The repressive and alleviating nature of FACT shapes the transcriptional landscape in 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

The conserved and essential histone chaperone FACT (Facilitates Chromatin Transcription) reorganizes nucleosomes during DNA transcription, replication and repair and ensures both, efficient elongation of polymerases and nucleosome integrity. In mammalian cells, FACT is a heterodimer, consisting of SSRP1 and SUPT16. Here, we show that in mouse embryonic stem cells (mESCs), depletion of FACT 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 control expression of these genes. On the other hand, a large number of down- regulated genes are controlled by the transcription factor Tfap2c suggesting a cooperation between this transcription factor and FACT. Finally, we demonstrate a role for FACT in cell fate determination and show that FACT depletion primes ES cells for the neuronal lineage.

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 polymerases in the cell and ensure both, efficient elongation of polymerases 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. Using motif analysis, we identified that down-regulated genes were enriched for binding sites of Tfpa2c (also referred to as AP-2γ), a transcriptional regulator previously implicated in developmental processes (Kuckenberg, Kubaczka, & Schorle, 2012; Qiao et al., 2012; Schemmer et al., 2013), whereas up-regulated genes show enrichment for TBP sites. 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) (Mayer et al., 2015). Using these approaches, we have identified a specific gene cluster comprising of key pluripotency factors and genes involved in embryogenesis/ neuronal development that are up-regulated upon FACT depletion, concomitantly with loss of nucleosome density just upstream of the transcription start site, suggesting that maintenance of nucleosome at this position by FACT is part of the mechanism how FACT impacts on the regulation of these genes. We also provide data indicating FACT in the maintenance of the pluripotent state by showing that its depletion leads to faster differentiation into the neuronal lineage.

Results

Occupancy of FACT correlates with increasing RNA levels

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 genes (*n* = 13,348). 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, a lower correlation was observed between FACT and H3K36me3 on a genome wide level. H3K36me3 has been shown to be able to directly recruit FACT to actively transcribed genes (Carvalho et al., 2013). We suspect that the enrichment of FACT subunits around the TSS might mask this potential correlation as 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 remained elevated when we focused on active promoter and enhancer regions (*n* = 52,329) (**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 (**Fig. 1c**).

Regulation of gene expression by FACT

To investigate how FACT orchestrates transcriptional regulation in ESCs, we depleted SSRP1 levels using short hairpin RNAs (shRNA – **Supplementary Fig. 2a**). Surprisingly, we observed an increase in mESC proliferation following *Ssrp1* knock-down (KD) as measured by proliferation rate using MTT 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 terminally differentiated cells, where FACT depletion has no effect on proliferation[8]. 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. Importantly, these findings suggest that FACT can act as activator, but also as repressor of gene expression. Given the surprising finding that FACT seems to serve as a repressor on a large set of genes and the fact that it serves as a transcription elongation factor, depletion of FACT might cause transcription to decrease genome-wide. Such a scenario would influence the RNA-seq results. We therefore repeated the experiment using ERCC spike-in controls (**Supplementary Figure 3**), which confirmed the repressive role of FACT. Next, we wanted to determine if the presence of FACT predicts whether genes are up- or downregulated

upon knock-down. 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**), demonstrating that the presence of FACT 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.

We also observed in the course of the analysis that depletion of FACT leads to differential splicing of genes and identified 356 Exon skipping/inclusion and 97 Intronic retention events following FACT depletion (**Supplementary Fig. 4b,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. 4d**), suggesting that genes involved in neuronal development might be influenced by FACT.

Depletion of FACT induces very specific changes in chromatin accessibility and transcriptional elongation

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 the 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. 5a**). 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 (**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 at the promoter (**Fig. 3a,b**). This difference in nucleosome occupancy at the promoter region is highly reproducible among the different replicates. As there is a high correlation between FACT occupancy and H3K4me3 (**Fig. 1a,b & Supplementary Fig. 4a**), we decided to split this gene class by H3K4me3 levels. We identify three gene clusters with diverse H3K4me3 levels that display distinctive nucleosomal patterns (**Supplementary Fig. 5b**).

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. 5c).** To confirm this difference in chromatin accessibility, we performed ATAC-seq in Control and *Ssrp1*-depleted ES cells (**Supplementary Fig. 5d-f**). In line with the observations of the MNase-seq experiments, we observe an increase in chromatin accessibility in the absence of FACT at the promoter region of up-regulated genes. In combination with the RNA-seq data, this reduction in nucleosome occupancy suggests that FACT might act as a repressor by enabling higher nucleosome density upstream of the TSS.

Over the last decade, it has become apparent that promoter proximal pausing of RNA Pol II plays an important role in regulating gene expression (Jonkers, Kwak, & Lis, 2014). *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). To explicate the interplay between the Pol II – FACT complex, RNA Pol II transcription and to integrate this with the observed changes in

nucleosome occupancy, we performed NET-seq (Mayer & Churchman, 2016) (**Supplementary Fig. 6b**). 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 the SSRP1-target regions in the Control state suggesting higher levels of Pol II pausing and mRNA levels in the presence of FACT (**Fig. 3c**). Nevertheless, in the *Ssrp1* KD state the SSRP1-bound regions maintained a higher slope, suggesting that pausing and elongation speed of Pol II are not controlled entirely by FACT (**Supplementary Fig. 6a**). To confirm this, we measured the travelling ratio of Pol II over down-regulated and up-regulated genes. Indeed, "up-regulated" FACT-bound genes show a lower travelling ratio, but no significant difference was observed among SSRP1-bound genes following FACT depletion, indicating that FACT is not involved into the release of Pol II towards successive elongation (**Fig. 3d**). Nevertheless, FACT-bound up-regulated promoters seem to exhibit an intrinsic, different type of transcriptional pausing. One piece of evidence that these genes might be regulated differently at the step of pausing comes from differences in Pol II S2ph (ChIP-seq) density profiles (**Supplementary Fig. 6c**). In conclusion, up- and down-regulated genes upon FACT depletion show different nucleosome occupancy pattern and GC-content. FACT depletion leads to specific alterations in nucleosome density, and does not greatly influence RNA Pol II release from promoters.

Tfap2c cooperates with FACT to enhance transcription in ES cells

Next, we sought to understand how FACT can act as activator and repressor. We started off by interrogating whether up- and down-regulated genes might be controlled by different sets of transcription factors. Indeed, motif analysis of the 500bp upstream of the TSS using the iRegulon (Janky et al., 2014) extension for Cytoscape revealed that differentially expressed

genes are enriched for different sets of transcription factors. FACT-bound, up-regulated genes were enriched for transcription factor binding motifs of Retinoic Acid Receptor Beta (RARB – 63/518, NES=4.575) and TATA-box binding protein (TBP, 124/518, NES=4.551). In yeast, FACT is genetically linked to TBP (Biswas et al., 2006) by enabling TBP binding to the TATA-box at promoters (Biswas, Yu, Prall, Formosa, & Stillman, 2005). Hence, we wanted to investigate if there is a direct interaction between TBP and FACT in ES cells. We did not observe any expression changes for RARB or TBP neither at an mRNA (**Fig. 2b**) nor proteome level (**Supplementary Table 8**). We then performed immunoprecipitations (IPs) against SSRP1 or SUPT16 followed by Mass spectrometry to fully interrogate protein-protein interactions (**Fig. 4b**). However, no direct interaction between FACT and RARB or TBP was observed.

Intriguingly, in the promoter of 308 out of 453 FACT-bound down-regulated genes we identified a binding motif for Tfap2 (**Fig. 4a**, NES=3.903). Tfap2 transcription factors regulate a variety of developmental processes (Kuckenberg et al., 2012). To investigate whether components of the Tfap2 transcription factor family might mediate FACT-dependent downregulation, we examined the RNA-seq dataset. We identified Tfap2c (also known as AP2-γ) mRNA levels as being significantly downregulated (**Fig. 2b**). However, in total proteome analysis (**Supplementary Table 8**) comparing scramble control and SSRP1 knock-down, we could not detect down-regulation of Tfap2c, indicating that the transcriptional effect was not translated yet into a protein level. Nevertheless, by interrogating the FACT interactome, we identified Tfap2c as one of the top interactors (**Fig.4b**). We therefore sought to determine whether binding of Tfap2c might be compromised on active promoters upon FACT depletion. IP for H3K4me3 followed by mass spectrometry, in both Control and SSRP1 KD conditions, revealed that apart from both FACT subunits, also Tfap2c exhibits reduced binding to H3K4me3 (**Fig. 4c**). ChIP-qPCR over promoters of predicted Tfap2c target genes (*Prdm16*, *Runx1*, *Wnt10a*) confirmed the reduced binding of Tfap2c in the *Ssrp1* KD state (**Fig. 4d**). These data support a co-operative model

between FACT and TFAP over those genes in ES cells. Interestingly, over-expression of Tfap2c in ES cells leads to trophectoderm development (Kuckenberg et al., 2012), whereas inhibition favours neuronal differentiation (Qiao et al., 2012).

Nucleosome loss upon FACT depletion upstream of the TSS coincides with antisense transcription

Down-regulation of genes in the absence of FACT might at least be partially explained by loss of Tfap2c recruitment to their promoters. To better comprehend the repressive functions of FACT, we turned to the identified clusters with different levels of H3K4me3 that gave rise to clusters with distinct changes in nucleosomal occupancy (**Supplementary Figure 5c**). Intriguingly, we observed that only the FACT-bound, high H3K4me3 regions exhibit a broad distribution of the chromatin mark (**Fig. 5a**); a result consistent with Set1 (responsible for H3K4me3 establishment) distribution over the gene body area. Key pluripotency factors (e.g. Oct4, Sox2, Nanog, and Klf4) and genes involved in embryogenesis belong to this gene cluster. No broad H3K4me3 and Set1 patterns, were identified in the “unchanged” or “down-regulated” gene classes (**Supplementary Fig. 7a,c**). In addition, we examined the distribution of several nucleosome remodellers (de Dieuleveult et al., 2016) around the promoter of all three gene classes (“Down-regulated”, “Unchanged”, “Up-regulated”) to determine whether they might cooperate with FACT and potentially be a predictor for the change in gene expression upon FACT depletion. No change was observed in occupancy for Chd1, Chd2, or Chd4, but we detected a striking difference in Ep400, Smarcad1, and Chd6/8/9 (**Supplementary Fig. 8**). This implicates that changes in the presence and occupancy degree of the latter ones might influence the transcriptional activity upon depletion of FACT levels.

We then focused more on the high and low H3K4me3 gene clusters in order to understand the consequences of Pol II directionality. H3K4me3 serves as a recruitment platform for TFs and RNA Pol II (Sims et al., 2007). NET-seq density plots identified that FACT targets displayed higher levels of promoter-recruited RNA Pol II at the TSS compared to FACT-unbound promoters, but also a strong and wide Pol II distribution similar to the H3K4me3 mark with little signal of divergent Pol II in the anti-sense strand (**Fig. 5b**). The broad distribution in the sense strand was also confirmed by employing published GRO-seq data (**Supplementary Fig. 7e**). Non-SSRP1 target regions exhibited an equal distribution both in convergent and divergent Pol II (H3K4me3 High – **Fig. 5b**). Upon knock-down of FACT, SSRP1 targets display higher levels of convergent transcription, and even more strikingly, an increase ($P < 10^{-27}$ and $P < 10^{-20}$) in divergent transcription compared to the non-SSRP1 targets. This occurred precisely at the locations in which nucleosomes were depleted upon knock-down of FACT ($P < 10^{-8}$, $P < 10^{-6}$) suggesting that the presence of FACT inhibits antisense transcription by maintaining higher nucleosome density upstream of the TSS. It is of note here as well that genes up-regulated and targeted by FACT exhibit significantly increased levels of nucleosome occupancy ($P < 10^{-4}$) upstream of the TSS compared to genes not bound by FACT (**Fig. 5b**).

Importantly, this increase in antisense transcription and loss of nucleosome occupancy is observed only in genes that are up-regulated upon FACT depletion (**Fig. 5b & Supplementary Fig. 7b, d**), indicating that these SSRP1-bound promoters favour uni-directional transcription by providing a nucleosome barrier upstream of the TSS. This is in line with the observation that these promoters harbour lower GC content and are enriched in TBP-binding sites. 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). In addition, similar observations have been reported upon depletion of FACT function in *S. cerevisiae* by using thermosensitive alleles of *spt6* and *spt16*, albeit in yeast upregulation of

sense/antisense transcription upon FACT depletion 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; Kaplan, Laprade, & Winston, 2003). Taken together, these 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. Given the gene ontology enrichment for neuronal terms in the mRNA isoform analysis, the enrichment for RARB motifs in up-regulated and Tfap2c in down-regulated genes, it was plausible that these gene expression changes might impact the differentiation potential of ES cells into the neuronal lineage. To test this, 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) that express key neurogenesis markers (Pax6, Nes, Tubb3) but still maintain FACT and key pluripotency factors at a high level (**Fig. 6a**). 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. 6b**) and are over-represented in pathways involved in neuronal development. To test, if these genes would follow the observed changes in nucleosome occupancy and induction in bi-directional transcription, we split these over-lapping genes according to FACT-bound status and interrogated nucleosome/ Pol II occupancy. Interestingly, all the neurogenesis-associated genes (based on gene ontology enrichment) found on this list are bound by FACT (n=38) and only these genes bound by FACT show a significant increase ($P < 10^{-15}$) in divergent anti-sense

transcription concomitantly with loss of a nucleosomal impediment ($P < 10^{-46}$) (**Fig. 6c**). This mechanism is also evident in lowly expressed/repressed genes that are up-regulated upon FACT depletion (**Supplementary Fig. 9)**.

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. 6d**).

Discussion

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, key pluripotency factors are up-regulated and lowly expressed developmental factors are further down-regulated. In addition, pro-proliferative genes are up-regulated, which results in the hyper-proliferation of the FACT-depleted ES cells. Moreover, the FACT-depleted gene signature has a large overlap with the 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 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[29]. 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.

The profile of FACT occupancy over active genes is reminiscent of a profile of the Ser5 phosphorylated form of RNA Pol II and might reflect binding of FACT to RNA Pol II. A similar profile for SSRP1 has been reported recently in HT1080 cells[8]. Interestingly, the profiles in mammalian cells differ substantially from the ones obtained in yeast cells, with the latter displaying high FACT occupancy over the gene body[30]. This might indicate that there are subtle differences between the function of FACT in *S. cerevisiae* and mammalian cells. All genes up-regulated upon depletion of FACT display a lower pausing index compared to down-regulated or unchanged genes, whereas RNA Pol II over FACT-bound promoters displays an even more decreased travelling ratio. However, FACT depletion does not influence these ratios, arguing that FACT has no effect on the release of paused Pol II and its successful progression towards transcriptional elongation.

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). Down-regulated genes harbour a motif for the transcription factor Tfap2c and using mass spectrometry approaches we could provide evidence that FACT and Tfap2c can cooperate. The observed decrease in binding of Tfap2c to promoters might also explain partly the priming of ES cells for the neuronal lineage upon FACT depletion as down-regulation of Tfap was demonstrated to accelerate neuronal differentiation (Qiao et al., 2012). A similar mechanistic link between developmentally important transcription factors and FACT has been reported for muscle development, where FACT can interact with myogenin (Lolis et al., 2013). Apart from its cooperation with transcription factors, it is tempting to speculate that another 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 just upstream of 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. This gene class also shows a much higher nucleosome density upstream of the TSS compared to all other gene classes analysed. 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. In order to address these questions, rapid depletion of FACT is required in

combination with time course experiments addressing loss of nucleosomes vs. upregulation of antisense transcription. However, it is clear that this observed effect is very specific to FACT-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. 5b**). 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. Such a scenario 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 the nucleosomal occupancy at the TSS and suppression of antisense transcription. In ES cells, genes repressed by FACT in this way encode key pluripotency factors and genes involved in embryogenesis, particularly in early neuronal differentiation, which is accelerated when FACT is depleted (**Fig. 6e**). Therefore, the elevated expression of both FACT subunits observed in undifferentiated cells might be required to fine-tune the transcriptional landscape, along with other nucleosome remodellers, in order to maintain this undifferentiated state.

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.

Acknowledgments

We would like to thank Ilian Attanassov of the Max Planck Institute for Biology of Ageing Proteomics Core Facility for Mass Spectrometry Analysis. 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 and members of the Tessarz laboratory for discussion and comments on the manuscript. This work was funded by the Max Planck Society.

Deposition of sequencing data

Data have been deposited in Gene Expression Omnibus (GEO) under accession numbers GSE 90906 (ChIP-seq, RNA-seq, chrRNA-seq, MNase-seq, ATAC-seq, and NET-seq).

Accession numbers and references of publicly available datasets. H3K4me3, H3K27me3, Pol II S5ph, H3K4me1, H3K27Ac, CTCF (ENCODE Consortium – E14 cell line); Ep400, Chd1, Chd2, Chd4, Chd6, Chd8, Chd9 (de Dieuleveult et al., 2016) : GSE64825; Set1(Cxxc1) (Denissov et al., 2014): GSM1258239; p53 (Li et al., 2012): GSE26360; Pol II S2ph (Brookes et al., 2016): GSM850470; GRO-seq (Min et al., 2011): GSE27037; Smarcad1 (Xiao et al., 2017) : GSE45338.

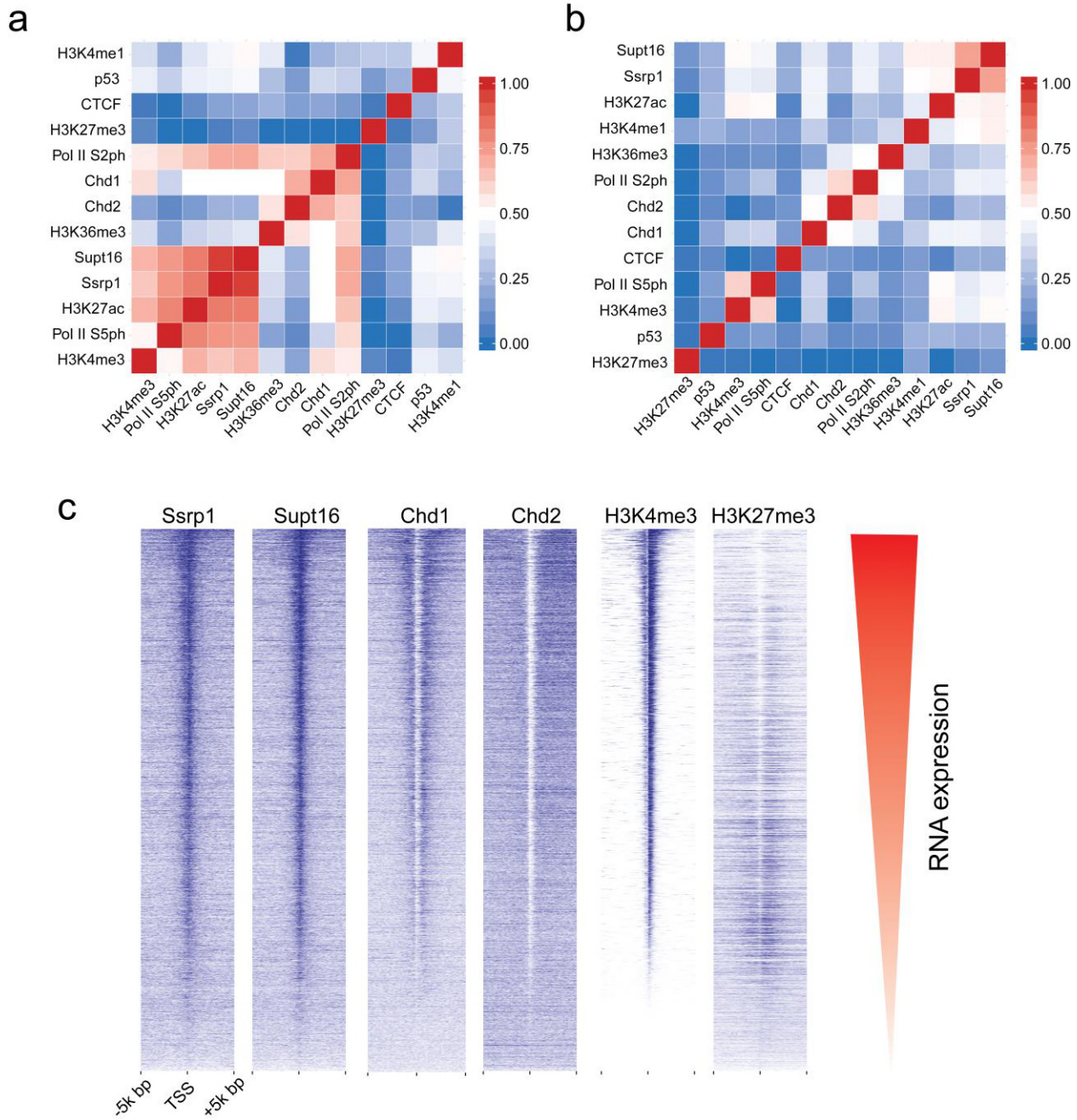


Figure 1: **Correlated occupancies across FACT-bound regions**. **a**, Heatmap representing Pearson's correlation between FACT subunits (SSRP1, SUPT16), and other factors over the gene body area of all uniquely annotated genes (*n* = 13,348). **b**, Same as **a** but for promoter/enhancer regions (*n* = 52,329) characterised by H3K27ac and/or H3K4me1 marks. **c,** Distribution of FACT and other factors (ChIP-seq tags indicated in blue) over the TSS of 13,348 unique RefSeq genes, sorted by H3K4me3 levels. Coinciding RNA expression levels are shown in red.

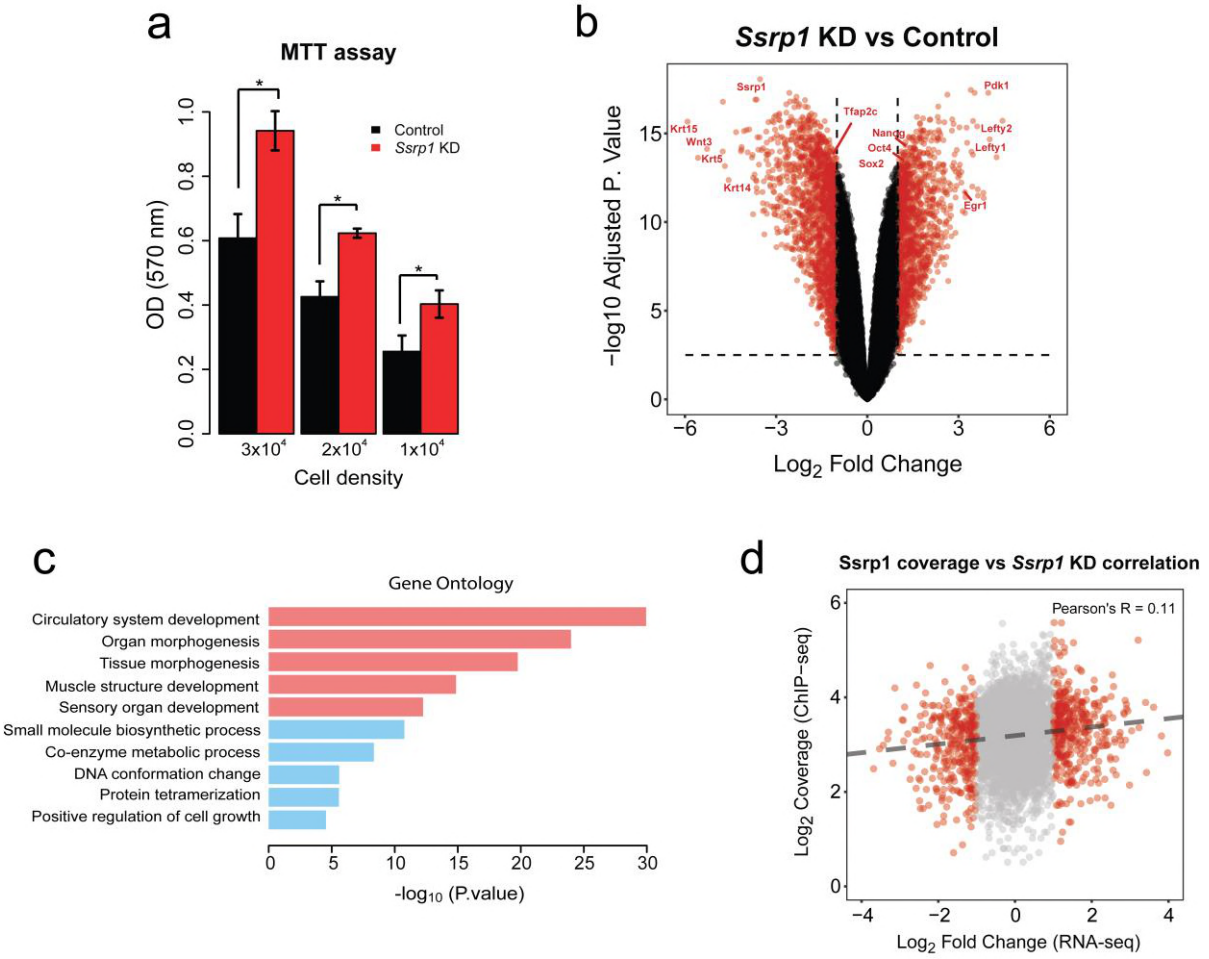
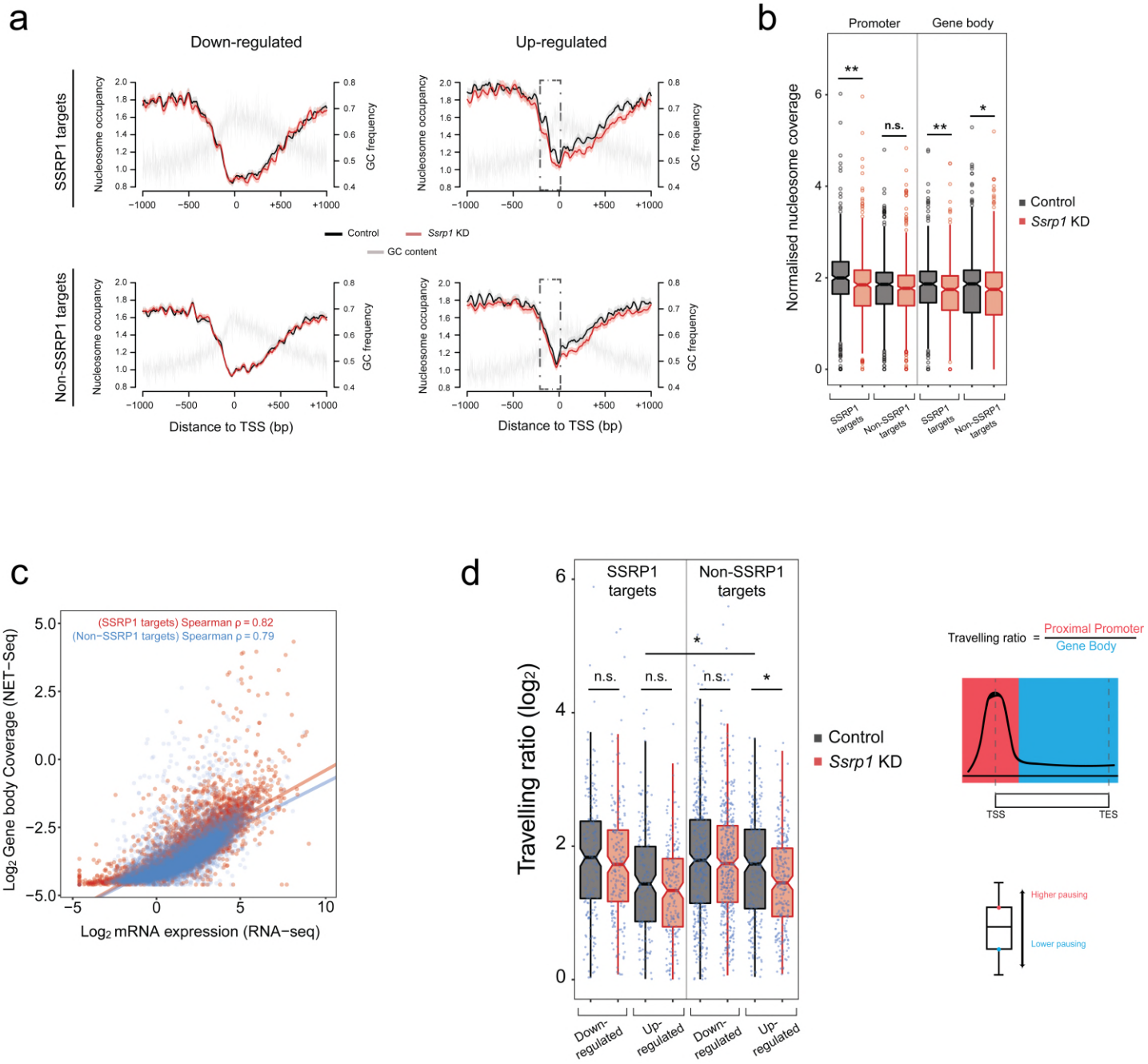


Figure 2: **Regulation of gene expression by FACT.** **a**, MTT assay assessing cell metabolic activity in mESCs at different cell densities following depletion of FACT levels. Values are mean and SE of three independent transfection experiments are displayed. Significance was calculated via a two-tailed *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).



452

453

454

455

456

Figure 3: **Regulation of gene expression by FACT through nucleosome deposition and Pol II pausing.** **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 a Wilcoxon rank test. **c.** 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, $P < 10^{-5}$). **d.** 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 as boxplots. The Wilcoxon rank test was used to calculate significance between Control and *Ssrp1* KD (* $p < 0.05$, n.s. =not significant).

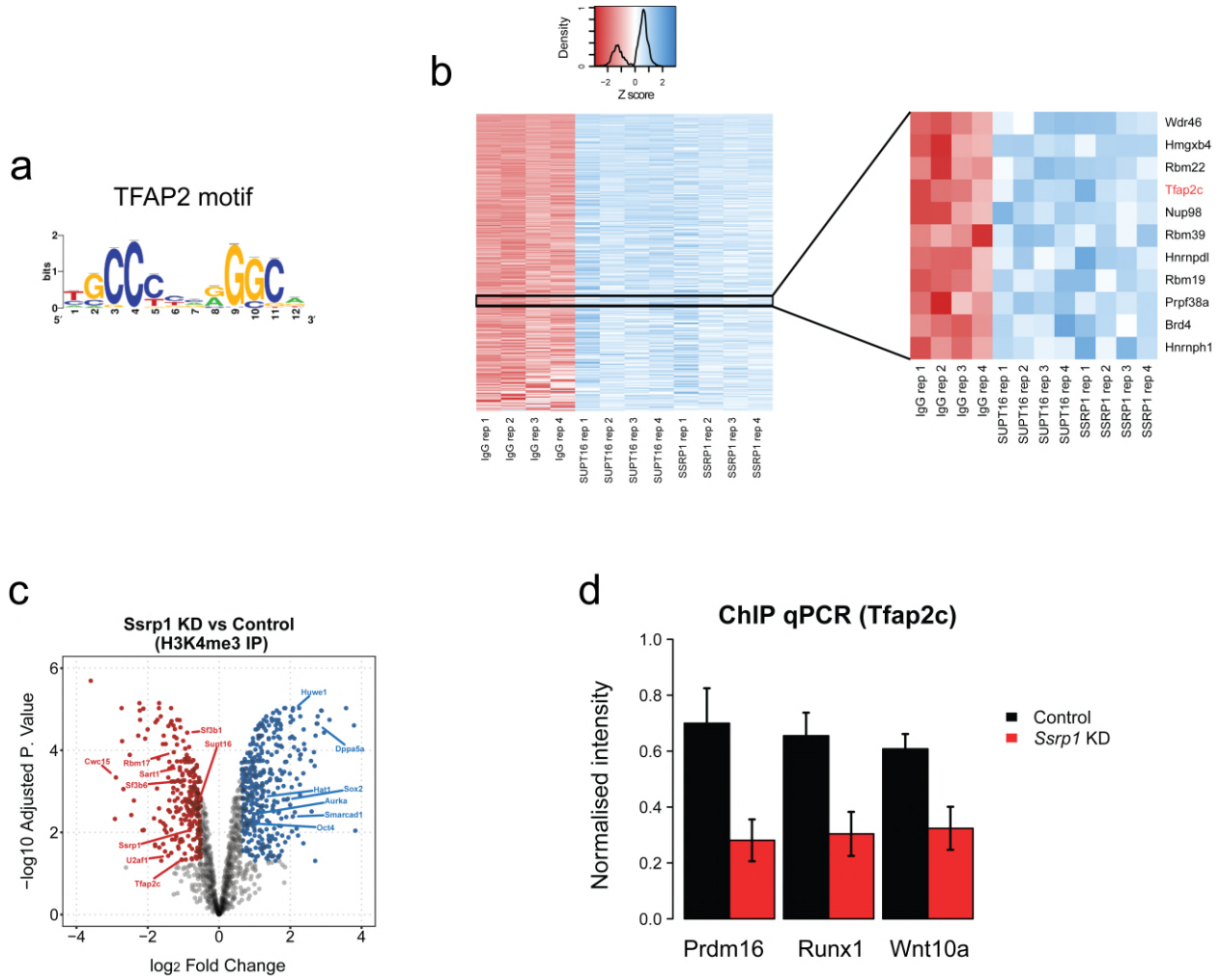
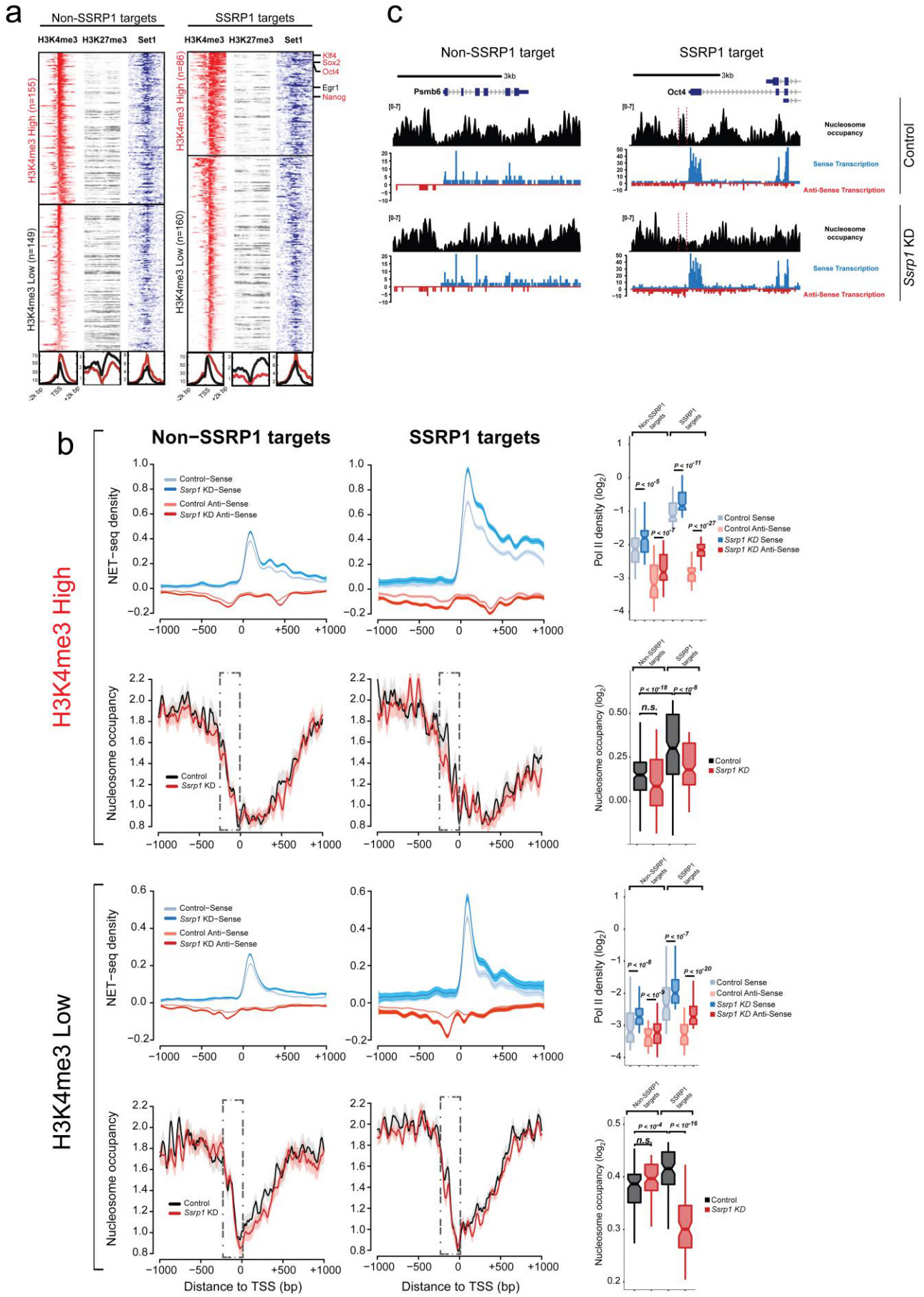
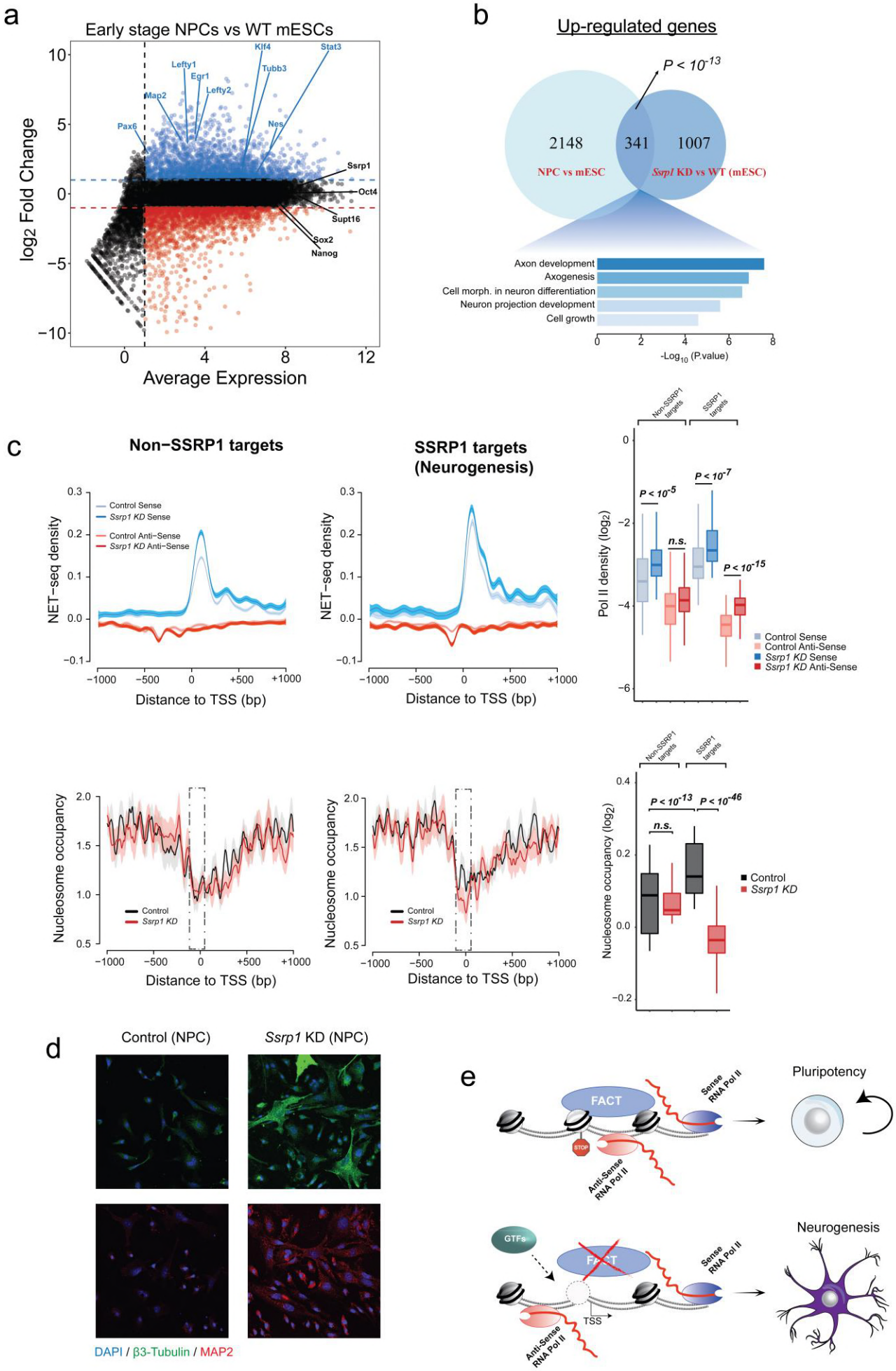


Figure 4: **Tfap2c co-operates with FACT to control gene expression.** **a**, Motif analysis for all down-regulated genes. TFAP2 is highly enriched over those gene promoters. **b,** Heatmap representing a fraction of the FACT whole protein interactome (n = 329). Highly enriched proteins are visualised in blue. **c,** Volcano plot of depleted/enriched proteins at H3K4me3 following *Ssrp1* depletion. **d,** Barplots measuring Tfap2c binding (ChIP-qPCR) over specific gene promoters, both in Control (black) and *Ssrp1* KD (red) conditions. Data were normalised to IgG mock control


a
Non-SSRP1 targets
SSRP1 targets
H3K4me3 H3K27me3 Set1
H3K4me3 High (n=155)
H3K4me3 Low (n=149)
H3K4me3 High (n=86)
H3K4me3 Low (n=160)
Klf4
Sox2
Oct4
Egr1
Nanog
c
Non-SSRP1 target
SSRP1 target
Psmb6
Oct4
Nucleosome occupancy
Sense Transcription
Anti-Sense Transcription
Control
Ssrp1 KD
b
Non−SSRP1 targets
SSRP1 targets
H3K4me3 High
H3K4me3 Low
NET−seq density
Control−Sense
Ssrp1 KD−Sense
Control Anti−Sense
Ssrp1 KD Anti−Sense
Nucleosome occupancy
Control
Ssrp1 KD
Distance to TSS (bp)
Pol II density (log2)
Nucleosome occupancy (log2)
Control Sense
Control Anti-Sense
Ssrp1 KD Sense
Ssrp1 KD Anti-Sense
n.s.

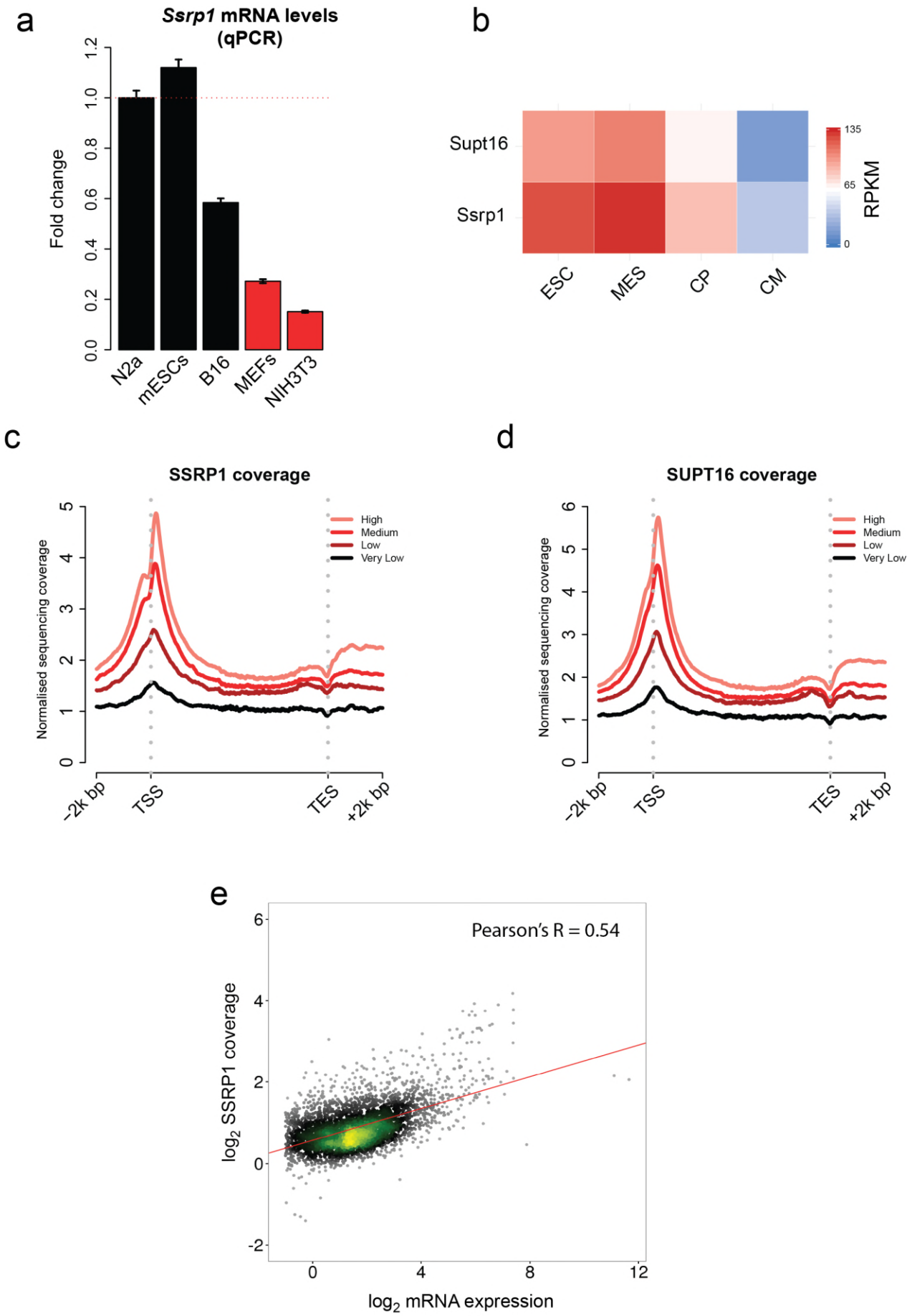
483

Figure 5: **Pausing and unidirectional passage of Pol II by FACT. a**, Occupancy heatmaps and profiles ("Up-regulated" class) for H3K4me3, H3K27me3, and Set1 split by chromatin (High H3K4me3 and Low H3K4me3) and FACT binding status (SSRP1 and Non-SSRP1 targets) in WT mESCs. Density was calculated relative to the TSS ± 2000 bp. In metaplot profiles, black and red lines represent H3K4me3 High and H3K4me3 Low chromatin respectively. **b**, NET-seq and nucleosome occupancy plots (Control and *Ssrp1* KD group) split by chromatin and FACT-bound status. Solid lines on the MNase-seq and NET-seq metaplots indicate the mean values, whereas the shading represents either the SE of the mean (MNase-seq) or the 95% confidence interval (NET-seq). Significant changes in Pol II/ nucleosome density were calculated using a Wilcoxon rank test. **c**, Transcriptional activity (NET-seq) and nucleosome occupancy (MNase-seq) of SSRP1 (*Oct4*) and non-SSRP1 (*Psmb6*) targets between Control and *Ssrp1* KD conditions. Nucleosomal loss in *Oct4* is indicated by a red-dotted line.


a
Early stage NPCs vs WT mESCs
log2 Fold Change
Average Expression
Klf4
Stat3
Lefty1
Egr1
Tubb3
Lefty2
Map2
Pax6
Nes
Ssrp1
Oct4
Supt16
Sox2
Nanog
b
Up-regulated genes
P < 10-13
2148
341
1007
NPC vs mESC
Ssrp1 KD vs WT (mESC)
Axon development
Axogenesis
Cell morph. in neuron differentiation
Neuron projection development
Cell growth
-Log10 (P.value)
c
Non−SSRP1 targets
SSRP1 targets (Neurogenesis)
NET−seq density
Control Sense
Ssrp1 KD Sense
Control Anti−Sense
Ssrp1 KD Anti−Sense
Distance to TSS (bp)
Non-SSRP1 targets
SSRP1 targets
Pol II density (log2)
P < 10-5
n.s.
P < 10-7
P < 10-15
Control Sense
Control Anti-Sense
Ssrp1 KD Sense
Ssrp1 KD Anti-Sense
Nucleosome occupancy
Control
Ssrp1 KD
Nucleosome occupancy (log2)
P < 10-13
P < 10-46
d
Control (NPC)
Ssrp1 KD (NPC)
DAPI / β3-Tubulin / MAP2
e
FACT
Sense RNA Pol II
Anti-Sense RNA Pol II
STOP
Pluripotency
GTFs
TSS
Neurogenesis

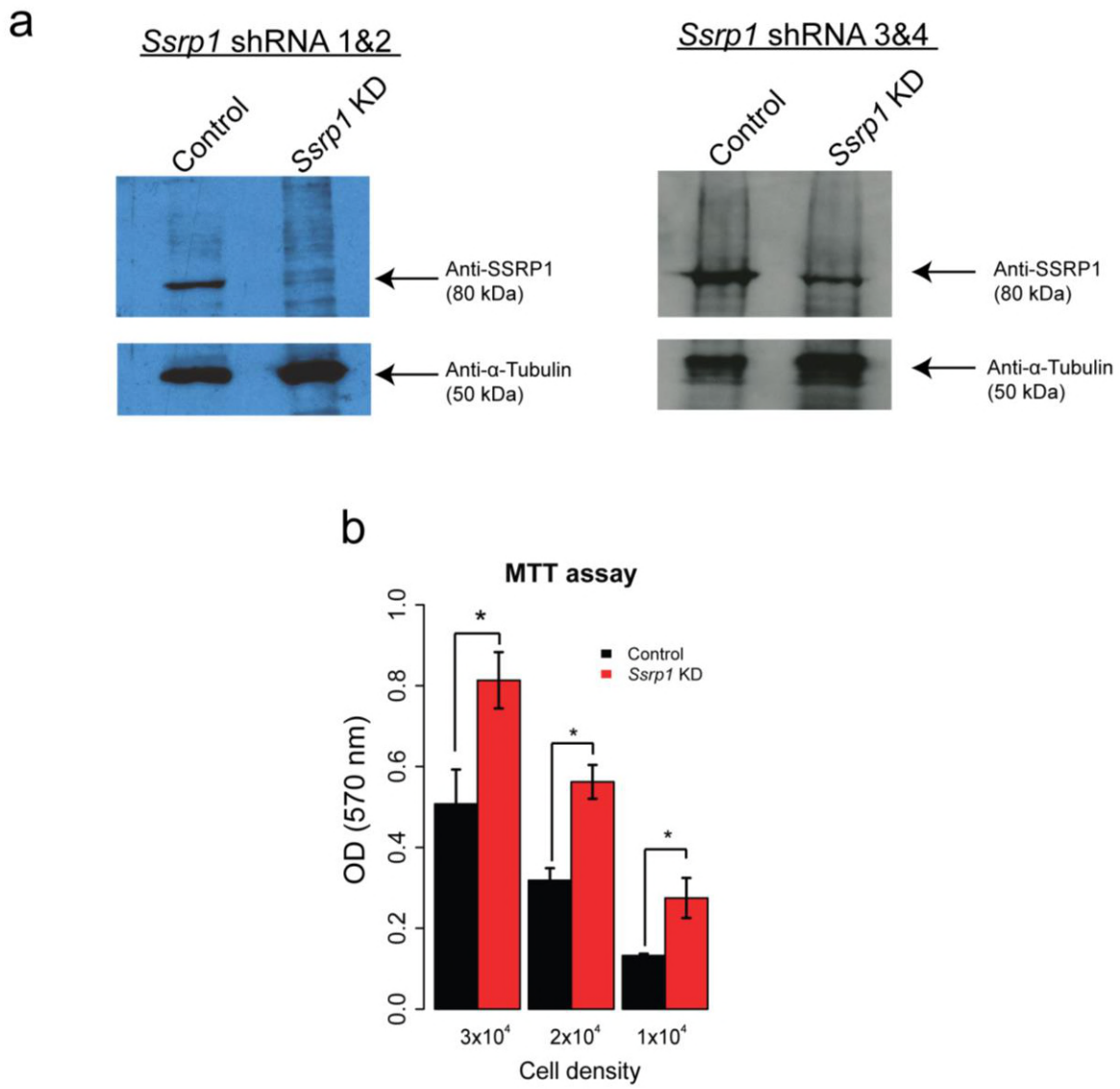
498

Figure 6: **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**, NET-seq and nucleosome occupancy plots (Control and *Ssrp1* KD group) split by chromatin and FACT-bound status. Solid lines on the MNase-seq and NET-seq metaplots indicate the mean values, whereas the shading represents either the SE of the mean (MNase-seq) or the 95% confidence interval (NET-seq). Significant changes in Pol II/ nucleosome density were calculated using a Wilcoxon rank test. **d**, Immunofluorescence (IF) analysis of early stage NPCs following *Ssrp1* depletion. (Blue) DAPI, nuclei; (Green) β3-Tubulin, neurons; (Red) MAP2, dendrites. **e**, 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 GTFs, increased gene expression, and ultimately, activation of neurogenesis cues.

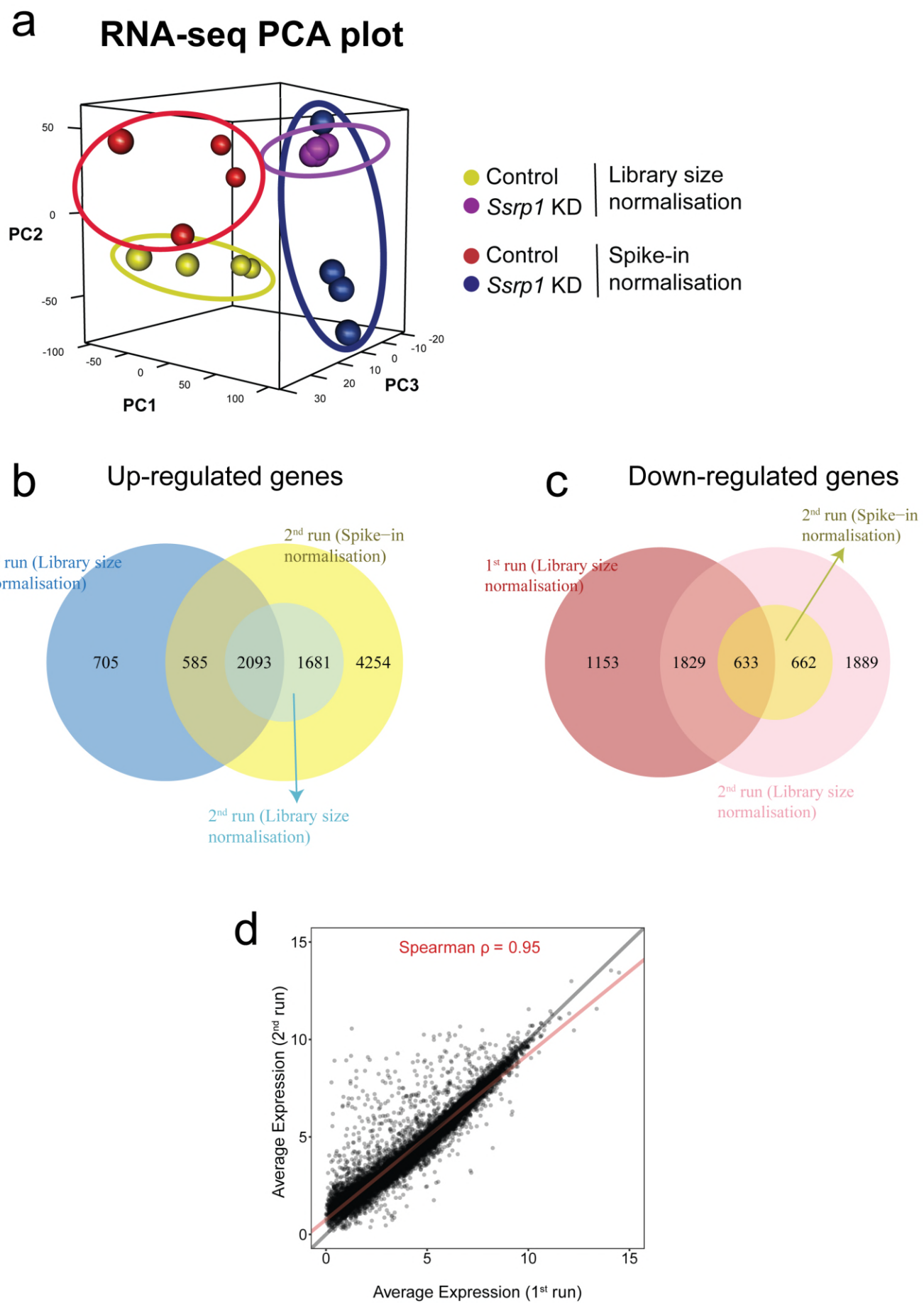

a
Ssrp1 mRNA levels (qPCR)
Fold change
0.0
0.2
0.4
0.6
0.8
1.0
1.2
N2a
mESCs
B16
MEFs
NIH3T3
b
Supt16
Ssrp1
ESC
MES
CP
CM
135
65
0
RPKM
c
SSRP1 coverage
Normalised sequencing coverage
High
Medium
Low
Very Low
-2k bp
TSS
TES
+2k bp
d
SUPT16 coverage
Normalised sequencing coverage
High
Medium
Low
Very Low
-2k bp
TSS
TES
+2k bp
e
Pearson's R = 0.54
log2 SSRP1 coverage
log2 mRNA expression

517

Supplementary Figure 1: FACT correlates with active gene expression. a, qPCR of *Ssrp1* levels among stem (mESCs), cancer (N2a, B16), and differentiated cell lines (MEFs, NIH3T3). Mean and SE of three biological replicates are shown. **b**, Heatmap assessing the mRNA levels (RPKM) of FACT (*Ssrp1, Supt16*) at different timepoints of differentiation of ES cells to cardiomyocytes (Wamstad et al., 2012) (ESC = Embryonic Stem Cells, MES = Mesodermal cells, CP = Cardiac Precursors, CM = cardiomyocytes).**c**, Distribution of SSRP1 relative to the TSS (± 2000 bp) and the TES (± 2000 bp) for four different gene classes ranked by level of RNA abundance (High, Medium, Low, Very Low). **d**, Same as **c** but for SUPT16. **e**, Scatterplot of the $\log_2$ SSRP1 coverage over $\log_2$ mRNA expression.

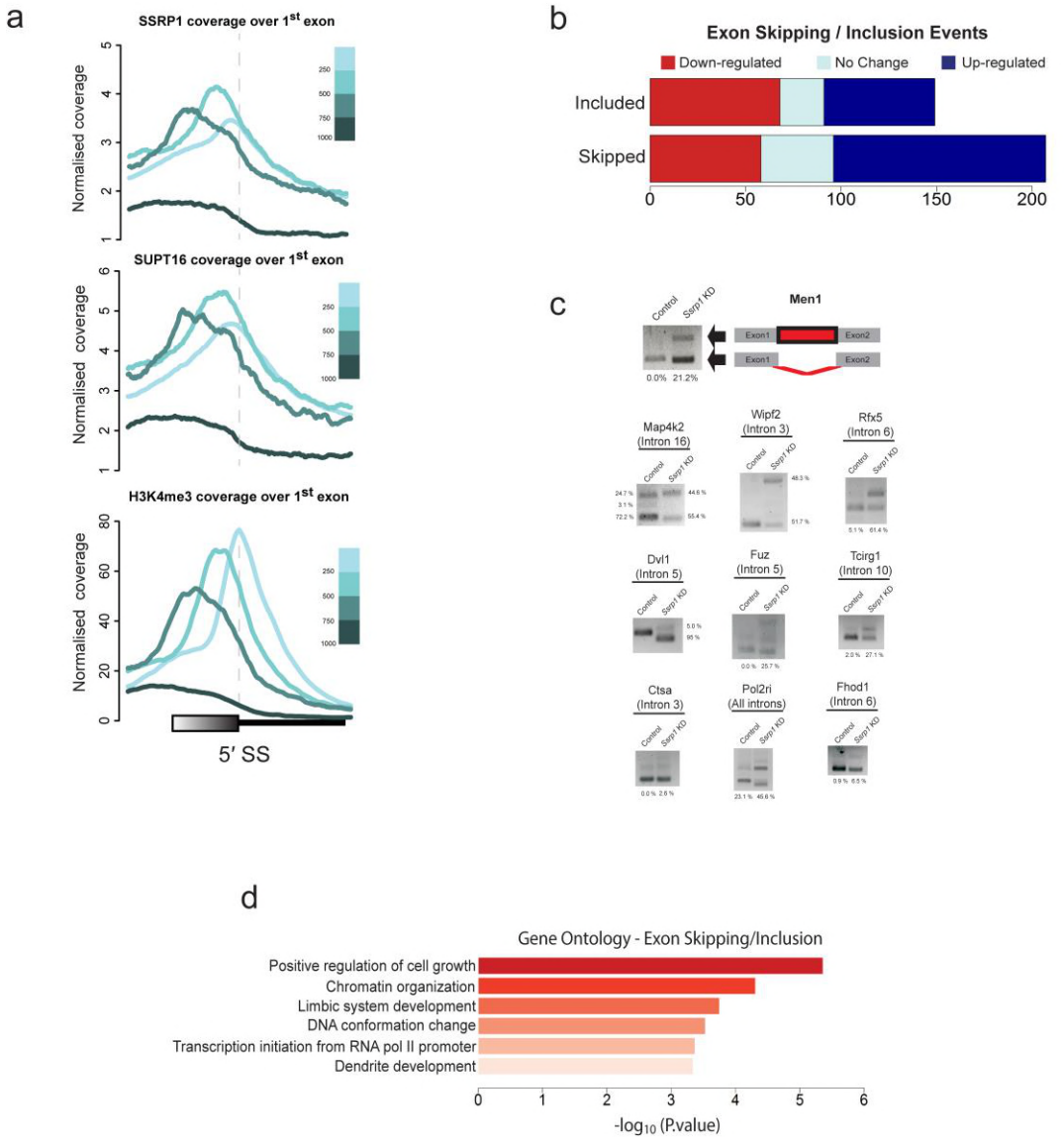


Supplementary Figure 2: FACT depletion promotes increased proliferation rate in mESCs. a, Western blots after transfection of mESCs with different combinations of *Ssrp1* shRNA vectors (shRNA 1&2 or shRNA 3&4). Anti-α-Tubulin was used as a reference. **b,** MTT assay following transfection with *Ssrp1*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

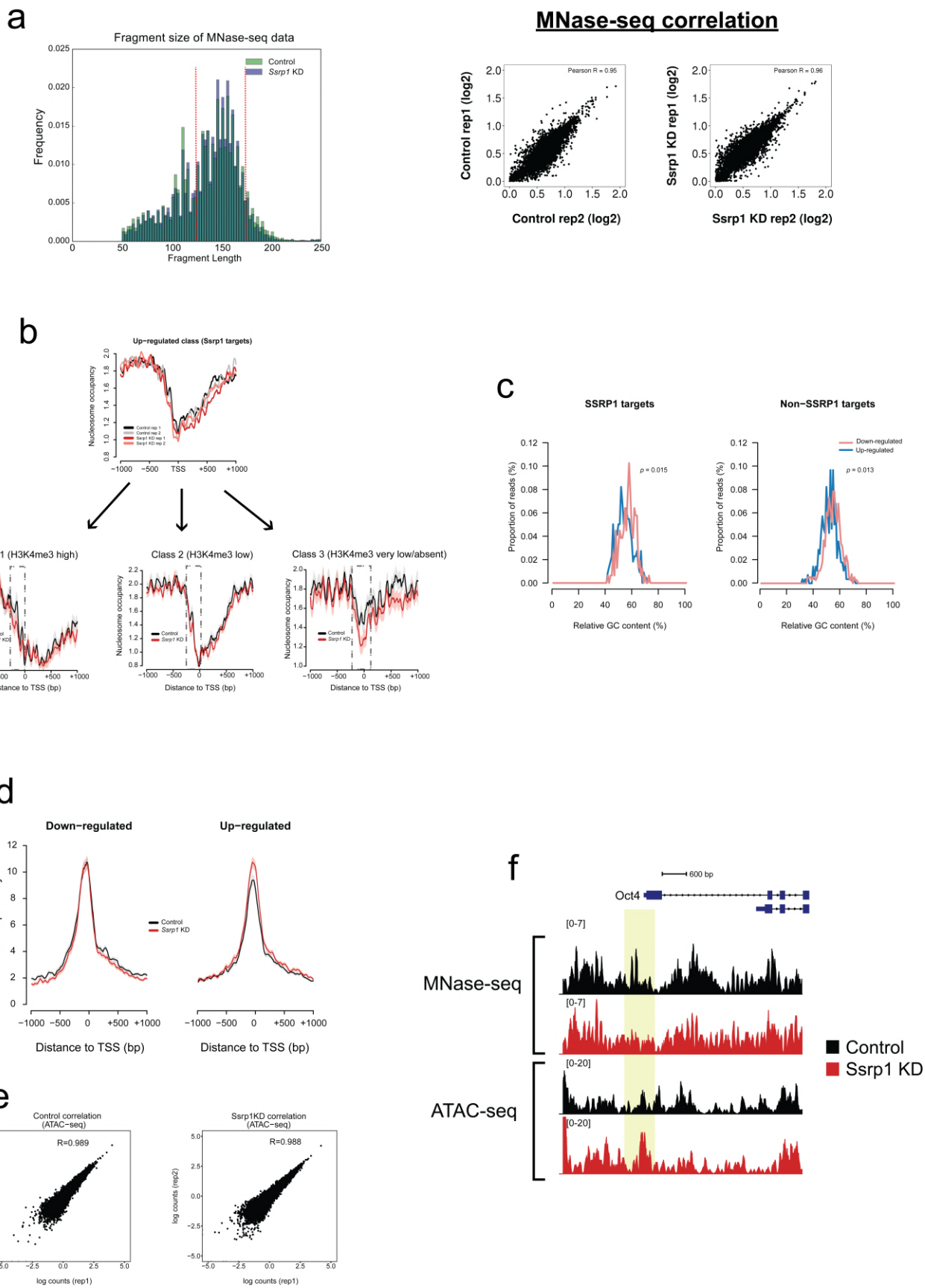

a
RNA-seq PCA plot
PC2
PC1
PC3
Control
Ssrp1 KD
Library size normalisation
Control
Ssrp1 KD
Spike-in normalisation
b
Up-regulated genes
1st run (Library size normalisation)
2nd run (Spike−in normalisation)
705
585
2093
1681
4254
2nd run (Library size normalisation)
c
Down-regulated genes
2nd run (Spike−in normalisation)
1st run (Library size normalisation)
1153
1829
633
662
1889
2nd run (Library size normalisation)
d
Spearman ρ = 0.95
Average Expression (2nd run)
Average Expression (1st run)

539

Supplementary Figure 3: Assessment of spike-in normalisation in RNA-seq data. a, PCA plot for all Controls and SSRP1-depleted RNA-seq files following either library size or spike-in normalisation. A spike-in normalisation causes higher divergence among the different replicates of the same condition. **b**, Venn diagrams showing the overlap of up-regulated genes following *Ssrp1* KD among the 1st (Library size normalisation) and 2nd RNA-seq (Library size or spike-in normalisation) runs. **c**, Same as **b**, but for down-regulated genes. **d**, Correlation plot between the average gene expression of the 1st run vs 2nd run. Spearman correlation is displayed in red. The red abline represents the slope of the plot. A high overlap in differential expression is observed between library size (both 1st and 2nd run) and spike-in normalisation. However, spike-in normalisation defines many genes (n=4,254) as being up-regulated in the *Ssrp1* KD state which probably arises due to the variance among replicates. We conclude that spike-in normalisation is not the most appropriate method for normalising RNA-seq data.

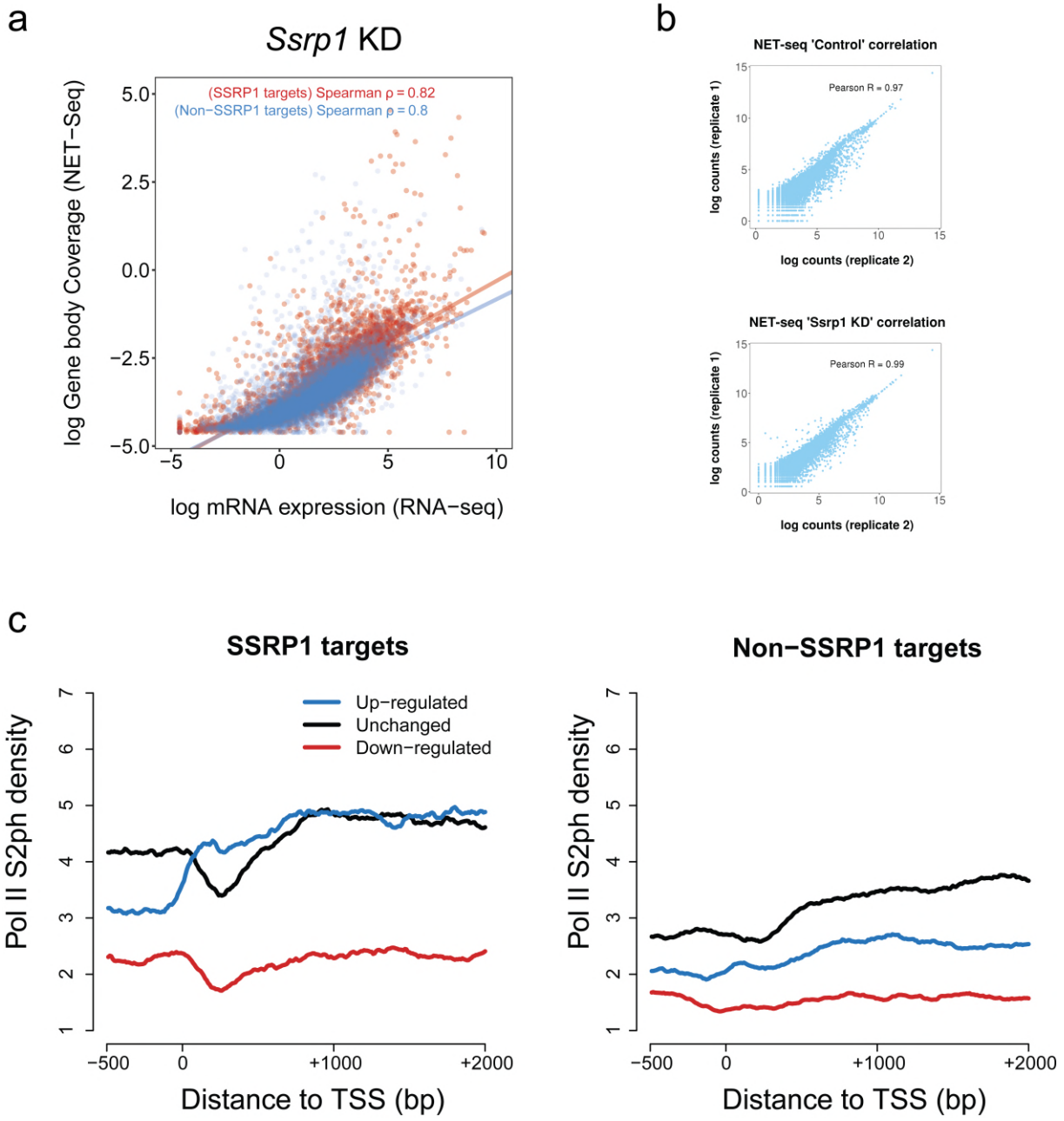


Supplementary Figure 4: FACT facilitates alternative splicing of RNA transcripts. a, Average distribution of SSRP1, SUPT16, and H3K4me3 aligned to the 5' Splice Site (5'SS) of all genes grouped by first exon length. **b**, Barplots representing the number of included/skipped exons categorized by their gene expression status (red: “Down-regulated”, cyan: “No change”, blue: “Up-regulated”). In total, we have identified 149 included and 207 skipped exon events in the *Ssrp1* KD group. **c**, Graphical representation of an intronic retention event (*Men1*) in the KD group. Also, analysis of intron inclusion events or isoform switches after FACT depletion. Unspliced transcript percentage was measured according to band intensity. **d,** Gene ontology analysis of transcripts (FACT-bound) that display alternative exon usage between the two conditions

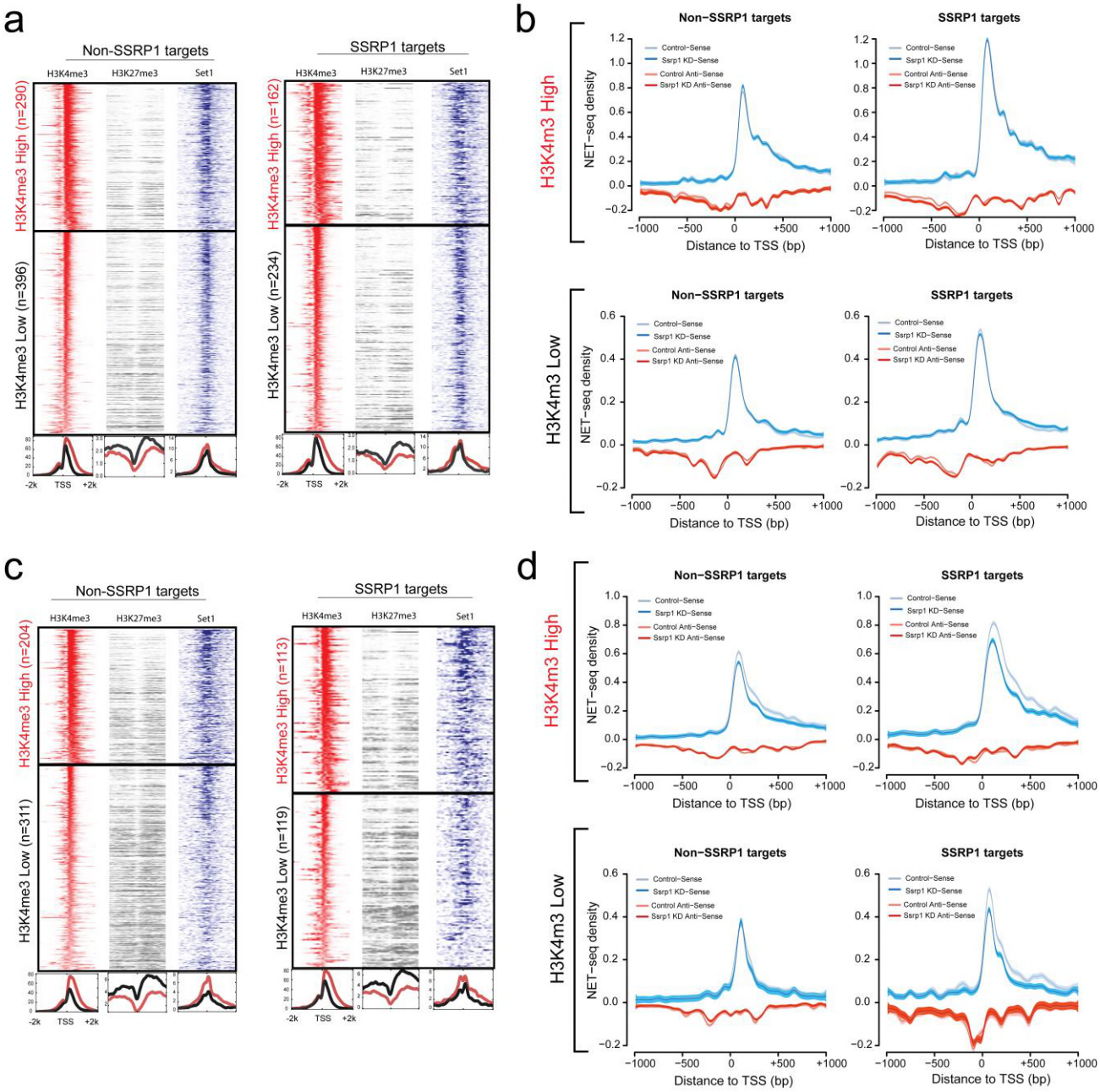

a
Fragment size of MNase-seq data
Control
Ssrp1 KD
Frequency
Fragment Length
MNase-seq correlation
Pearson R = 0.95
Pearson R = 0.96
Control rep1 (log2)
Control rep2 (log2)
Ssrp1 KD rep1 (log2)
Ssrp1 KD rep2 (log2)
b
Up−regulated class (Ssrp1 targets)
Nucleosome occupancy
TSS
Class 1 (H3K4me3 high)
Class 2 (H3K4me3 low)
Class 3 (H3K4me3 very low/absent)
Distance to TSS (bp)
c
SSRP1 targets
Non−SSRP1 targets
Down−regulated
Up−regulated
Proportion of reads (%)
Relative GC content (%)
d
Down−regulated
Up−regulated
ATAC−seq density
Control
Ssrp1 KD
Distance to TSS (bp)
e
Control correlation (ATAC−seq)
Ssrp1KD correlation (ATAC−seq)
R=0.989
R=0.988
log counts (rep1)
log counts (rep2)
f
600 bp
Oct4
MNase-seq
ATAC-seq
[0-7]
[0-20]
Control
Ssrp1 KD

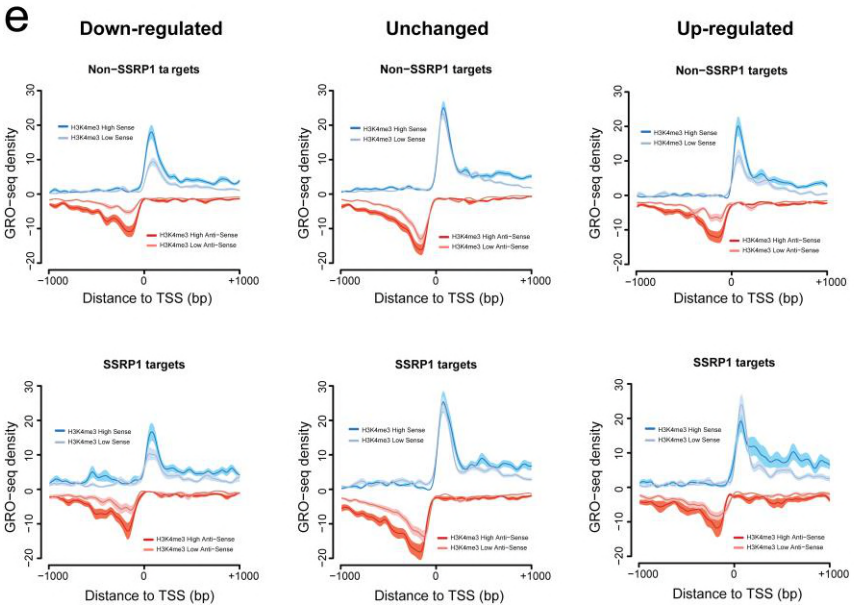
567

Supplementary Figure 5: 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. Correlation scatterplots (MNase-seq) accessing replicate reproducibility in each condition. Pearson's correlation is indicated at the top of each plot. **b,** Nucleosome occupancy metaplots for each replicate for the composite metaplot in **Figure 3a**. This nucleosome occupancy at the promoter region derives from three distinctive gene clusters of diverse gene expression. **c,** GC content frequency of all "Up-regulated" and "Down-regulated" genes. Data are also split by FACT-bound dependency. **d,** Metaplot of open chromatin assessed by ATAC-seq among Down-regulated and Up-regulated genes both in Control and *Ssrp1* KD conditions. **e,** Correlation scatterplots (ATAC-seq) assessing replicate reproducibility in each condition. Pearson's correlation is indicated at the top of each plot. **f,** Interrogation of nucleosome occupancy (MNase-seq) and chromatin accessibility (ATAC-seq) over a *Oct4* for Control and *Ssrp1* KD conditions. Changes in nucleosome occupancy and chromatin accessibility are highlighted in yellow.

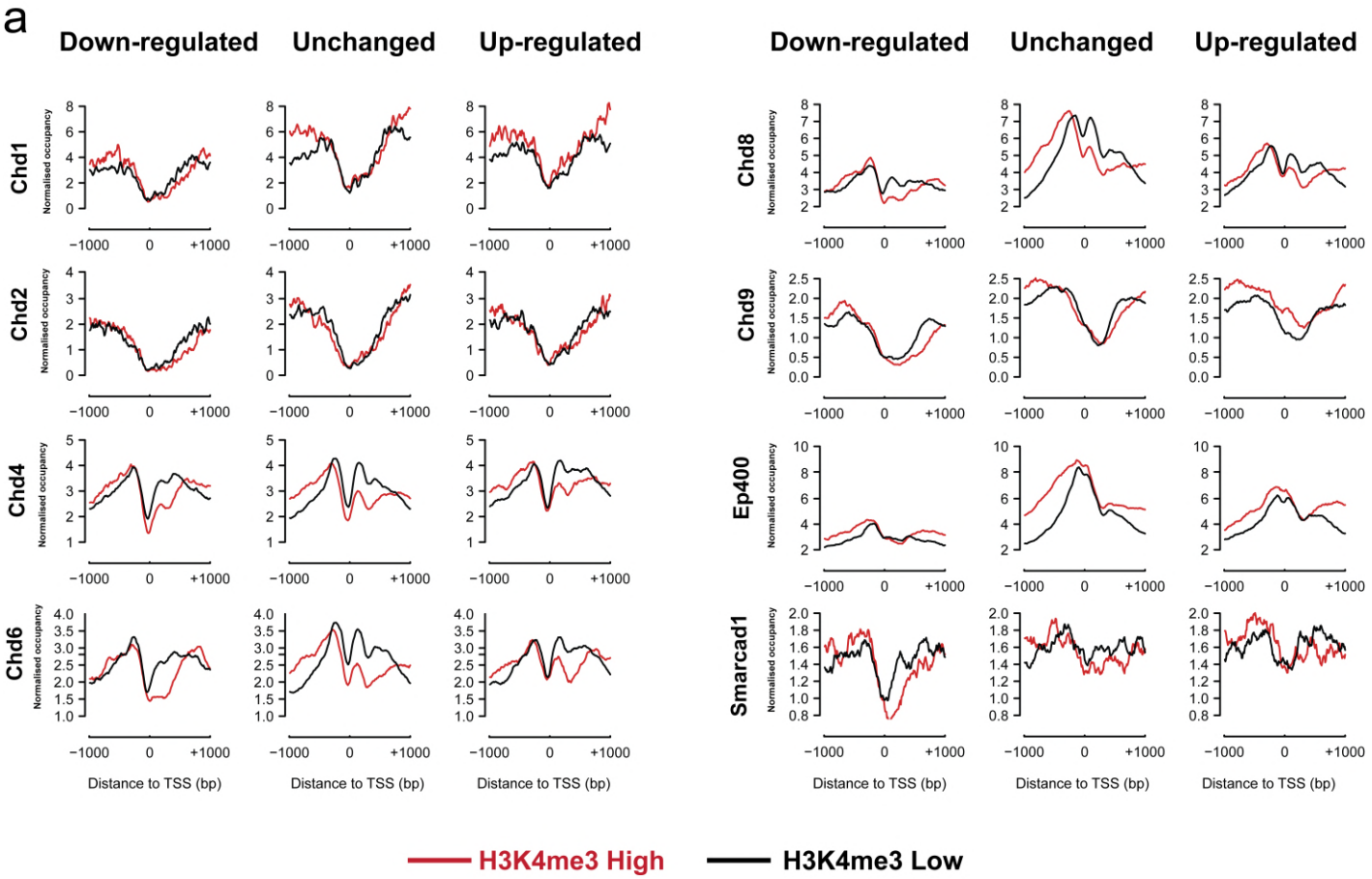


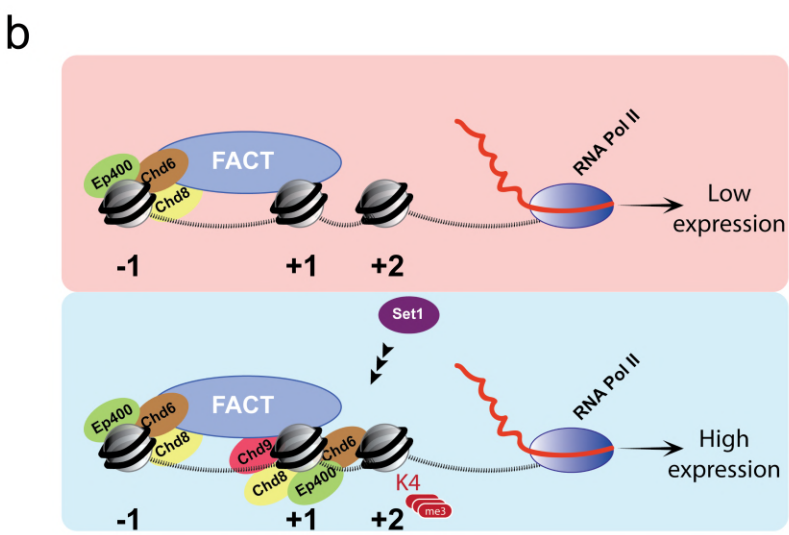
Supplementary Figure 6: Distinctive Pol II elongation progression over a specific FACT-bound gene class. a, Scatterplots of log gene body coverage (NET-seq) versus log mRNA expression (RNA-seq) for SSRP1 (n=4,576) and non-SSRP1 (n=8,844) target regions in the *Ssrp1* KD state (Z-score = 7.2, $P < 1 \times 10^{-5}$). **b,** Correlation scatterplots (NET-seq) assessing replicate reproducibility in each condition. Pearson's correlation is indicated at the top of each plot. **c,** Distribution of elongating Pol II variant (S2 ph – WT ES cells) over the TSS (± 2000 bp) of SSRP1- and Non-SSRP1-bound regions split by transcriptional regulation ("Up-regulated", "No Change", "Down-regulated"). The "Up-regulated" SSRP1-bound gene class exhibits a different occupancy pattern


a
Non-SSRP1 targets
SSRP1 targets
H3K4me3
H3K27me3
Set1
H3K4me3 High (n=290)
H3K4me3 Low (n=396)
H3K4me3 High (n=162)
H3K4me3 Low (n=234)
-2k
TSS
+2k
b
Non−SSRP1 targets
SSRP1 targets
H3K4m3 High
H3K4m3 Low
NET−seq density
Control−Sense
Ssrp1 KD−Sense
Control Anti−Sense
Ssrp1 KD Anti−Sense
Distance to TSS (bp)
c
H3K4me3 High (n=204)
H3K4me3 Low (n=311)
H3K4me3 High (n=113)
H3K4me3 Low (n=119)
d

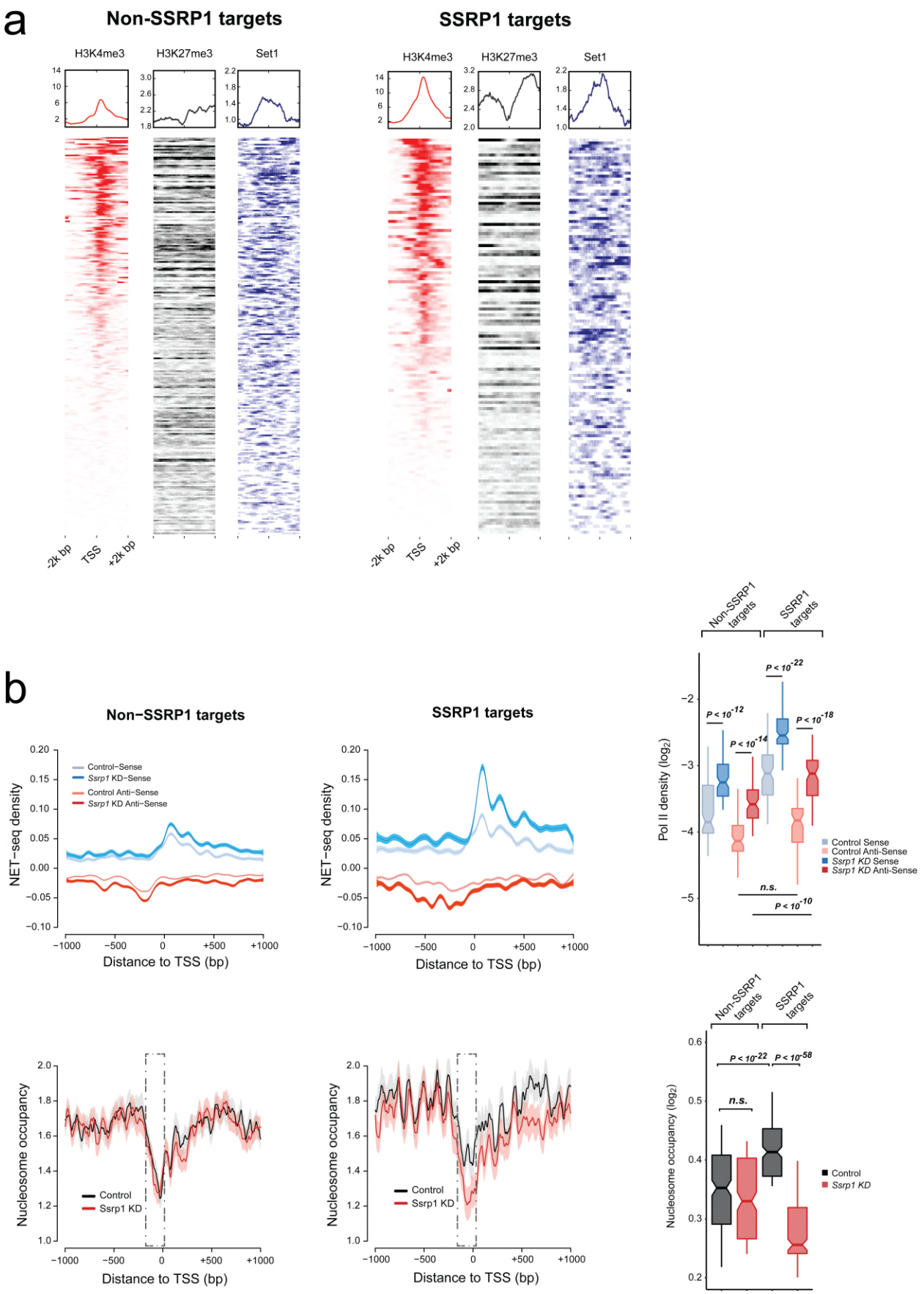

e
Down-regulated
Unchanged
Up-regulated
Non−SSRP1 ta rgets
Non−SSRP1 targets
SSRP1 targets
GRO−seq density
H3K4me3 High Sense
H3K4me3 Low Sense
H3K4me3 High Anti−Sense
H3K4me3 Low Anti−Sense
Distance to TSS (bp)

Supplementary Figure 7: Pol II pausing over other gene classes. a) Occupancy heatmaps and profiles ("Unchanged" class) for H3K4me3, H3K27me3, and Set1 split by chromatin (High H3K4me3 and Low H3K4me3) and FACT binding status (SSRP1 and Non-Ssrp1 targets). Occupancy was calculated relative to the TSS ± 2000 bp. In metaplot profiles, black and red lines represent H3K4me3 High and H3K4me3 Low chromatin respectively. **b)** NET-seq density plots for the "No Change" class split by FACT binding status (SSRP1 and Non-Ssrp1 targets) and condition (Control and *Ssrp1* KD). Occupancy was calculated relative to the TSS ± 2000 bp. **c)** Same as **a** but for the "Down-regulated" class. **d)** Same as **b** but for the "Down-regulated" class. **e)** GRO-seq density plots derived from WT mESCs for the three different gene classes split by FACT binding status (SSRP1 and Non-Ssrp1 targets). Solid lines on the NET-seq and GRO-seq metaplots indicate the mean values, whereas the shading represents the 95% confidence interval (NET-seq) or SE (GRO-seq).


a
Down-regulated
Unchanged
Up-regulated
Chd1
Chd2
Chd4
Chd6
Chd8
Chd9
Ep400
Smarcad1
Normalised occupancy
Distance to TSS (bp)
H3K4me3 High
H3K4me3 Low

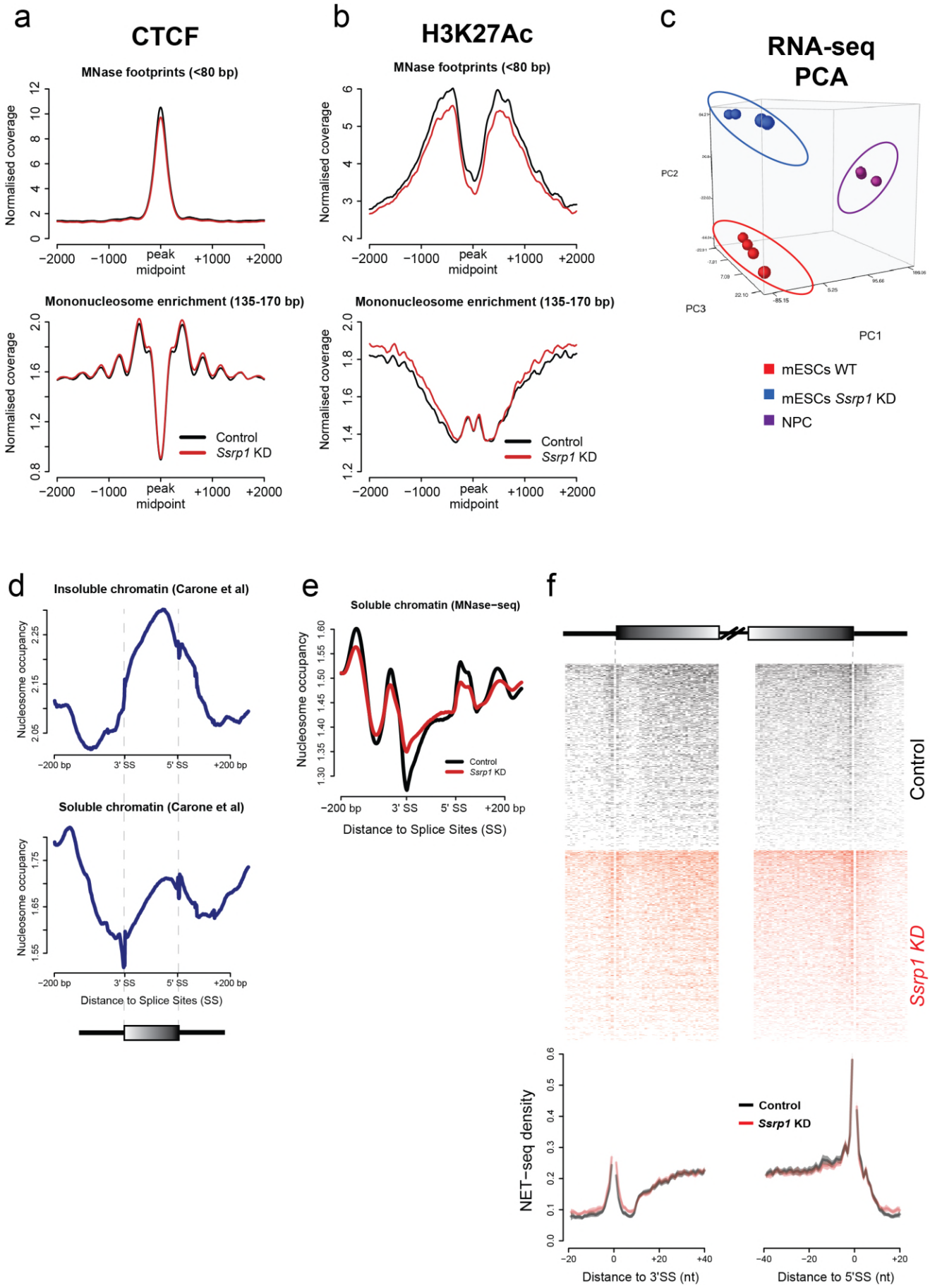

b
FACT
Ep400
Chd6
Chd8
RNA Pol II
Low expression
-1
+1
+2
Set1
Chd9
K4
me3
High expression

Supplementary Figure 8: Distribution of different nucleosome remodellers over the TSS of FACT-bound promoters. Datasets are split by their relative chromatin (H3K4me3 High, H3K4me3 Low) and gene expression status ("Down-regulated", "Unchanged", "Up-regulated"). Model of how FACT and chromatin remodellers might cooperate to ensure proper transcription of their target genes. The whole model was derived by comparing the H3K4me3 High class between the "Up-regulated" (High expression) and "Down-regulated" (Low expression) gene clusters. In the upper model, the presence of FACT along with other remodellers (Ep400, Chd6, and Ch8) on the -1 nucleosome enables a conformational chromatin organisation that keeps lineage-specific genes at a low expression level. In the 2nd model, the presence of FACT enables binding of Set1 and ultimately establishment of a broad H3K4me3 mark. In addition, the Ep400-Chd6-Chd8-Chd9 complex in the +1 nucleosome probably marks those promoters for up-regulation in the absence of FACT.


a
Non-SSRP1 targets
SSRP1 targets
H3K4me3
H3K27me3
Set1
-2k bp
TSS
+2k bp
b
Non−SSRP1 targets
SSRP1 targets
Control−Sense
Ssrp1 KD−Sense
Control Anti−Sense
Ssrp1 KD Anti−Sense
NET−seq density
Distance to TSS (bp)
Pol II density (log2)
Control Sense
Control Anti-Sense
Ssrp1 KD Sense
Ssrp1 KD Anti-Sense
n.s.
Nucleosome occupancy
Control
Ssrp1 KD
Nucleosome occupancy (log2)

632

Supplementary Figure 9: Distribution of chromatin marks, nucleosomes, and Pol II over lowly expressed/ repressed genes. a, Occupancy heatmaps and profiles (**Supplementary Figure 5b** – Class 3) for H3K4me3, H3K27me3, and Set1 split by FACT binding status (SSRP1 and Non-SSRP1 targets). Occupancy was calculated relative to the TSS ± 2000 bp. **b**, NET-seq and nucleosome occupancy plots (Control and *Ssrp1* KD group) split by chromatin and FACT-bound status. Solid lines on the MNase-seq and NET-seq metaplots indicate the mean values, whereas the shading represents either the SE of the mean (MNase-seq) or the 95% confidence interval (NET-seq). Significant changes in Pol II/ nucleosome density were calculated using a Wilcoxon rank test.


a
CTCF
MNase footprints (<80 bp)
Normalised coverage
peak midpoint
Mononucleosome enrichment (135-170 bp)
Control
Ssrp1 KD
b
H3K27Ac
MNase footprints (<80 bp)
Mononucleosome enrichment (135-170 bp)
c
RNA-seq PCA
PC1
PC2
PC3
mESCs WT
mESCs Ssrp1 KD
NPC
d
Insoluble chromatin (Carone et al)
Soluble chromatin (Carone et al)
Nucleosome occupancy
3' SS
5' SS
Distance to Splice Sites (SS)
e
Soluble chromatin (MNase−seq)
f
Control
Ssrp1 KD
NET−seq density
Distance to 3'SS (nt)
Distance to 5'SS (nt)

642

Supplementary Figure 10: Quality control assessing MNase-/NET-seq integrity. a, CTCF peak midpoints were used as a reference and nucleosome occupancy for short (<80 bp) and long (135-170 bp) MNase footprints was plotted. **b**, Same as **a** but for H3K27Ac. Both nucleosome profiles over CTCF and H3K27Ac sites are consistent to previous studies(Carone et al., 2014; Teif et al., 2012). **c**, PCA plot showing clustering of different replicates (RNA-seq) presented in this manuscript. **d**, MNase-seq datasets of soluble (higher intron occupancy) and insoluble (higher exon occupancy) chromatin retrieved from Carone *et al*.**e**, Mean nucleosomal density (207,232 exons) of our insoluble MNase-seq dataset. The soluble nucleosome profile in **"c"** is highly consistent to our MNase treated samples for both conditions where nucleosomal occupancy on introns is similar or higher compared to the exons. Identical occupancy has also been observed by chemical mapping of nucleosomes (Voong et al., 2016). **f**. NET-seq heatmaps and density plots over 41,356 exons with the highest Pol II coverage. Solid lines on the NET-seq meta-exon plots indicate the mean values, whereas the shading represents the 95% confidence interval.

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. ERCC spike-in controls (Ambion-ExFold RNA spike-in mixes) were added during library preparation. An assessment of spike-in normalisation can be viewed in **Supplementary Fig. 10**. All of the RNA-seq data presented in this manuscript have been normalised to the total library size. Significant genes with an absolute logFC > 1 and Adjusted P.Value < 0.01 were considered as differentially expressed

(Supplementary Table 1). The “No Change” gene class (n=2,179) was obtained from genes with an Adjusted P.Value > 0.05. The diffSplice function implemented in limma was used to identify differentially spliced exons between the two conditions (Supplementary Table 2). Significant exons with an FDR < 0.001 were considered as differentially spliced. Retention Introns were identified using the MISO(Katz, Wang, Airoldi, & Burge, 2010) (Mixture of Isoforms) probabilistic framework (Supplementary Table 3).

Retention intron events. We verified the presence of retained introns in the *Ssrp1* KD by randomly selecting ten intron retention events. The FastStart SYBR Green Master (Roche) was 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

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* = 13,348) from UCSC (blacklisted regions were removed). The mESC promoter/enhancer regions identified in Shen *et al*(Shen et al., 2012) were used for the generation of **Figure 1b**. H3K4me3 ChIP-seq tag densities were split by k-means clustering into three categories in order to remove genes with low H3K4me3 density.

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. 10**). 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. 9**.

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. 10c,d**, 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” and “Trypsin/P,” allowing for cleavage after lysine and arginine, also when followed by proline. The minimum number of peptides and razor peptides for protein identification was 1; the minimum number of unique peptides was 0. Protein identification was performed at a peptide spectrum matches and protein false discovery rate of 0.01. The “second peptide” option was on in order to identify co-fragmented peptides. Successful identifications were transferred between the different raw files using the “Match between runs” option, using a match time window of 0.7 min. Label-free quantification (LFQ)(Jurgen Cox, Hein, Luber, & Paron, 2014) was performed using an LFQ minimum ratio count of 2.

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

FACT subunits in ES cells can be found in Supplementary Table 7. The complete list of differential protein expression between Control and *Ssrp1* KD can be found Supplementary Table 8.

NET-seq library preparation. 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. 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 3d** the Proximal Promoter region was defined as -30 bp and +250 bp around the TSS. For **Figure 3d,e**, 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/).

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

Biswas, D., Dutta-biswas, R., Mitra, D., Shibata, Y., Strahl, B. D., Formosa, T., & Stillman, D. J. (2006). Opposing roles for Set2 and yFACT in regulating TBP binding at promoters, *25*(19), 4479–4489. http://doi.org/10.1038/sj.emboj.7601333

Biswas, D., Yu, Y., Prall, M., Formosa, T., & Stillman, D. J. (2005). The Yeast FACT Complex Has a Role in Transcriptional Initiation †, *25*(14), 5812–5822. http://doi.org/10.1128/MCB.25.14.5812

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

Denissov, S., Hofemeister, H., Marks, H., Kranz, A., Ciotta, G., Singh, S., … Stewart, A. F. (2014). Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development*, *141*(3), 526 LP-537. Retrieved from http://dev.biologists.org/content/141/3/526.abstract

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

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

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

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

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

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

Janky, R., Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christiaens, V., … Aerts, S. (2014). iRegulon: From a Gene List to a Gene Regulatory Network Using Large Motif and Track Collections. *PLOS Computational Biology*, *10*(7), e1003731. Retrieved from https://doi.org/10.1371/journal.pcbi.1003731

Jonkers, I., Kwak, H., & Lis, J. T. (2014). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *eLife*, *3*(e02407), 1–25. http://doi.org/http://dx.doi.org/10.7554/eLife.02407

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

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

Katz, Y., Wang, E. T., Airoldi, E. M., & Burge, C. B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nature Methods*, *7*(12), 1009–

15. http://doi.org/10.1038/nmeth.1528

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

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

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

Kuckenberg, P., Kubaczka, C., & Schorle, H. (2012). The role of transcription factor Tcfap2c/TFAP2C in trophectoderm development. *Reproductive BioMedicine Online*, *25*(1), 12–20. http://doi.org/10.1016/j.rbmo.2012.02.015

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

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

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

Lolis, A. A., Londhe, P., Beggs, B. C., Byrum, S. D., Tackett, A. J., & Davie, J. K. (2013). Myogenin recruits the histone chaperone facilitates chromatin transcription (FACT) to promote nucleosome disassembly at muscle-specific genes. *Journal of Biological Chemistry*, *288*(11), 7676–7687. http://doi.org/10.1074/jbc.M112.426718

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

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

Min, I. M., Waterfall, J. J., Core, L. J., Munroe, R. J., Schimenti, J., Lis, J. T., … Lis, J. T. (2011). Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes & Development*, *25*(7), 742–754. http://doi.org/10.1101/gad.2005511

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

Qiao, Y., Zhu, Y., Sheng, N., Chen, J., Tao, R., Zhu, Q., … Jing, N. (2012). AP2γ regulates neural and epidermal development downstream of the BMP pathway at early stages of ectodermal patterning. *Cell Research*, *22*(11), 1546–1561. http://doi.org/10.1038/cr.2012.122

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

Ramirez, F., Ryan, D. P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., … Manke, T.

(2016). deepTools2: a next generation web server for deep-sequencing data analysis. *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

Schemmer, J., Araúzo-Bravo, M. J., Haas, N., Schäfer, S., Weber, S. N., Becker, A., … Schorle, H. (2013). Transcription Factor TFAP2C Regulates Major Programs Required for Murine Fetal Germ Cell Maintenance and Haploinsufficiency Predisposes to Teratomas in Male Mice. *PLoS ONE*, *8*(8), 1–15. http://doi.org/10.1371/journal.pone.0071113

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., 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

Shen, Y., Yue, F., McCleary, D. F., Ye, Z., Edsall, L., Kuan, S., … Ren, B. (2012). A map of the cis-regulatory sequences in the mouse genome. *Nature*, *488*(7409), 116–120. http://doi.org/10.1038/nature11243

Sims, R. J., Millhouse, S., Chen, C. F., Lewis, B. A., Erdjument-Bromage, H., Tempst, P., … Reinberg, D. (2007). Recognition of Trimethylated Histone H3 Lysine 4 Facilitates the Recruitment of Transcription Postinitiation Factors and Pre-mRNA Splicing. *Molecular Cell*, *28*(4), 665–676. http://doi.org/10.1016/j.molcel.2007.11.010

Teif, V. B., Vainshtein, Y., Caudron-Herger, M., Mallm, J.-P., Marth, C., Höfer, T., & Rippe, K. (2012). Genome-wide nucleosome positioning during embryonic stem cell development. *Nature Structural & Molecular Biology*, *19*(11), 1185–1192. http://doi.org/10.1038/nsmb.2419

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

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

Voong, L. N., Xi, L., Sebeson, A. C., Xiong, B., Wang, J. P., & Wang, X. (2016). Insights into Nucleosome Organization in Mouse Embryonic Stem Cells through Chemical Mapping. *Cell*, *167*(6), 1555–1570.e15. http://doi.org/10.1016/j.cell.2016.10.049

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

Xiao, S., Lu, J., Sridhar, B., Cao, X., Yu, P., Zhao, T., … Zhong, S. (2017). SMARCAD1 Contributes to the Regulation of Naive Pluripotency by Interacting with Histone Citrullination. *Cell Reports*, *18*(13), 3117–3128. http://doi.org/10.1016/j.celrep.2017.02.070

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