924

925

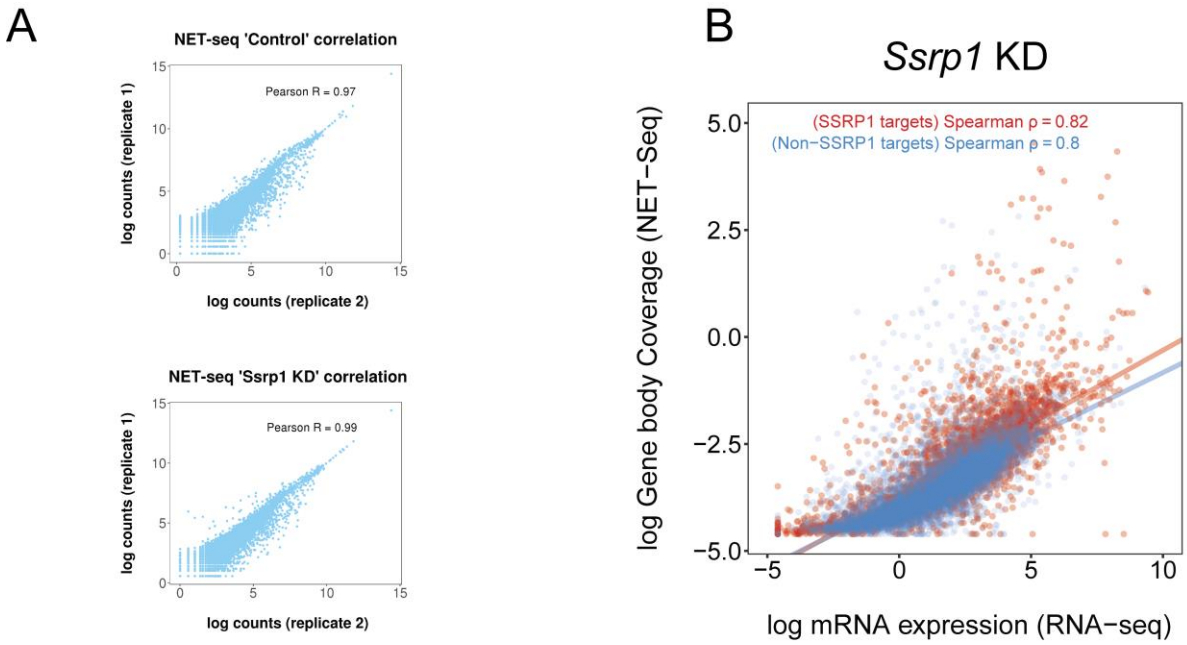
926

927

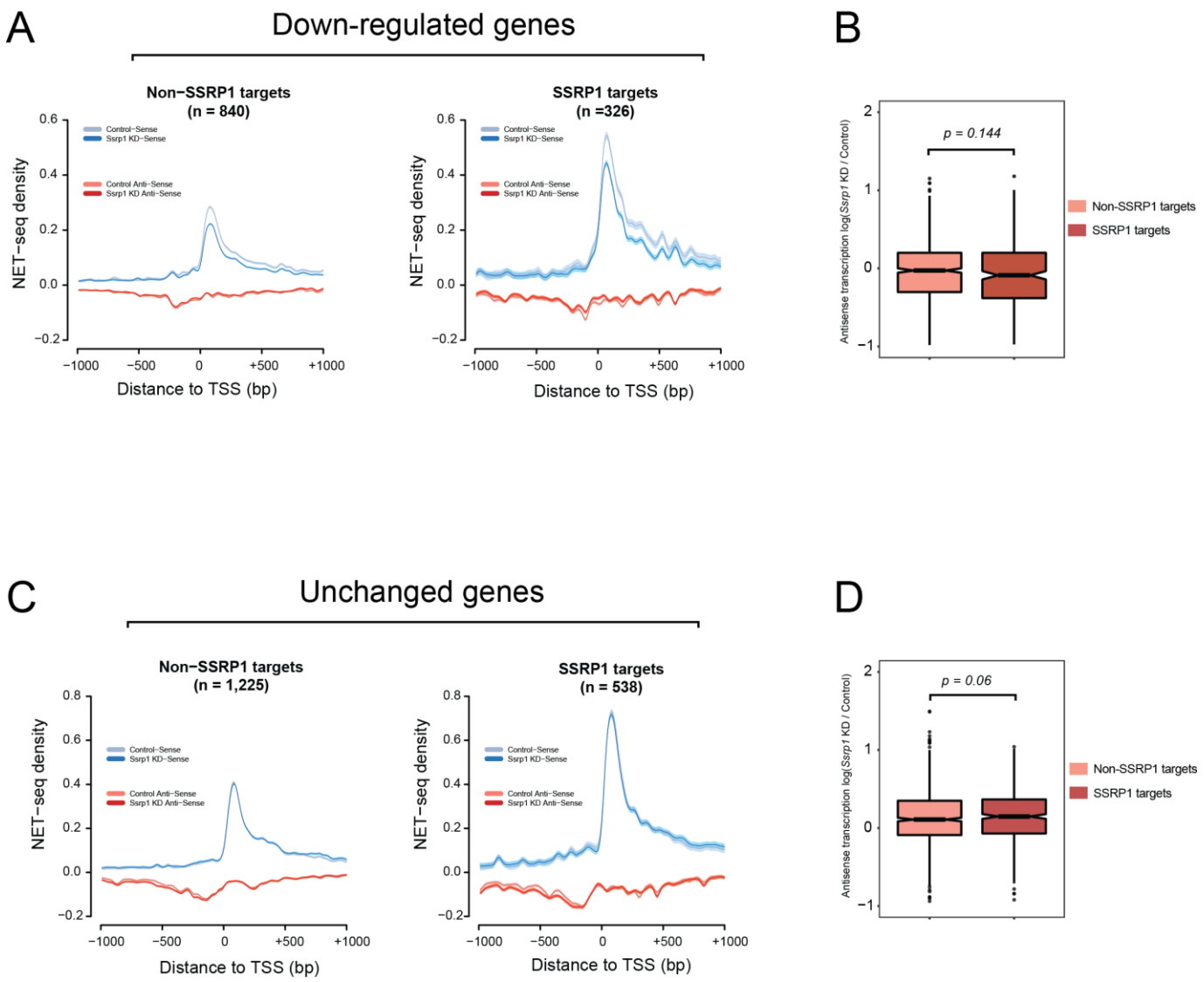
928

929

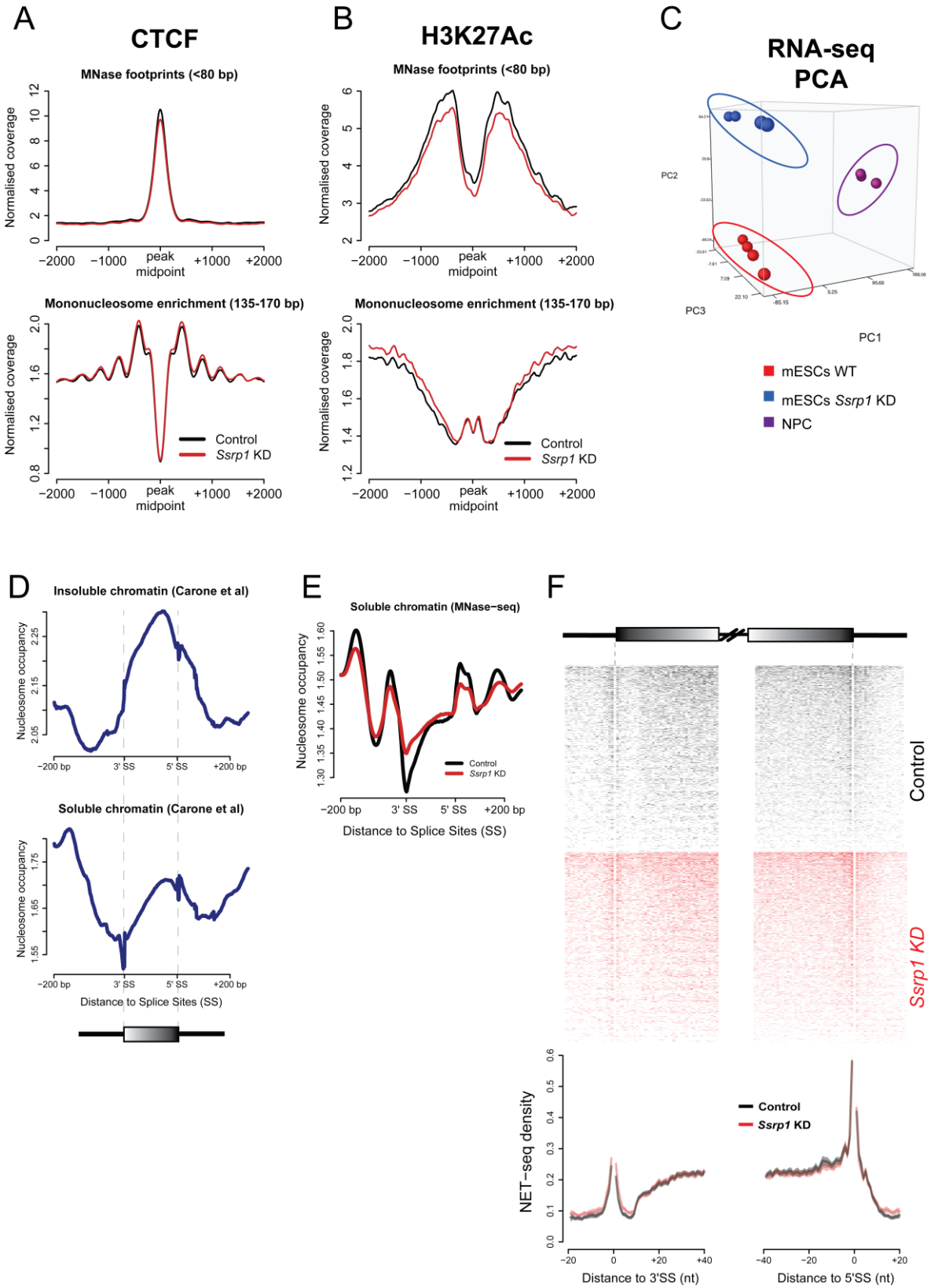
Supplementary Figure 4: Reproducibility assessment of MNase / ATAC-seq datasets. A, Histogram showing the sequenced paired-end fragments lengths extracted from Control and *Ssrp1* depleted conditions. Fragments between 135-170 bp (indicated in red) have been computationally selected and used to plot mono-nucleosomal occupancy over promoter regions. **B,** Correlation scatterplots (MNase-seq) accessing replicate reproducibility in each condition. Pearson's correlation is indicated at the top of each plot. **C,** Nucleosome occupancy metaplots for each replicate for the composite metaplot in **Figure 4A** (Up-regulated; SSRP1 targets). This nucleosome occupancy at the promoter region derives from three distinctive gene clusters of diverse gene expression and H3K4me3 levels. **D,** GC content frequency of all "Up-regulated" and "Down-regulated" genes. Data are also split by FACT-bound dependency. **E,** Correlation scatterplots (ATAC-seq) accessing replicate reproducibility in each condition. Pearson's correlation is indicated at the top of each plot.



Supplementary Figure 5: Assessment of NET-seq datasets. A, Correlation scatterplots (NET-seq) assessing replicate reproducibility in each condition. Pearson's correlation is indicated at the top of each plot. **B,** Scatterplots of log gene body coverage (NET-seq) versus log mRNA expression (RNA-seq) for SSRP1 (n=4,576) and non-SSRP1 (n=8,844) target regions in the *Ssrp1* KD state (Z-score = 7.2, $P < 1 \times 10^{-5}$)



Supplementary Figure 6: Pol II pausing over other gene classes. A, NET-seq density plots (Control and *Ssrp1* KD group) of Down-regulated genes split by FACT-bound status (Non-SSRP1 and SSRP1 targets). Solid lines indicate mean values, whereas the shading represents the 95% confidence interval. **B,** Boxplots assessing fold change (*Ssrp1* KD vs Control) in anti-sense transcription (NET-seq) in a window 1000 bp upstream of the TSS. The Welch's Two-sided t-test was used to calculate significance between Non-SSRP1 and SSRP1 targets. **C,** Same as **(A)** but for genes whose expression does not change (Unchanged genes) in the *Ssrp1* KD condition. **D,** Same as **(B)** but for Unchanged genes.



959

960 **Supplementary Figure 7: Quality control assessing MNase-/NET-seq integrity. A**, CTCF
961 peak midpoints were used as a reference and nucleosome occupancy for short (<80 bp) and
962 long (135-170 bp) MNase footprints was plotted. **B**, Same as **(A)** but for H3K27Ac. Both
963 nucleosome profiles over CTCF and H3K27Ac sites are consistent to previous studies (Carone
964 et al., 2014; Teif et al., 2012). **C**, PCA plot showing clustering of different replicates (RNA-seq)
965 presented in this manuscript. **D**, MNase-seq datasets of soluble (higher intron occupancy) and
966 insoluble (higher exon occupancy) chromatin retrieved from Carone *et al*. **E**, Mean nucleosomal
967 density (207,232 exons) of our insoluble MNase-seq dataset. The soluble nucleosome profile in
968 **"C"** is highly consistent to our MNase treated samples for both conditions where nucleosomal
969 occupancy on introns is similar or higher compared to the exons. Identical occupancy has also
970 been observed by chemical mapping of nucleosomes (Voong et al., 2016). **f**. NET-seq
971 heatmaps and density plots over 41,356 exons with the highest Pol II coverage. Solid lines on
972 the NET-seq meta-exon plots indicate the mean values, whereas the shading represents the
973 95% confidence interval.

974

975

976

977

978