Published online 7 September 2020 *Nucleic Acids Research, 2020, Vol. 48, No. 18* **10211–10225**
doi: 10.1093/nar/gkaa732

Histone chaperone FACT represses retrotransposon MERVL and MERVL-derived cryptic promoters

Fuquan Chen[1,2,†], Weiyu Zhang[1,2,†], Dan Xie[1,2,†], Tingting Gao[1,2], Zhiqiang Dong[1,3] and Xinyi Lu[1,2,*]

[1]State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300350, People's Republic of China, [2]College of Pharmacy, Nankai University, Tianjin 300350, People's Republic of China and [3]College of Life Sciences, Nankai University, Tianjin 300307, People's Republic of China

Received February 03, 2020; Revised August 19, 2020; Editorial Decision August 20, 2020; Accepted August 24, 2020

ABSTRACT

Endogenous retroviruses (ERVs) were usually silenced by various histone modifications on histone H3 variants and respective histone chaperones in embryonic stem cells (ESCs). However, it is still unknown whether chaperones of other histones could repress ERVs. Here, we show that H2A/H2B histone chaperone FACT plays a critical role in silencing ERVs and ERV-derived cryptic promoters in ESCs. Loss of FACT component Ssrp1 activated MERVL whereas the re-introduction of Ssrp1 rescued the phenotype. Additionally, Ssrp1 interacted with MERVL and suppressed cryptic transcription of MERVL-fused genes. Remarkably, Ssrp1 interacted with and recruited H2B deubiquitinase Usp7 to Ssrp1 target genes. Suppression of *Usp7* caused similar phenotypes as loss of *Ssrp1*. Furthermore, Usp7 acted by deubiquitinating H2Bub and thereby repressed the expression of MERVL-fused genes. Taken together, our study uncovers a unique mechanism by which FACT complex silences ERVs and ERV-derived cryptic promoters in ESCs.

INTRODUCTION

Transposable elements (TEs) contribute to >40% of the mammalian genome (1). Approximately 20% of TEs in the mouse and human genome are endogenous retroviruses (ERVs) (1,2). During preimplantation development, ERVs are actively expressed across different stages (3–5). ERV III family members MERVL and MTA are highly expressed in totipotent mouse embryos and are gradually repressed since four-cell stage (3,4). In comparison, ERV II family members are activated from 8-cell stage until blastocyst (3). These ERVs are silenced timely before proceeding to the next developmental stage (6). They are usually silenced by various epigenetic routes such as heterochromatin associated histone marks (H3K9me3, H3K9me2 and DNA methylation) (7–9) and epigenetic modifiers in embryonic stem cells (ESCs) (10–16). These histone modifications occurred on specific histone variants deposited by histone chaperones (17,18). Histone variant H3.3 and corresponding chaperone ATRX/DAXX are essential to the maintenance of H3K9me3, which is critical to the silencing of Class I/II ERVs (19–23). Histone chaperone CAF-1 mediates the deposition of histone variant H3.1/H3.2 and can strengthen ERV repression via H3K9me3 and H4K20me3 by recruiting NuRD complex and Setdb1 (24–26). While most of these studies have been focusing on the role of histone H3 and corresponding chaperones in repressing ERVs, the contributions of other histones and corresponding chaperones to ERV regulation are less well understood.

Another important histone chaperone is FACT (facilitates chromatin transcription) complex, which is a well-characterized histone chaperone of H2A/H2B (27). It is associated with actively transcribed genes and promotes the displacement of H2A/H2B from chromatin to facilitate RNA polymerase II-mediated transcription (28). However, disruption of FACT expression only leads to modest changes in gene expression (29). Instead, recent evidence shows that FACT is important to preserve the chromatin structure (30). Hence, it is possible for FACT to act as a repressor of transcription. It was reported that FACT could inhibit the expression of cryptic transcripts, antisense transcripts, and subtelomeric genes (31–36). Moreover, a rich source of cryptic promoters originates from TEs including ERVs (8,37–38). Therefore, we proposed a hypothesis that FACT complex might suppress ERVs and ERV-driven cryptic transcription in ESCs.

MATERIALS AND METHODS

Cell culture

Mouse E14 and J1 ESCs were cultured on plates coated with 0.2% gelatin (G1890, Sigma) in Dulbecco's modified Eagle's

*To whom correspondence should be addressed. Tel: +86 22 8535 8316; Fax: +86 22 85358316; Email: luxy@nankai.edu.cn
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

medium (DMEM, Hyclone) supplemented with 15% fetal bovine serum (FBS, SH30070.03, Hyclone), 2 mM Glutamine (Gibco), 1% penicillin-streptomycin(P1400, Solarbio), 100 nM non-essential amino acids (Gibco), 0.1 mM β-mercaptoethanol (M3148–250, Sigma) and 10 ng/ml LIF (Z03077, GenScript) at 37°C under 5% CO_2. HEK 293T cells were cultured as previously described (3).

shRNA knockdown and inhibitor treatment

For gene knockdown, short hairpin RNAs (shRNAs) for Luciferase (control) or target genes were designed with the siRNA Selection Program (39). The shRNAs were cloned into a pSuper-puro plasmid. Targeting sequences of shRNAs were listed in Supplementary Table S1. ESCs were transfected with plasmids using Polyjet (SL100688, SignaGen), according to the manufacturer's protocol. Transfected ESCs were selected with 1 μg/ml puromycin from 24 h after transfection until harvested. For Usp7 inhibitor treatment, ESCs were treated with 10 μM P22077 (T2424, TargetMol) for 24 h.

Generation of *Ssrp1* knockout cell lines

The CRISPR/Cas9 system (Addgene #48138) was used following published protocols (40). Briefly, single guide RNAs (sgRNAs) targeting two distinct regions of the mouse *Ssrp1* gene were designed according to published protocols (40). ESCs were transfected with 2 μg *Ssrp1* gRNA using Polyjet transfection reagent (SL100688, SignaGene). GFP+ ESCs were sorted by flow cytometry and cultured for 6–8 days. Genomic DNA from the individual colony was obtained and validated by polymerase chain reaction (PCR) and mutation sites were sequenced by primers listed in Supplementary Table S1.

RNA extraction, reverse transcription and quantitative PCR (qPCR)

Total RNA was extracted from cells using RNAiso Plus (B9109, TAKARA BIO INC), and treated with DNase I for genomic DNA removal. cDNA synthesis was performed using a Transcriptor First Strand cDNA Synthesis Kit (4897030001, Roche), according to the manufacturer's instructions. Quantitative PCR (qPCR) was performed using Hieff qPCR SYBR Green Master Mix (H97410, Yeasen) and a QuantStudio 6 Real-Time PCR System (Life Technologies). Primer sequences for qPCR analysis are listed in Supplementary Table S1.

Establishment of overexpression ESC line

The full-length and truncations of *Ssrp1* were cloned by PCR and inserted into the pCAG-3HA vector. Ssrp1-ΔHMG mutant contained amino acid (a.a.) 1–480 of Ssrp1. Ssrp1 ΔSSrecog contained a.a. 285–708 of Ssrp1. Ssrp1 ΔRtt106 contained a.a. 1–284 and a.a. 481–708 of Ssrp1. The plasmids for transfection were purified by kit (PD1211, Biomiga). *Ssrp1*$^{-/-}$ ESCs were transfected with respective overexpression plasmids with Polyjet (SL100688, SignaGen), and ESCs were selected under 800 μg/ml hygromycin B for two weeks to obtain a stable cell line.

Western blot analysis

Cells were lysed in RIPA lysis buffer (R0010, Solarbio) with protease inhibitor PMSF (P0100, Solarbio), and protein concentration was measured using bicinchoninic acid (BCA) Protein Assay Kit (PC0020, Solarbio). The samples were loaded to an sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gel and transferred to polyvinylidene fluoride (PVDF) membrane (Millipore), which was blocked in 5% milk and probed with primary antibodies, and subsequent HRP–linked secondary antibodies (Santa Cruz). HRP activity was detected by Luminol HRP Substrate (WBKLS0100, Millipore). Primary antibodies used are anti-Flag (F1804, Sigma), anti-HA (sc7392, Santa Cruz), anti-Gapdh (KM9002, Sungene), anti-H3 (17168–1-AP, Proteintech), anti-Ssrp1 (sc-74536, Santa Cruz) and anti-Supt16 (#12191, Cell Signaling Technology). Secondary antibodies used are goat anti-rabbit IgG-HRP (sc-2004, Santa Cruz) and anti-mouse IgG-HRP (sc-516102, Santa Cruz).

RNA-seq and data analysis

Total RNA was extracted from wild-type (WT) ESCs and *Ssrp1*$^{-/-}$ ESCs in RNAiso Plus (B9109, TAKARA BIO INC) and DNA was removed with DNase I (Thermo Scientific). Total RNA (4 μg) was used for RNA-seq library preparation. Each library was prepared and sequenced to obtain at least 6 Gb data by GENEWIZ Inc. China. For strand-specific RNA-seq, 1 μg total RNA with RIN value above nine was used for rRNA depletion with mouse Ribo-Zero rRNA Removal Kit (Illumina) and subsequent library preparation. Cutadapt was used to remove adapter sequences and low-quality 3′ end sequences. Hisat2 (41) was used to map reads to mouse mm10 genome assembly and genes were annotated according to the Ensembl database (42). The parameter '–rna-strandness RF' was used for mapping strand-specific RNA-seq reads. The number of reads mapped to genes was determined with featureCounts, and the parameter '-s 2' was added for counting strand-specific RNA-seq reads. The number of reads mapped to TE repNames and loci was generated by squire. Genes and TEs with mean count number <5 were filtered out. DESeq2 (43) was used to analyzed differentially expressed genes, TE repNames and TE locus according to negative binomial generalized linear models. Genes, TE repNames and TE loci with expression fold change > 1.5 and adjusted *P*-value < 0.05 were considered to be with significant expression change. R package RDAVID Web Service (44) was used for gene ontology (GO) and KEGG analysis. Gene set enrichment analysis (GSEA) was done by gseapy. For analysis of MERVL-fusion transcripts, stringtie (45) was used to assemble RNA-Seq reads into potential transcripts. If the exons of assembled transcripts overlapped with MERVL, they were considered MERVL-fusion transcripts.

Chromatin immunoprecipitation

ChIP was performed essentially as described previously (3). Briefly, ESCs were cross-linked by 1% formaldehyde for 10 min at room temperature and quenched by 200 mM glycine for 5 min. Soluble chromatin was obtained after cell lysis

and sonication. Then samples were pre-cleared and incubated with 3–5 μg antibody loaded protein G MagBeads (L00274, GenScript) at 4°C overnight. Subsequently, the immunoprecipitated DNA was decrosslinked, purified and analyzed by qPCR. ChIP-qPCR was done with anti-Ssrp1 antibody (ab137034, Abcam), anti-USP7 antibody (A300–033A, Bethyl) or anti-H2Bub antibody (#5546, Cell Signaling Technology).

ChIP-seq and data analysis

ChIP using anti-Ssrp1 (ab137034, Abcam) or anti-USP7 (A300–033A, Bethyl) was performed as described above. ChIP DNA (2 ng) was used for the ChIP-seq library which was prepared by Novogene Corporation. Each library was sequenced by Novogene Corporation, resulting in pair-ended 150 bp reads with ChIP input as control. The adapter sequences and low-quality 3′ ends of reads were removed with Cutadapt and mapped to the mouse mm10 genome assembly using Bowtie2 (46). The ChIP-seq correlation coefficient was determined by Deeptools (47). Due to the close correlation between Ssrp1 ChIP duplicate, we merged the two Ssrp1 duplicates for downstream analysis. Enrichment of ChIP-seq signal was also generated by Deeptools. Ssrp1 binding peaks were predicted by Macs2 (48). MERVL-int center information was inferred from RepeatMasker annotation file from UCSC genome browser (49).

Immunostaining and flow cytometry analysis

Flow cytometry analysis after immunostaining was done according to a published protocol (3). In brief, the ESCs were harvested and fixed in pre-cooled 70% ethanol. Cells were stained with the MERVL-gag antibody after permeabilization and blocking. The secondary antibody used was Alexa Fluor 594 goat anti-rabbit IgG. Cells were analyzed by flow cytometry (BD LSRFortessa).

Protein co-immunoprecipitation (co-IP) and mass spectrometry

E14 ESCs expressing 3HA-Ssrp1 and control E14 ESCs expressing pCAG-3HA were used were washed with pre-cooled PBS and lysed in 1 ml pre-cooled lysis buffer (20 mM Tris [pH = 7.5], 150 mM NaCl, 0.5% Triton X-100, 20 mM KCl, 1% Glycine, 1.5 mM $MgCl_2$ and protease inhibitor cocktail). The cell lysate was pre-cleared and immunoprecipitated with anti-HA magnetic beads (b26202, Bimake). Beads were washed three times with lysis buffer and twice with wash buffer (50 mM Tris [pH = 8.0], 300 mM NaCl, 1% Triton X-100, 1 mM ethylenediaminetetraacetic acid (EDTA)). Proteins were eluted in elution buffer (50 mM Tris [pH = 7.5], 1% SDS, 10 mM EDTA), and loaded to 10% SDS-PAGE for silver-staining and mass spectrometry. For mass spectrometry, proteins were digested, labeled and loaded to EASY-nLC 1200 UHPLC system coupled with a Q Exactive HF-X mass spectrometer (Thermo Scientific). The resulting spectra from each fraction were searched by Proteome Discoverer 2.2 (PD 2.2, Thermo Scientific). E14 ESCs expressing Flag-Usp7 and control pLCH72 -Flag was used for Usp7/Ssrp1 co-IP with anti-Flag magnetic beads (b26102, Bimake). For co-IP of Ssrp1 mutants and Usp7, 293T cells were co-transfected with Usp7-Flag/pCAG-3HA, Usp7-Flag/3HA-Ssrp1 ΔRtt106, Usp7-Flag/3HA-Ssrp1-ΔSSrecog or Usp7-Flag/3HA-Ssrp1-HMG. For co-IP of Supt16 and Usp7, Usp7 was immunoprecipitated with Usp7 antibody (A300–033A, Bethyl) whereas normal rabbit IgG (sc-2027, Santa Cruz Biotechnology) was used as control. *Ssrp1*$^{-/-}$ ESC overexpressing Supt16-Flag or control vector were used for Supt16/Usp7 co-IP with anti-Flag magnetic beads (b26102,Bimake). Proteins were loaded to 10% or 12% SDS-PAGE together with 1.25% input for western blot. For analysis of mass spectrometry results, unique peptides were used for the determination of protein enrichment in corresponding samples. Cpm in E14 ESCs was used as the expression level of the corresponding genes. Proteins in GO:0005856 (cytoskeleton) were removed from the protein list before analysis.

Luciferase assay

WT and *Ssrp1*$^{-/-}$ ESCs were transfected with pGL4.23 vector containing MERVL or empty vector. A total of 200 ng of the reporter construct was transfected into each well of E14 ESC in a 24-well plate together with 10 ng of pCMV-Renilla. After transfection, the cells were cultured for 38 h. The medium was changed 1 day after transfection. Luciferase and Renilla activity was determined by the Dual-Luciferase Reporter Assay System (#E1910, Promega) according to the manufacturer's instructions.

Statistical analysis

qPCR results were analyzed by the Student's *t*-test and overlapping enrichment significance was calculated by Fisher's exact test. Significant differences were defined as $*P < 0.05$, $**P < 0.01$ or $***$ $P < 0.001$. For the determination of differentially expressed genes and TEs, the Wald test was used. For GSEA, the Kolmogorov-Smirnov statistic was used.

RESULTS

FACT complex modulates ERVs in ESCs

FACT complex consists of Ssrp1 and Supt16. To test our hypothesis that FACT represses ERVs, we depleted *Ssrp1* and *Supt16* separately with two shRNAs for each gene in E14 ESCs (Figure 1A and B). The depletion of either *Ssrp1* or *Supt16* induced the activation of MERVL and RLTR1B expression while the expression of IAPEz, LINE1 and SINE B1 was not affected (Figure 1A and B), suggesting that both components of FACT complex are critical to ERV repression. To further prove the function of the FACT complex, we designed gRNAs targeting exon 3 of *Ssrp1* gene (Supplementary Figure S1A). We established two *Ssrp1* knockout ESC lines with 37 or 160 bp deletion in exon 3, which induced frameshift mutations and premature termination of Ssrp1 protein at 35 or 43 amino acids (a.a.) during translation (Figure 1C; Supplementary Figure S1B and C). In these two knockout cell lines, *Ssrp1* was not detected by qPCR with primers targeting deleted regions (Figure 1D). Our *Ssrp1*$^{-/-}$ ESCs completely lost Ssrp1 protein (Figure

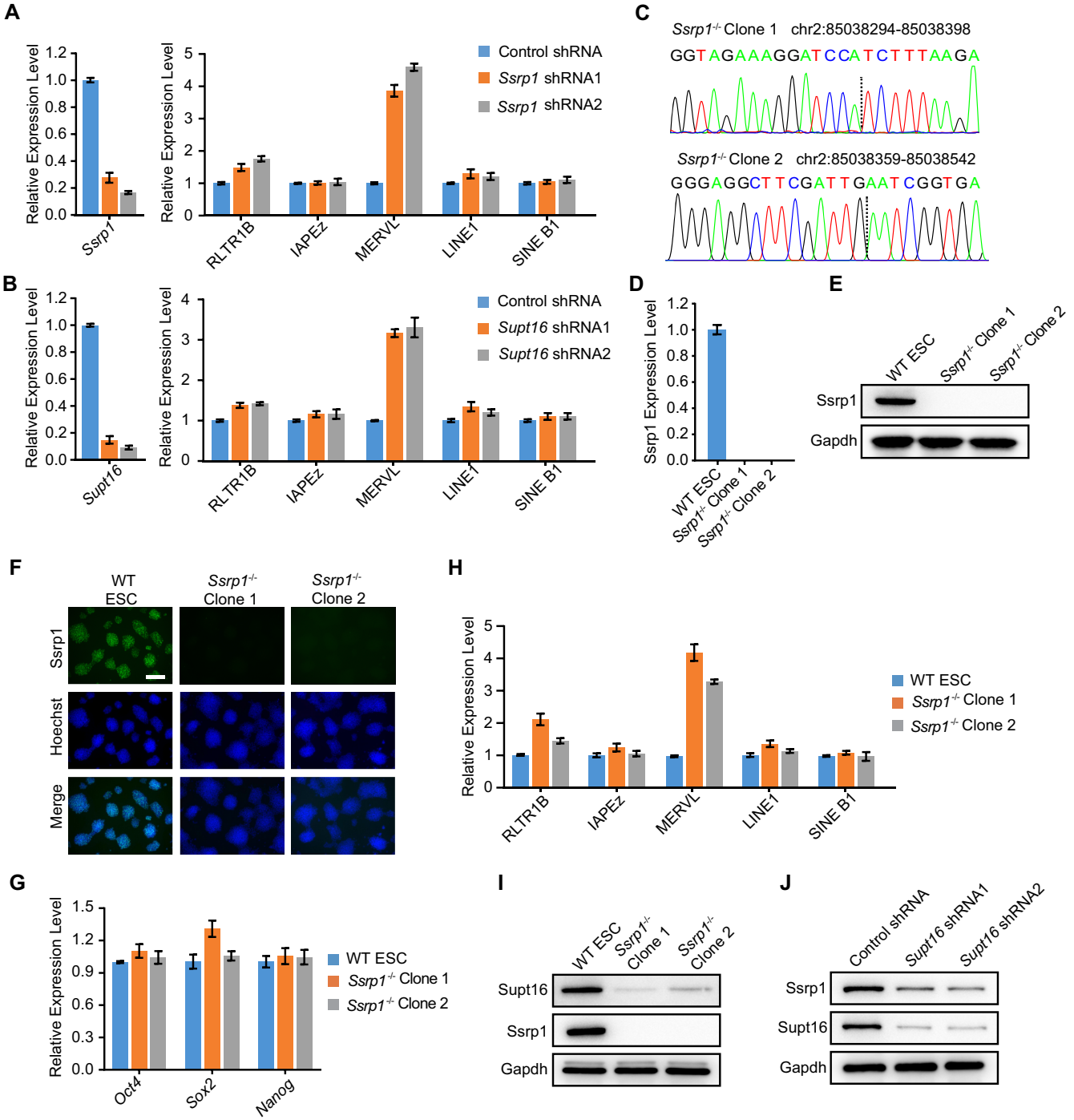

Figure 1. FACT complex represses ERVs. (**A**) qPCR analysis of the expression of *Ssrp1* and retrotransposons after *Ssrp1* depletion in E14 ESCs. Data represent mean ± s.e.m., $n = 3$ biological replicates. (**B**) The expression levels of *Supt16* and retrotransposons after *Supt16* depletion in E14 ESCs, as measured by qPCR and normalized to *Gapdh* levels. Biological triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. (**C**) DNA sequencing results of mutation sites in two *Ssrp1*$^{-/-}$ ESC lines. (**D**) RT–qPCR analysis of the expression of *Ssrp1* using specific primers located near the sgRNA sites in WT and *Ssrp1*$^{-/-}$ ESCs. Biological triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. (**E**) Western blot analysis of Ssrp1 protein in WT and *Ssrp1*$^{-/-}$ ESCs. Gapdh was used as a loading control. (**F**) Immunofluorescence staining of Ssrp1 protein in WT ESC and *Ssrp1*$^{-/-}$ ESCs. Ssrp1 was stained in green. DNA was stained by Hoechst 33342. Scale bar, 50 μm. (**G**) qPCR analysis of pluripotent genes (*Oct4*, *Sox2* and *Nanog*) expression in WT ESCs and *Ssrp1*$^{-/-}$ ESCs. (**H**) qPCR analysis of expression of *Ssrp1* and retrotransposons in WT ESCs and *Ssrp1*$^{-/-}$ ESCs. The results were presented as mean ± s.e.m from three biological replicates. (**I**) Western blot analysis of Ssrp1, Supt16 and Gapdh protein levels in WT ESCs and *Ssrp1*$^{-/-}$ ESCs. (**J**) Western blot analysis of Supt16 and Ssrp1 protein levels in ESCs expressing *Supt16* shRNA or control shRNA. Gapdh was included as a loading control.

1E and F) without demonstrating changes in the expression of pluripotency genes (Oct4, Sox2 and Nanog) (Figure 1G) and cell morphology (Supplementary Figure S1D). Knockout of *Ssrp1* similarly induced upregulation of ERVs including RLTR1B and MERVL (Figure 1H), in support of the results from depletion of *Supt16* or *Ssrp1* in ESCs. Also, loss of Ssrp1 protein coincided with the reduction of Supt16 protein level in ESCs (Figure 1I), which was previously observed in other cell types upon *Ssrp1* depletion (50). Likewise, the depletion of *Supt16* destabilized Ssrp1 protein (Figure 1J). Hence, the knockout of *Ssrp1* alone impaired the function of the FACT complex. To exclude the possibility that the upregulation of MERVL is a cell type-specific phenotype after *Ssrp1* loss, we depleted *Ssrp1* and *Supt16* in another ESC line (J1). In agreement with results in E14 ESCs, depletion of either *Ssrp1* or *Supt16* in J1 ESCs caused the upregulation of MERVL (Supplementary Figure S1E and F). Together, these results demonstrate that FACT complex members are critical to the proper repression of ERVs such as MERVL in ESCs.

Overexpression of *Ssrp1* rescues MERVL expression

Since the *Ssrp1* knockout led to the reduction of both Ssrp1 and Supt16 proteins, we next examined which subunit of FACT complex was important to the regulation of MERVL. We overexpressed *Ssrp1* or *Supt16* independently in *Ssrp1*$^{-/-}$ ESCs (Figure 2A and B). Overexpression of Ssrp1 alone almost completely restored MERVL expression (Figure 2C). In contrast, Supt16 overexpression only modestly rescued the phenotype of *Ssrp1* knockout (Figure 2D). This could be because that Ssrp1 is essential to the stability of Supt16 protein, but *Ssrp1* had been deleted and could not be rescued by Supt16 overexpression. Therefore, both components of the FACT complex are central to the suppression of MERVL in ESCs.

To test which domain of Ssrp1 is important for its function, we generated Ssrp1 mutants with individual domain deleted (Figure 2E). We overexpressed these Ssrp1 mutants in *Ssrp1*$^{-/-}$ ESCs to rescue the expression of MERVL (Supplementary Figure S2A and B). We found that Ssrp1 mutant protein without the HMG domain completely abolished the function of Ssrp1 in repressing MERVL (Figure 2F). Ssrp1 is able to bind to DNA (51) and the HMG domain of Ssrp1 is known to function as a DNA binding domain (52). Thus, the interaction of Ssrp1 with DNA should be necessary for Ssrp1 to repress MERVL. The deletion of either the SSrecog domain or Rtt106 domain of Ssrp1 protein partially impaired the rescue efficiency of MERVL (Figure 2G and H), suggesting the participation of these domains in ERV repression. In sum, these results indicate that the integrities of Ssrp1 protein and FACT complex are required for FACT to repress MERVL.

Genome-wide suppression of genes and ERVs by FACT complex

To further investigate whether Ssrp1 represses other genes or TEs in ESCs, we analyzed the transcriptome of *Ssrp1*$^{-/-}$ ESCs by RNA-seq. Surprisingly, more genes were upregulated (895 genes) than downregulated (213 genes) after the loss of *Ssrp1* (Figure 3A and Supplementary Table S2), although Ssrp1 is traditionally associated with actively transcribed genes (53). Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of Ssrp1 target genes identified downregulation of various cellular metabolic pathways related to the biosynthesis of amino acids, antibiotics and glycolysis whereas upregulated KEGG terms were enriched of pathways related to cell cycle, p53 signaling, viral infection and ubiquitination (Figure 3B and C). GO analysis revealed biological processes similar to KEGG analysis, including cell cycle and protein ubiquitination for upregulated genes and glycolytic process for downregulated genes (Supplementary Figure S3A and B). These data prove the FACT complex as a regulator of cellular metabolism, cell cycle and protein ubiquitination in ESCs.

Interestingly, loss of *Ssrp1* induced the upregulation of other TEs in ESCs (Figure 3D). Among the upregulated TEs, MERVL (MT2-Mm and MERVL-int) loci were the most strongly activated (Figure 3D and E; Supplementary Table S3). The number of activated MERVL loci was the highest among all upregulated TEs as well (Figure 3E), substantiating that the expression upregulation of MERVL is a genome-wide phenomenon rather than the consequence of activation of a particular MERVL locus. In contrast, downregulated TEs did not show a genome-wide expression reduction of different loci (Figure 3F). To inspect how Ssrp1 regulates the expression of genes and TEs, we performed Ssrp1 ChIP-seq. We found that promoter regions and gene body (exons and introns) accommodated most Ssrp1 binding sites (17.2% for promoter, 18.9% for exons and 32.1% for introns) (Figure 3G and Supplementary Figure S3C). Importantly, Ssrp1 binding was well enriched on MERVL (Figure 3H). This was confirmed by our ChIP-qPCR results (Figure 3I). Ssrp1 interacted with MERVL but not other TEs we examined (Figure 3H and I), suggesting that Ssrp1 directly represses MERVL expression. Altogether, these analyses show that Ssrp1 is able to repress the genes and TEs in ESCs.

Ssrp1 represses MERVL and MERVL-driven fusion transcripts

Since MERVL is highly expressed in 2-cell (2C) embryos (4) and activated after *Ssrp1* knockout (Figure 3D and E), we asked whether 2C embryo genes were activated at the same time. Indeed, 2C genes were enriched in upregulated transcriptome after *Ssrp1* knockout (Figure 4A). This is consistent with the fact that loss of *Ssrp1* similarly led to an increased percentage of MERVL+ 2-cell like population (Figure 4B). Next, we asked how the 2C genes were activated. It was found that ERVs fused with mRNAs to become chimeric transcripts (4,54–56). MERVL (MT2_Mm) was also able to act as a promoter (Supplementary Figure S4A) and drove fused transcripts in 2C embryos and 2C-like cells (4). Notably, after the loss of *Ssrp1*, a significant number of upregulated transcripts overlapped with TEs (Fisher's exact test, $P = 2.87e\text{-}7$) (Figure 4C), but no overlap was found for downregulated transcripts (data not shown), implying a number of transcripts gained TE-fused exons. The majority (59.3%) of TE-derived upregulated transcripts are fused with MERVL (MT2-Mm and

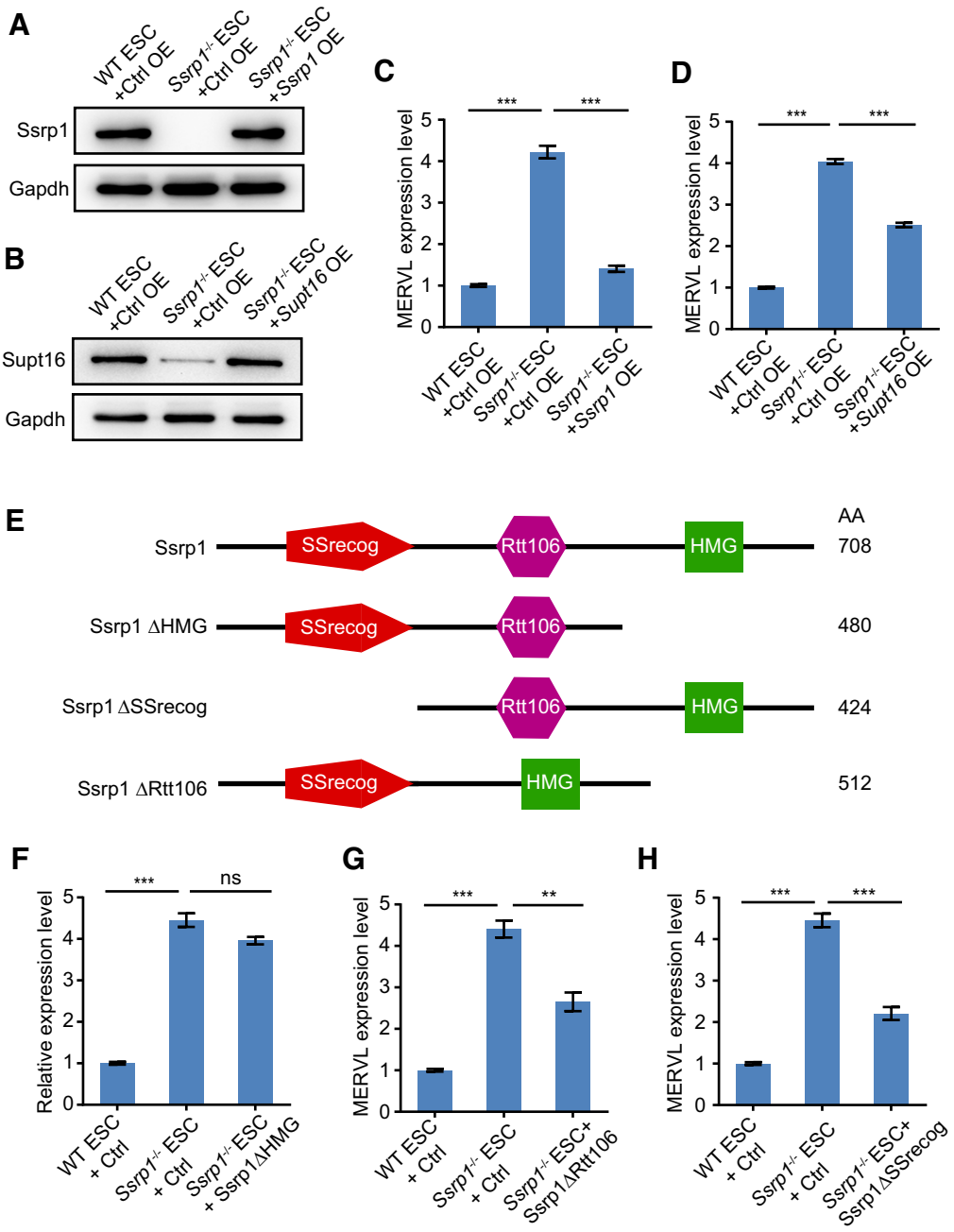

Figure 2. Restoration of MERVL expression by the introduction of Ssrp1 or Supt16. (**A** and **B**) Immunoblot analysis of the expression of Ssrp1 (A) or Supt16 (B) after overexpression of *Ssrp1* or *Supt16* in *Ssrp1*$^{-/-}$ ESCs. Gapdh was included as a loading control. OE, overexpression. Ctrl, control vector. (**C**) qPCR analysis of MERVL expression in *Ssrp1*$^{-/-}$ ESC after overexpression of *Ssrp1*. Biological triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. Significant differences were determined by Student's *t*-test and defined as ***$P < 0.001$. (**D**) qPCR analysis of MERVL expression in *Ssrp1*$^{-/-}$ ESC after overexpression of *Supt16*. Ctrl, control vector overexpression. Biological triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. Significant differences were determined by Student's *t*-test and defined as ***$P < 0.001$. ns, non-significant. (**E**) A schematic summary of Ssrp1 mutants used for functional rescue. The length of each mutant form is indicated at the right in amino acids (AA). Δ, deletion. (**F**–**H**) qPCR analysis of MERVL expression after overexpression of Ssrp1 ΔHMG (F), Ssrp1 ΔRtt106 (G), or Ssrp1ΔSSrecog (H) in *Ssrp1*$^{-/-}$ ESCs. qPCR results are normalized to Gapdh. Biological triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. Ctrl, control vector overexpression. Significant differences were determined by Student's *t*-test and defined as **$P < 0.01$ or *** $P < 0.001$.

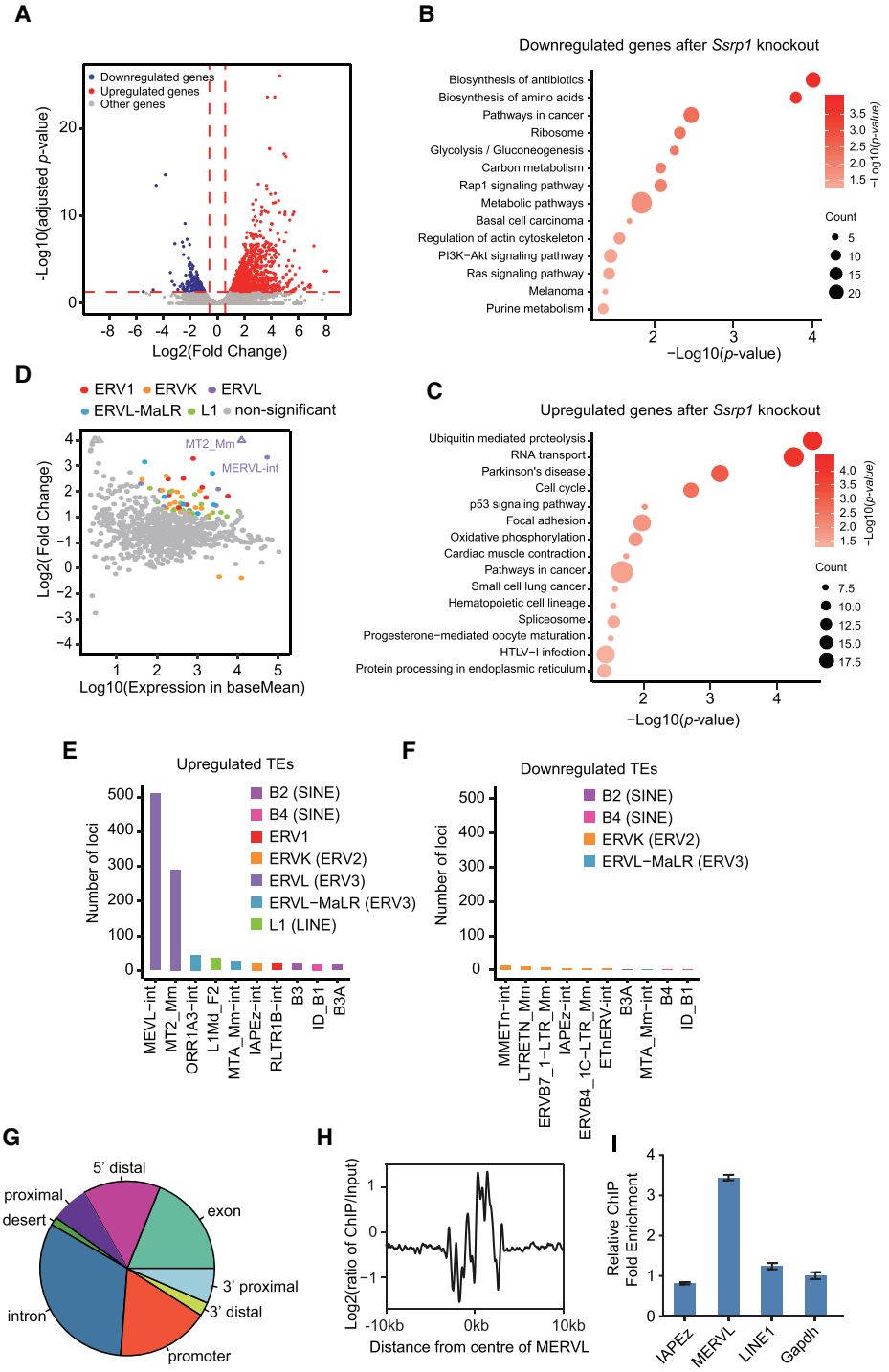

Figure 3. FACT complex genome-wide regulates TE transcription. (**A**) The volcano plot of gene expression in *Ssrp1*$^{-/-}$ ESCs versus WT ESCs. Significantly upregulated genes were labeled in red and significantly downregulated genes were labeled in blue. Horizontal red dash line marked adjusted *P*-value (Wald test) 0.05 and vertical lines marked expression fold change 1.5. (**B** and **C**) KEGG analysis of pathways related to downregulated genes (B) and upregulated genes (C) after *Ssrp1* knockout in ESCs. The analysis was done in DAVID. Color gradient indicated significance in −log10 (*P*-value) and dot size indicated the number of genes in the corresponding pathway. (**D**) A scatter diagram shows a transcriptome analysis of TE expression after *Ssrp1* knockout. The result from Squire was used to plot the diagram. Colored dots indicate TE with significant expression change ($P < 0.05$, Wald test). Triangles represent TEs with log2 (fold change) > 4. (**E** and **F**) The top 10 TEs with the highest number of loci upregulated (E) or downregulated (F) after *Ssrp1* loss. The subfamily type of each TE was labeled in brackets. (**G**) Locations of Ssrp1 peaks relative to the nearest transcription units (Promoter, 2 kb around transcriptional start sites; 5' proximal, 2–10 kb upstream of the gene; 5' distal, 10–100 kb upstream of the gene; 3' proximal, 0–10 kb downstream of the gene; 3' distal, 10–100 kb downstream of the gene; Gene desert, >100 kb away from the nearest gene). (**H**) Ssrp1 binding profile around the center of MERVL locus. The ChIP-seq signal was calculated as the log2 ratio of the normalized number of reads relative to the input. (**I**) ChIP-qPCR analysis of Ssrp1 binding on different retrotransposons. ChIP-qPCR data were normalized to input and Gapdh. Biological triplicate data ($n = 3$ extracts) are presented as mean ± s.e.m.

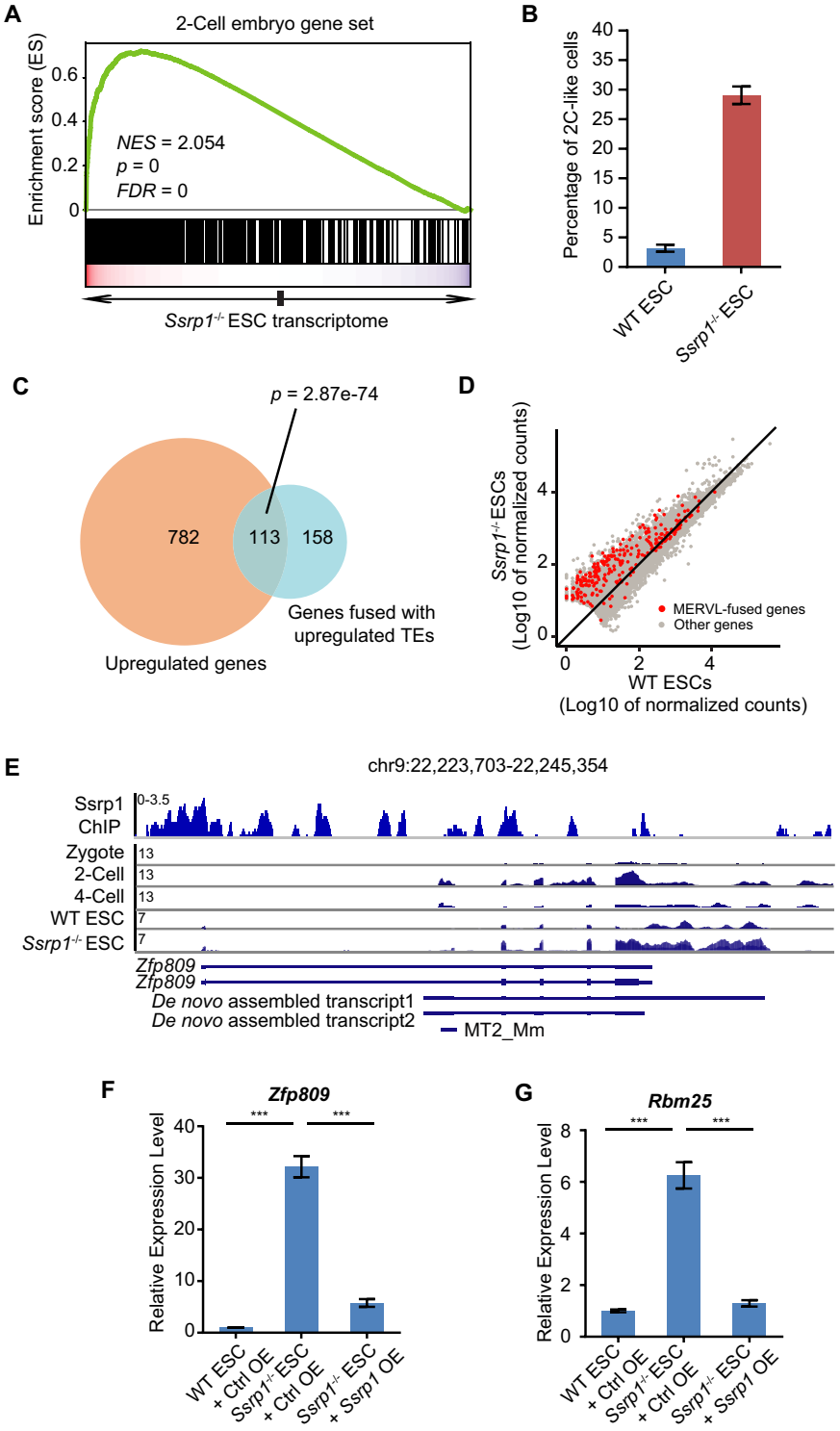

Figure 4. FACT represses MERVL-driven fusion transcripts. (**A**) GSEA analysis of upregulated genes after *Ssrp1* knockout for the enrichment of 2C genes. Red, upregulated genes; blue, downregulated genes; NES, normalized enrichment scores; FDR, false discovery rate. The Kolmogorov–Smirnov statistic was used for calculation of *P*-value. (**B**) Flow cytometry analysis of the MERVL-gag+ population within WT ESCs or *Ssrp1*$^{-/-}$ ESCs. Biological triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. (**C**) The overlap between genes whose TSS are overlapped with upregulated TEs and upregulated genes in *Ssrp1*$^{-/-}$ ESCs. The *P*-value was calculated by Fisher's exact test. (**D**) Dot plot of all expressed genes in WT and *Ssrp1*$^{-/-}$ ESCs. Genes with alternative transcript(s) overlapped with MERVL were labeled in red. (**E**) ChIP-seq signal of Ssrp1 binding on *Zfp809* in WT ESC and RNA-seq signal across embryo of the zygote, 2-cell and 4-cell, WT and *Ssrp1*$^{-/-}$ ESCs. The signal of multiple samples was merged. Genes and MERVL tracks were indicated below. (**F** and **G**) RT-qPCR analysis of expression of representative genes *Zfp809* (F) and *Rbm25* (G) in WT ESCs, *Ssrp1*$^{-/-}$ ESCs and *Ssrp1*$^{-/-}$ ESCs overexpressing *Ssrp1* or ctrl vector. Biological triplicate data ($n = 3$ extracts) are presented as mean ± s.e.m. Significant differences were determined by Student's *t*-test and defined as *** $P < 0.001$.

MERVL-int) (Supplementary Figure S4B). It is noteworthy that most MERVL-fused transcripts (210 out of 249) were activated after the loss of *Ssrp1* (Figure 4D). This is exemplified by 2C gene *Zfp809*, which was bound by Ssrp1 and upregulated after *Ssrp1* knockout (Figure 4E). We further validated the activation of MERVL-fused transcripts (*Zfp809* and *Rbm25*) by qPCR after the loss of *Ssrp1* (Figure 4F and G; Supplementary Figure S4C). These transcripts were again repressed after the re-acquisition of Ssrp1 expression in *Ssrp1*$^{-/-}$ ESCs (Figure 4F and G). Overall, these results suggest that Ssrp1 represses MERVL-fused transcripts.

To study whether FACT complex also represses cryptic transcripts without MERVL fusion, we overlapped genes with alternative transcription start site (TSS) and upregulated genes. We found that 728 out of 895 upregulated genes were transcribed from alternative TSS after the loss of *Ssrp1* (Supplementary Table S4). Cryptic transcripts were more strongly activated than wildtype transcripts from these genes (Supplementary Figure S4D), implying cryptic transcription contributed to the upregulation of these genes. Intriguingly, transcriptional activation of cryptic transcripts fused with TEs was stronger than other cryptic transcripts (Supplementary Figure S4E). Additionally, the activation level of the first exons of genes with alternative TSS was lower than that of other upregulated genes (Supplementary Figure S4F). The median of expression fold change for first exons of genes with alternative TSS was 1.09, which is lower than 1.5-fold upregulation cutoff for upregulated genes and close to no expression change (fold change = 1). In addition, the expression fold changes of first exons were lower than those of the corresponding full-length genes (Supplementary Figure S4G). These data suggest that Ssrp1 also represses the cryptic transcription besides those driven by MERVL and those cryptic transcription events tend to be initiated beyond the first exons.

Besides cryptic transcription, disruption of FACT function activates antisense transcription in yeasts (31,34). Hence, we asked whether antisense transcription was activated in *Ssrp1*$^{-/-}$ ESCs. Our strand-specific RNA-seq showed that 78 antisense transcripts were upregulated while 65 of them were downregulated (Supplementary Figure S4H). Among the upregulated antisense transcripts, 64 of them fused with TEs, of which 15 contained MERVL and 8 contained RLTR1B (Supplementary Figure S4I). These results suggest that mouse FACT also represses the expression of antisense transcripts, especially those fused with TEs.

Ssrp1 acts through Usp7 to repress MERVL-fused transcripts

We hypothesized that FACT complex repressed MERVL by working with epigenetic regulators. To investigate how Ssrp1 exerts its repression role through epigenetic regulation, we performed Ssrp1 co-immunoprecipitation following by mass spectrometry analysis. As expected, the top protein that interacted with Ssrp1 was Supt16, approving the results of our proteomic analysis (Figure 5A; and Supplementary Table S5). Of note, Ssrp1 strongly interacted with Usp7 (Figure 5A), which is known as an H2B deubiquitinase (57–61). In addition, Ssrp1 binding profile was positively correlated with that of Usp7, but not other known repressors of MERVL (Chaf1a, Lsd1, Hp1g, Sumo2, H3K9me3 and H3K9me2) (Figure 5B). Our co-immunoprecipitation results confirmed the direct interaction between Ssrp1 and Usp7 in both ESCs and 293T cells (Figure 5C and D; Supplementary Figure S5A and B). This interaction was mediated by both the SSrecog domain and Rtt106 domain, but not HMG-box domain of Ssrp1 (Supplementary Figure S5C–E). Importantly, depletion of *Usp7* by shRNA treatment or inhibition of Usp7 deubiquitinase activity resulted in the upregulation of MERVL (Figure 5E and F; Supplementary Figure S5F–H and Table S6). What is more, *Usp7* depletion activated 2-cell genes (Supplementary Figure S5I and Table S7). More importantly, RNA-seq analysis after *Usp7* depletion revealed a similar change of KEGG pathways in comparison to *Ssrp1* knockout (Figure 3B and C; Supplementary Figure S5J and K). Upregulated genes in *Ssrp1*$^{-/-}$ ESCs were enriched in upregulated transcriptome after *Usp7* depletion (Figure 5G). Furthermore, we found enrichment of Usp7 binding on Ssrp1 peaks (Figure 5H), and this enrichment was decreased after *Ssrp1* loss (Figure 5H). This was similarly observed on MERVL. Less Usp7 became associated with MERVL after *Ssrp1* loss (Figure 5I and J). Supt16 also interacted with Usp7 (Supplementary Figure S5L and M), however, overexpression of Supt16 alone could not rescue the recruitment of Usp7 to MERVL after the loss of *Ssrp1* (Figure 5K), suggesting that an intact FACT complex is required for Usp7 to interact with MERVL. Altogether, these results support the mechanism that FACT represses MERVL by recruiting Usp7.

Usp7 represses MERVL-fused transcripts through controlling H2Bub deposition

Since Usp7 was recruited to MERVL by Ssrp1, we examined the role of Usp7 in regulating the production of MERVL-fused transcripts. We found that the depletion of *Usp7* resulted in the upregulation of MERVL-fused transcripts (Figure 6A). Usp7 is known to de-ubiquitinate histone H2B (57–61). Thus, we speculated that the reduced association of Usp7 with MERVL after *Ssrp1* loss might lead to the gain of H2Bub. The inhibition of Usp7 activity caused an increment of H2Bub level in ESCs (Supplementary Figure S6A). The loss of Ssrp1 influenced the interaction of Usp7 with MERVL (Figure 5I). This was accompanied by the enhanced enrichment of H2Bub on MERVL-fused genes and MERVL consensus in *Ssrp1*$^{-/-}$ ESCs (Figure 6B). Similar to *Ssrp1* depletion, loss of *Usp7* (60) led to increased H2Bub association with MERVL and MERVL-fused genes (Supplementary Figure S6B and C). In support of our finding, depletion of H2B ubiquitinase Rnf20 with two independent shRNAs partially rescued MERVL expression and decreased fusion transcript expression (Figure 6C and D). Thus, we conclude that the dissociation of Usp7 activates MERVL-fused transcripts by allowing H2Bub deposition.

DISCUSSION

In summary, we propose a model that Ssrp1 recruits Usp7 to remove H2Bub and represses transcription from MERVL

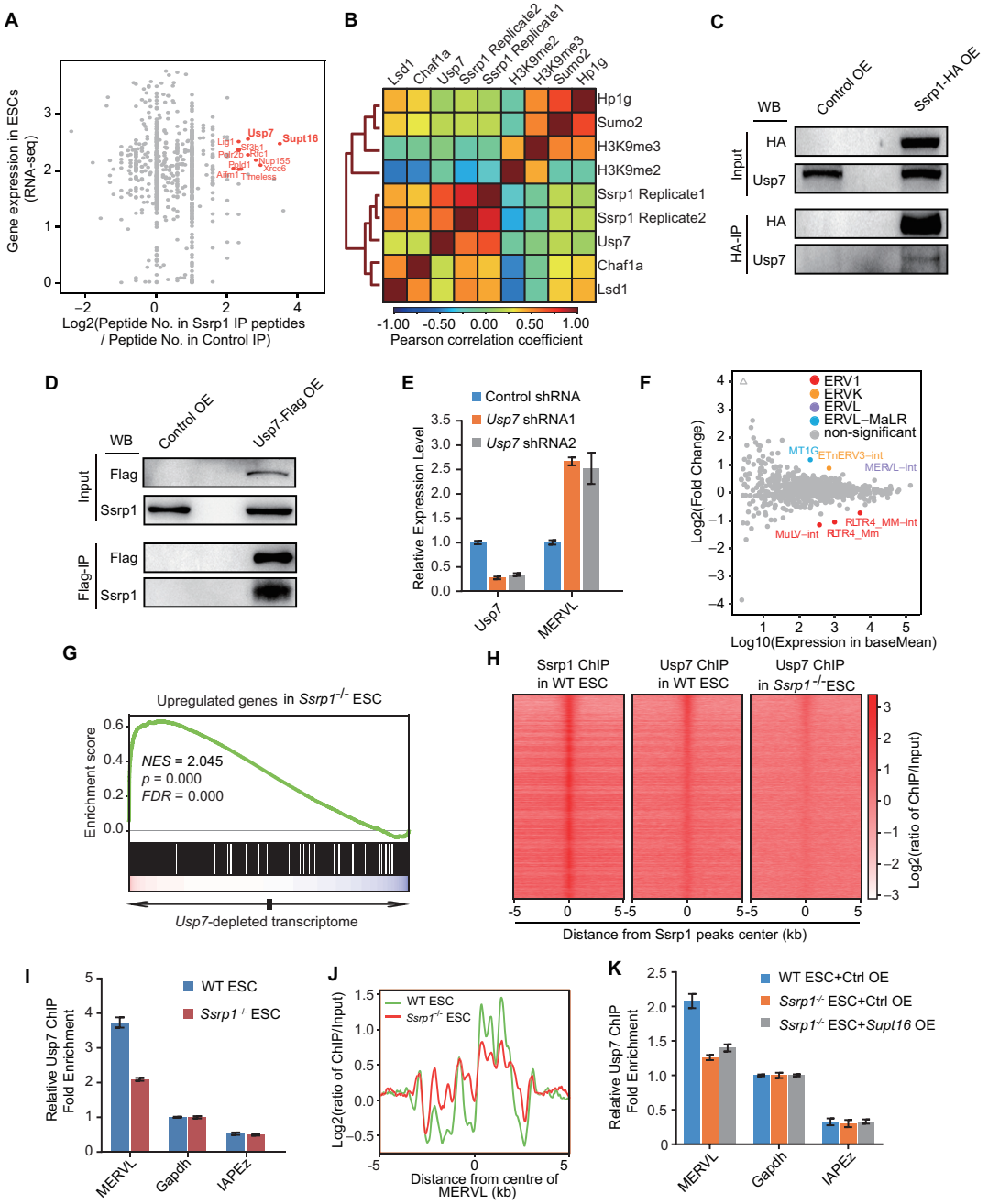

Figure 5. Ssrp1 recruits Usp7 to repress MERVL expression. (**A**) Mass spectrometry analysis of Ssrp1-associated proteins after Ssrp1 co-IP. *Y*-axis indicates the transcriptional level of corresponding proteins and the *X*-axis indicates the binding strength of proteins to Ssrp1. (**B**) Correlation heatmap of Ssrp1 duplicate binding profile and Usp7 binding profile together with known histone marks/regulators of ERVs. Pearson's correlation coefficient was used to estimate the strength of the correlation. The ChIP-seq signal was calculated as the log2 ratio of normalized reads relative to the input. (**C**) Western blot analysis of Ssrp1-HA/Usp7 co-immunoprecipitation in ESCs overexpressing control empty vector or HA-tagged Ssrp1. IP was done with anti-HA magnetic beads. 1.25% input was loaded as control. (**D**) Western blot analysis of Usp7-Flag/Ssrp1 co-immunoprecipitation in ESCs overexpressing control empty vector or Flag-tagged Usp7. IP was done with anti-Flag magnetic beads. 1.25% input was loaded as control. IP: immunoprecipitation; OE: overexpression. (**E**) RT-qPCR analysis of the expression of *Usp7* and MERVL after *Usp7* depletion in ESCs. Data are shown as mean ± s.e.m. ($n = 3$). (**F**) Scatter plot of TE expression after *Usp7* depletion. Squire results were used to plot the diagram. Colored dots indicate TEs with significant expression change ($P < 0.05$, Wald test). Triangles represent TEs with log2 (fold change) > 4. (**G**) GSEA analysis of enrichment of genes repressed by Ssrp1 in the upregulated transcriptome of ESCs with *Usp7* depleted. Red, upregulated genes; blue, downregulated genes. NES: normalized enrichment scores; FDR: false discovery rate. The Kolmogorov–Smirnov statistic was used for calculation of *P*-value. (**H**) Enrichment heatmap of Ssrp1 binding in WT ESCs, Usp7 binding in WT ESCs and Usp7 binding in $Ssrp1^{-/-}$ ESCs around binding regions of Ssrp1. The regions are sorted by the Ssrp1 binding strength. The ChIP-seq signal was calculated as the log2 ratio of normalized reads relative to the input. (**I**) ChIP-qPCR analysis of Usp7 enrichment on MERVL in WT ESCs and $Ssrp1^{-/-}$ ESCs. ChIP-qPCR data were normalized to input and Gapdh locus. Biological triplicate data ($n = 3$ extracts) are presented as mean ± s.e.m. (**J**) Usp7 binding profile around the center of MERVL locus in WT ESCs and $Ssrp1^{-/-}$ ESCs. The ChIP-seq signal was calculated as the log2 ratio of the normalized number of reads relative to the input. (**K**) ChIP-qPCR analysis of Usp7 enrichment on MERVL in $Ssrp1^{-/-}$ ESCs after overexpression of Supt16. Biological triplicate data ($n = 3$ dishes) are presented as mean ± s.e.m. Ctrl: control vector; OE: overexpression.

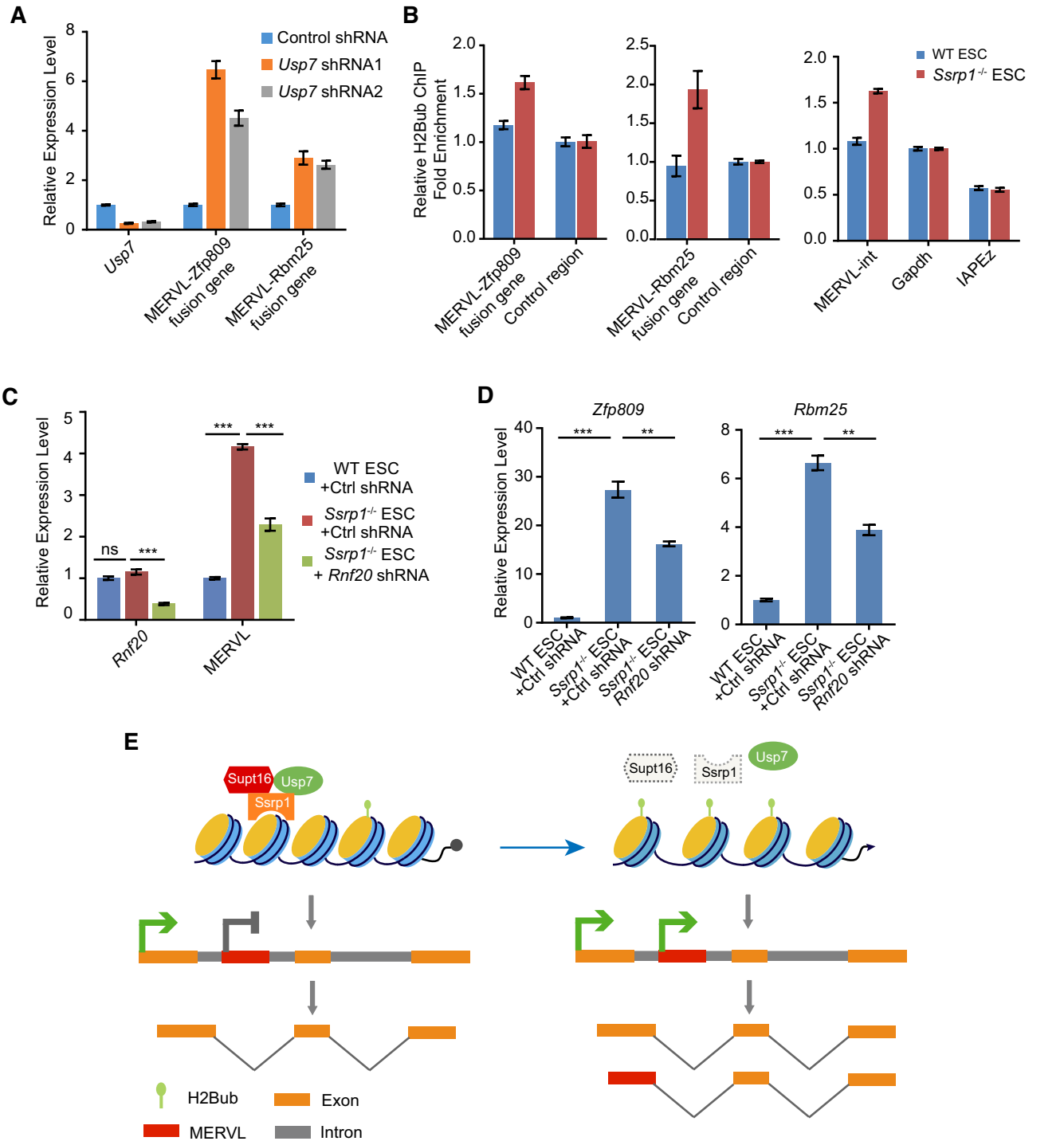

Figure 6. Ssrp1 and Usp7 suppress transcription of MERVL-fused transcripts by removing H2Bub. (**A**) qPCR analysis of expression of MERVL-fused transcripts after *Usp7* depletion. The data are represented as mean ±s.e.m. from three biological replicates. (**B**) ChIP-qPCR analysis of H2Bub enrichment on the specific loci of MERVL and MERVL consensus in WT ESCs and *Ssrp1*$^{-/-}$ ESCs. ChIP-qPCR data were normalized to input and that of the control region. Biological triplicate data (n = 3 extracts) are presented as mean ± s.e.m. (**C**) qPCR analysis of MERVL expression after *Rnf20* depletion in *Ssrp1*$^{-/-}$ ESCs. The data are represented as mean ± s.e.m. from three biological replicates. Ctrl: control. Significant differences were determined by Student's *t*-test and defined as *** $P < 0.001$. ns: non-significant. (**D**) qPCR analysis of expression of MERVL-fused transcripts after *Rnf20* depletion in WT and *Ssrp1*$^{-/-}$ ESCs. The data are represented as mean ± s.e.m. from three biological replicates. Significant differences were determined by Student's *t*-test and defined as *** $P < 0.001$ or ** $P < 0.01$. (**E**) Schematic of FACT complex function in the coordinated repression of MERVL and 2-cell cleavage genes in ESCs. In WT ESCs, Ssrp1 binds to MERVL, interacts with Supt16 and Usp7 and decreases H2Bub deposition on MERVL, thus represses the expression of MERVL and MERVL-derived cryptic promoters. In the absence of Ssrp1, Usp7 is dissociated from chromatin, and H2Bub is gained to activate the expression of MERVL and MERVL-driven cryptic transcription. Box with dotted lines, the genes with depleted expression.

and MERVL-fused genes, whereas, in the absence of Ssrp1, Usp7 dissociates from chromatin, thereby MERVL and MERVL-fused genes, including 2-cell genes, are activated (Figure 6E). Ssrp1 is usually associated with genes under active transcription, however, whether Ssrp1 activates gene expression directly has been questioned (62). Previous evidence hint that Ssrp1 is involved in transcription repression of certain genes (31–36). Our results supported the FACT complex as both a transcription repressor and an activator (Figure 3A). We showed that Ssrp1 interacted with Usp7, which is known to be associated with heterochromatin and mediate gene silencing (57). Usp7 deubiquitinates and stabilizes heterochromatin-associated proteins, including Uhrf1 and Dnmt1 (63–66). Usp7 also deubiquitinates telomeric shelterin component TPP1 and prolongs its stability (67). FACT complex could repress the expression of antisense transcripts and subtelomeric genes as well (31). Recently, FACT was found to be loaded to heterochromatin (68). The recruitment of Usp7 to MERVL supported an additional repression role of Ssrp1 (Figure 5I) besides its gene activation role. However, Ssrp1 may not repress MERVL through a canonical mechanism. It was noted that the binding profiles of Ssrp1 and Usp7 were positively correlated and they formed a distinct cluster different from known heterochromatin associated regulators (H3K9me2 and H3K9me3) of MERVL (Figure 5B), implying that Ssrp1 and Usp7 may modulate non-heterochromatic MERVL loci. Therefore, we discover a previously unrecognized ERV repression mechanism mediated by the FACT complex and Usp7.

Previously, it was documented that FACT repressed cryptic transcription (35–36,69) and corresponding mechanisms were well-studied in yeast cells. Yeast FACT is associated with coding regions of active genes (35,70–73), where they maintain proper chromatin structure. Disrupting FACT function causes nucleosome loss and mislocalization of the histone variants such as CENP-A, H2A.X and H2A.Z (74–76). Localization of histones and associated active histone marks are also affected by the disruption of FACT function (30). These lead to global activation of antisense transcription and de-repression of cryptic transcription in yeast (34–36,69). In comparison to the observations in yeast, loss of FACT complex did induce cryptic transcription in mouse ESCs (Supplementary Figure S4D–G). However, the activation of cryptic transcription was restricted to a subset of genes rather than a global effect in mouse ESCs (Figure 3A and Supplementary Figure S4D–G). Unlike in yeast (34), we did not observe genome-wide de-repression of antisense transcription either after the loss of *Ssrp1* in mouse ESCs (Supplementary Figure S4H). Furthermore, disruption of FACT complex caused ∼50% of yeast genes to show expression change (73), but *Ssrp1* deletion only caused a minority of genes (1108 genes) to change expression in mouse ESCs (Figure 3A), suggesting a restricted role of FACT during transcription regulation in ESCs. The modest gene expression changes caused by FACT depletion are also observed in other mammalian cell types (29,77). In contrast to the global activation of antisense transcripts in yeasts with FACT mutation (31,34), loss of FACT in mouse ESCs only upregulated a small number of antisense transcripts, of which most fused with TEs (Supplementary Figure S4H and I). These discrepancies in FACT functions between yeast and mammalian cells are probably due to species differences in genome structure and transcription regulation. In the future, it will be interesting to investigate whether changes in histone variants and histone marks contributed to the differences after FACT loss in mammalian and yeast cells.

In addition to the roles of FACT described in yeast, we showed that FACT repressed cryptic transcription through Usp7 by regulating H2Bub in mouse ESCs. Consistently, it was recently found that FACT mutant yeast demonstrated an increased level of H2Bub (78). The depletion of Ubp10, yeast counterpart of H2Bub deubiquitinase Usp7, activates cryptic transcription reporter in yeast strains with weak FACT mutation (*spt16–11*) (78). Moreover, the severe FACT mutation alone leads to massive cryptic transcription (35–36,69), which is further enhanced in combination with *Spt6* mutation (69). Similar to yeast FACT, loss of FACT complex alone activated cryptic transcription in ESCs (Figure 3A and Supplementary Figure S4D–G). In certain chromatin context, FACT and histone H2B ubiquitination work together to facilitate RNA polymerase II-mediated transcription and genome stability (79). However, it is recently discovered that that Ubp10 depends on FACT to efficiently cleave H2Bub from nucleosomes whereas the coupling of Ubp10 and FACT links cycles of H2B ubiquitination and deubiquitination in yeast (78). Therefore, Usp7 that is recruited by FACT complex (Figure 5H–J) probably plays a similar role to Ubp10 during ubiquitination and deubiquitination cycles in mammalian cells. The observation that disruption of FACT and Ubp10 function caused a global increase of H2Bub level (78) is in agreement with the genome-scale disruption of transcription after disturbing FACT function in yeast (34–36,69,73). In contrast, only a small percentage of genes were influenced after depletion of *Ssrp1* or *Usp7* in mouse ESCs (Figure 3A), although the recruitment of Usp7 was reduced on both MERVL and other Ssrp1 binding regions (Figure 5H-J). This means that the presence of other factors interacting with FACT complex or Usp7 may restrict the transcription regulatory role of FACT and Usp7 to certain chromatin contexts. Thus, it is worth to investigate which factor(s) are responsible for the loci specific control of mammalian FACT and Usp7 activity in the future study.

It is noteworthy that the depletion of *Usp7* activated MERVL to a lesser extent in comparison to *Ssrp1* knockout (Figure 1H and 5F). Hence, FACT may repress MERVL and MERVL-fused genes through other mechanisms besides the recruitment of Usp7. Other FACT-associated proteins (Figure 5A and Supplementary Table S5) may subsequently contribute to the repression of MERVL and MERVL-fused genes. Another possible reason is that other pathways repressing cryptic transcription initiation may be regulated by the FACT complex itself. The activation of MERVL-fused genes meant these genes need to switch from canonical promoters to MERVL-derived cryptic promoters (Figure 4E). FACT interacts with RNA polymerase II (Figure 5A and Supplementary Table S5) and maintains RNA polymerase II pausing at the proximal promoter region (80), which prevents downstream cryptic transcription initiation (81). Besides MERVL, other TEs are capable to work as cryptic promoters to drive expression of oncogenes in hu-

man cancers (37), implying a potential role of FACT and other repressors of cryptic transcription in tumorigenesis.

It is intriguing to see that FACT sets a barrier to the acquisition of totipotent cell fate by ESCs (Figure 4A). This is consistent with the notion that the FACT complex represses cell fate reprogramming (82), although the FACT complex is not essential to pluripotency maintenance (Figure 1G) (33,83). Here, we show that the FACT impedes 2-cell like fate conversion through suppressing cryptic promoters derived from TEs. Given that TE-derived promoters drive a considerable number of cell type-specific transcription during development and in various tissues (4,84), it is plausible for us to extrapolate that FACT may support cell fate fixation by repressing the activation of TE-derived promoters.

In conclusion, our study demonstrates that FACT complex recruits Usp7 to repress MERVL and MERVL-fused 2C genes in ESCs by impeding the ubiquitination of H2Bub. Our insights into the repressive role of FACT will inform the regulation of TE-derived cryptic promoters during mammalian development and in diseases.

DATA AVAILABILITY

High-throughput sequencing data generated from this study have been deposited in Gene Expression Omnibus under GSE141788 and GSE155863 for RNA-seq and GSE141791 for ChIP-seq. Data analyzed in this study are available in Gene Expression Omnibus under accession GSE77440 (H3K9me3 ChIP-seq) (85), GSE54412 (H3K9me2 ChIP-seq) (86), GSE99022 (Lsd1 ChIP-seq) (87), GSE70799 (Chaf1a and Sumo2 ChIP-seq) (26), GSE97945 (HP1g ChIP-seq) (88), GSE90045 (H2Bub ChIP-seq after *Usp7* knockout) (60) and GSE90893 (H3K4me3 and H3K36me3 ChIP-seq) (89).

The mass spectrometry proteomics data of Ssrp1 immunoprecipitation have been deposited to the ProteomeXchange Consortium (90) via the PRIDE (91) partner repository with the dataset identifier PXD017351.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Jian Shen for his assistance in flow cytometry analysis. We thank Daolei Dou and Xiangshuai Zhao from the instrument platform of State Key Laboratory of Medicinal Chemical Biology for their technical assistance.
Author Contributions: X.L. designed research; F.C., D.X. and T.G. performed most experiments; W.Z performed bioinformatics analysis; Z.D. established *Ssrp1*$^{-/-}$ ESCs lines; X.L., F.C. and W.Z. wrote the manuscript.

FUNDING

National Key Research and Development Program of China [2018YFA0107000]; National Natural Science Foundation of China [31671352, 31871488]; Fundamental Research Funds for the Central Universities. Funding for open access charge: National Key Research and Development Program of China.

Conflict of interest statement. None declared.

REFERENCES

1. Cordaux,R. and Batzer,M.A. (2009) The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.*, **10**, 691–703.
2. Stocking,C. and Kozak,C.A. (2008) Murine endogenous retroviruses. *Cell. Mol. Life Sci.*, **65**, 3383–3398.
3. Zhang,W., Chen,F., Chen,R., Xie,D., Yang,J., Zhao,X., Guo,R., Zhang,Y., Shen,Y., Goke,J. *et al.* (2019) Zscan4c activates endogenous retrovirus MERVL and cleavage embryo genes. *Nucleic Acids Res.*, **47**, 8485–8501.
4. Macfarlan,T.S., Gifford,W.D., Driscoll,S., Lettieri,K., Rowe,H.M., Bonanomi,D., Firth,A., Singer,O., Trono,D. and Pfaff,S.L. (2012) Embryonic stem cell potency fluctuates with endogenous retrovirus activity. *Nature*, **487**, 57–63.
5. Goke,J., Lu,X., Chan,Y.S., Ng,H.H., Ly,L.H., Sachs,F. and Szczerbinska,I. (2015) Dynamic transcription of distinct classes of endogenous retroviral elements marks specific populations of early human embryonic cells. *Cell Stem Cell*, **16**, 135–141.
6. Wang,C., Liu,X., Gao,Y., Yang,L., Li,C., Liu,W., Chen,C., Kou,X., Zhao,Y., Chen,J. *et al.* (2018) Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nat. Cell Biol.*, **20**, 620–631.
7. Bulut-Karslioglu,A., De La Rosa-Velazquez,I.A., Ramirez,F., Barenboim,M., Onishi-Seebacher,M., Arand,J., Galan,C., Winter,G.E., Engist,B., Gerle,B. *et al.* (2014) Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol. Cell*, **55**, 277–290.
8. Karimi,M.M., Goyal,P., Maksakova,I.A., Bilenky,M., Leung,D., Tang,J.X., Shinkai,Y., Mager,D.L., Jones,S., Hirst,M. *et al.* (2011) DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. *Cell Stem Cell*, **8**, 676–687.
9. Maksakova,I.A., Thompson,P.J., Goyal,P., Jones,S.J., Singh,P.B., Karimi,M.M. and Lorincz,M.C. (2013) Distinct roles of KAP1, HP1 and G9a/GLP in silencing of the two-cell-specific retrotransposon MERVL in mouse ES cells. *Epigenet. Chromatin*, **6**, 15.
10. Deniz,O., de la Rica,L., Cheng,K.C.L., Spensberger,D. and Branco,M.R. (2018) SETDB1 prevents TET2-dependent activation of IAP retroelements in naive embryonic stem cells. *Genome Biol.*, **19**, 6.
11. He,J., Fu,X., Zhang,M., He,F., Li,W., Abdul,M.M., Zhou,J., Sun,L., Chang,C., Li,Y. *et al.* (2019) Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells. *Nat. Commun.*, **10**, 34.
12. Lu,F., Liu,Y., Jiang,L., Yamaguchi,S. and Zhang,Y. (2014) Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev.*, **28**, 2103–2119.
13. Macfarlan,T.S., Gifford,W.D., Agarwal,S., Driscoll,S., Lettieri,K., Wang,J., Andrews,S.E., Franco,L., Rosenfeld,M.G., Ren,B. *et al.* (2011) Endogenous retroviruses and neighboring genes are coordinately repressed by LSD1/KDM1A. *Genes Dev.*, **25**, 594–607.
14. Thompson,P.J., Dulberg,V., Moon,K.M., Foster,L.J., Chen,C., Karimi,M.M. and Lorincz,M.C. (2015) hnRNP K coordinates transcriptional silencing by SETDB1 in embryonic stem cells. *PLos Genet.*, **11**, e1004933.
15. Rowe,H.M., Jakobsson,J., Mesnard,D., Rougemont,J., Reynard,S., Aktas,T., Maillard,P.V., Layard-Liesching,H., Verp,S., Marquis,J. *et al.* (2010) KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature*, **463**, 237–240.
16. Matsui,T., Leung,D., Miyashita,H., Maksakova,I.A., Miyachi,H., Kimura,H., Tachibana,M., Lorincz,M.C. and Shinkai,Y. (2010) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature*, **464**, 927–931.
17. Grover,P., Asa,J.S. and Campos,E.I. (2018) H3-H4 histone chaperone pathways. *Annu. Rev. Genet.*, **52**, 109–130.
18. Hammond,C.M., Stromme,C.B., Huang,H., Patel,D.J. and Groth,A. (2017) Histone chaperone networks shaping chromatin function. *Nat. Rev. Mol. Cell Biol.*, **18**, 141–158.
19. Elsasser,S.J., Noh,K.M., Diaz,N., Allis,C.D. and Banaszynski,L.A. (2015) Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature*, **522**, 240–244.
20. He,Q., Kim,H., Huang,R., Lu,W., Tang,M., Shi,F., Yang,D., Zhang,X., Huang,J., Liu,D. *et al.* (2015) The Daxx/Atrx complex

protects tandem repetitive elements during DNA hypomethylation by promoting H3K9 trimethylation. *Cell Stem Cell*, **17**, 273–286.
21. Voon,H.P., Hughes,J.R., Rode,C., De La Rosa-Velazquez,I.A., Jenuwein,T., Feil,R., Higgs,D.R. and Gibbons,R.J. (2015) ATRX plays a key role in maintaining silencing at interstitial heterochromatic loci and imprinted genes. *Cell Rep.*, **11**, 405–418.
22. Sadic,D., Schmidt,K., Groh,S., Kondofersky,I., Ellwart,J., Fuchs,C., Theis,F.J. and Schotta,G. (2015) Atrx promotes heterochromatin formation at retrotransposons. *EMBO Rep.*, **16**, 836–850.
23. Hoelper,D., Huang,H., Jain,A.Y., Patel,D.J. and Lewis,P.W. (2017) Structural and mechanistic insights into ATRX-dependent and -independent functions of the histone chaperone DAXX. *Nat. Commun.*, **8**, 1193.
24. Hatanaka,Y., Inoue,K., Oikawa,M., Kamimura,S., Ogonuki,N., Kodama,E.N., Ohkawa,Y., Tsukada,Y. and Ogura,A. (2015) Histone chaperone CAF-1 mediates repressive histone modifications to protect preimplantation mouse embryos from endogenous retrotransposons. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 14641–14646.
25. Ishiuchi,T., Enriquez-Gasca,R., Mizutani,E., Boskovic,A., Ziegler-Birling,C., Rodriguez-Terrones,D., Wakayama,T., Vaquerizas,J.M. and Torres-Padilla,M.E. (2015) Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat. Struct. Mol. Biol.*, **22**, 662–671.
26. Yang,B.X., El Farran,C.A., Guo,H.C., Yu,T., Fang,H.T., Wang,H.F., Schlesinger,S., Seah,Y.F., Goh,G.Y., Neo,S.P. *et al.* (2015) Systematic identification of factors for provirus silencing in embryonic stem cells. *Cell*, **163**, 230–245.
27. Winkler,D.D. and Luger,K. (2011) The histone chaperone FACT: structural insights and mechanisms for nucleosome reorganization. *J. Biol. Chem.*, **286**, 18369–18374.
28. Hsieh,F.K., Kulaeva,O.I., Patel,S.S., Dyer,P.N., Luger,K., Reinberg,D. and Studitsky,V.M. (2013) Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 7654–7659.
29. Li,Y., Zeng,S.X., Landais,I. and Lu,H. (2007) Human SSRP1 has Spt16-dependent and -independent roles in gene transcription. *J. Biol. Chem.*, **282**, 6936–6945.
30. Jeronimo,C., Poitras,C. and Robert,F. (2019) Histone recycling by FACT and Spt6 during transcription prevents the scrambling of histone modifications. *Cell Rep.*, **28**, 1206–1218.
31. Murawska,M., Schauer,T., Matsuda,A., Wilson,M.D., Pysik,T., Wojcik,F., Muir,T.W., Hiraoka,Y., Straub,T. and Ladurner,A.G. (2020) The chaperone FACT and histone H2B Ubiquitination Maintain S. pombe genome architecture through genic and subtelomeric functions. *Mol. Cell*, **77**, 501–513.
32. Nielsen,M., Ard,R., Leng,X., Ivanov,M., Kindgren,P., Pelechano,V. and Marquardt,S. (2019) Transcription-driven chromatin repression of Intragenic transcription start sites. *PLos Genet.*, **15**, e1007969.
33. Mylonas,C. and Tessarz,P. (2018) Transcriptional repression by FACT is linked to regulation of chromatin accessibility at the promoter of ES cells. *Life Sci. Alliance*, **1**, e201800085.
34. Feng,J., Gan,H., Eaton,M.L., Zhou,H., Li,S., Belsky,J.A., MacAlpine,D.M., Zhang,Z. and Li,Q. (2016) Noncoding transcription is a driving force for nucleosome instability in spt16 mutant cells. *Mol. Cell. Biol.*, **36**, 1856–1867.
35. Mason,P.B. and Struhl,K. (2003) The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol. Cell. Biol.*, **23**, 8323–8333.
36. Kaplan,C.D., Laprade,L. and Winston,F. (2003) Transcription elongation factors repress transcription initiation from cryptic sites. *Science*, **301**, 1096–1099.
37. Jang,H.S., Shah,N.M., Du,A.Y., Dailey,Z.Z., Pehrsson,E.C., Godoy,P.M., Zhang,D., Li,D., Xing,X., Kim,S. *et al.* (2019) Transposable elements drive widespread expression of oncogenes in human cancers. *Nat. Genet.*, **51**, 611–617.
38. Brocks,D., Schmidt,C.R., Daskalakis,M., Jang,H.S., Shah,N.M., Li,D., Li,J., Zhang,B., Hou,Y., Laudato,S. *et al.* (2017) DNMT and HDAC inhibitors induce cryptic transcription start sites encoded in long terminal repeats. *Nat. Genet.*, **49**, 1052–1060.
39. Yuan,B., Latek,R., Hossbach,M., Tuschl,T. and Lewitter,F. (2004) siRNA Selection Server: an automated siRNA oligonucleotide prediction server. *Nucleic Acids Res.*, **32**, W130–W134.
40. Ran,F.A., Hsu,P.D., Wright,J., Agarwala,V., Scott,D.A. and Zhang,F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.*, **8**, 2281–2308.
41. Kim,D., Paggi,J.M., Park,C., Bennett,C. and Salzberg,S.L. (2019) Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.*, **37**, 907–915.
42. Howe,K.L., Contreras-Moreira,B., De Silva,N., Maslen,G., Akanni,W., Allen,J., Alvarez-Jarreta,J., Barba,M., Bolser,D.M., Cambell,L. *et al.* (2020) Ensembl genomes 2020-enabling non-vertebrate genomic research. *Nucleic Acids Res.*, **48**, D689–D695.
43. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
44. Fresno,C. and Fernandez,E.A. (2013) RDAVIDWebService: a versatile R interface to DAVID. *Bioinformatics*, **29**, 2810–2811.
45. Pertea,M., Pertea,G.M., Antonescu,C.M., Chang,T.C., Mendell,J.T. and Salzberg,S.L. (2015) StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.*, **33**, 290–295.
46. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
47. Ramirez,F., Dundar,F., Diehl,S., Gruning,B.A. and Manke,T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.*, **42**, W187–W191.
48. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W. *et al.* (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.
49. Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M. and Haussler,D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.
50. Safina,A., Garcia,H., Commane,M., Guryanova,O., Degan,S., Kolesnikova,K. and Gurova,K.V. (2013) Complex mutual regulation of facilitates chromatin transcription (FACT) subunits on both mRNA and protein levels in human cells. *Cell Cycle*, **12**, 2423–2434.
51. Liu,Y., Zhou,K., Zhang,N., Wei,H., Tan,Y.Z., Zhang,Z., Carragher,B., Potter,C.S., D'Arcy,S. and Luger,K. (2020) FACT caught in the act of manipulating the nucleosome. *Nature*, **577**, 426–431.
52. McCullough,L.L., Connell,Z., Xin,H., Studitsky,V.M., Feofanov,A.V., Valieva,M.E. and Formosa,T. (2018) Functional roles of the DNA-binding HMGB domain in the histone chaperone FACT in nucleosome reorganization. *J. Biol. Chem.*, **293**, 6121–6133.
53. Chang,H.W., Valieva,M.E., Safina,A., Chereji,R.V., Wang,J., Kulaeva,O.I., Morozov,A.V., Kirpichnikov,M.P., Feofanov,A.V., Gurova,K.V. *et al.* (2018) Mechanism of FACT removal from transcribed genes by anticancer drugs curaxins. *Sci. Adv.*, **4**, eaav2131.
54. Peaston,A.E., Evsikov,A.V., Graber,J.H., de Vries,W.N., Holbrook,A.E., Solter,D. and Knowles,B.B. (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. *Dev. Cell*, **7**, 597–606.
55. Franke,V., Ganesh,S., Karlic,R., Malik,R., Pasulka,J., Horvat,F., Kuzman,M., Fulka,H., Cernohorska,M., Urbanova,J. *et al.* (2017) Long terminal repeats power evolution of genes and gene expression programs in mammalian oocytes and zygotes. *Genome Res.*, **27**, 1384–1394.
56. Brind'Amour,J., Kobayashi,H., Richard Albert,J., Shirane,K., Sakashita,A., Kamio,A., Bogutz,A., Koike,T., Karimi,M.M., Lefebvre,L. *et al.* (2018) LTR retrotransposons transcribed in oocytes drive species-specific and heritable changes in DNA methylation. *Nat. Commun.*, **9**, 3331.
57. van der Knaap,J.A., Kumar,B.R., Moshkin,Y.M., Langenberg,K., Krijgsveld,J., Heck,A.J., Karch,F. and Verrijzer,C.P. (2005) GMP synthetase stimulates histone H2B deubiquitylation by the epigenetic silencer USP7. *Mol. Cell*, **17**, 695–707.
58. Maertens,G.N., El Messaoudi-Aubert,S., Elderkin,S., Hiom,K. and Peters,G. (2010) Ubiquitin-specific proteases 7 and 11 modulate Polycomb regulation of the INK4a tumour suppressor. *EMBO J.*, **29**, 2553–2565.
59. Sarkari,F., Sanchez-Alcaraz,T., Wang,S., Holowaty,M.N., Sheng,Y. and Frappier,L. (2009) EBNA1-mediated recruitment of a histone H2B deubiquitylating complex to the Epstein-Barr virus latent origin of DNA replication. *PLoS Pathog.*, **5**, e1000624.

60. Liefke,R., Karwacki-Neisius,V. and Shi,Y. (2016) EPOP interacts with elongin BC and USP7 to modulate the chromatin landscape. *Mol. Cell*, **64**, 659–672.
61. Jin,Q., Martinez,C.A., Arcipowski,K.M., Zhu,Y., Gutierrez-Diaz,B.T., Wang,K.K., Johnson,M.R., Volk,A.G., Wang,F., Wu,J. *et al.* (2019) USP7 cooperates with NOTCH1 to drive the oncogenic transcriptional program in T-cell leukemia. *Clin. Cancer Res.*, **25**, 222–239.
62. Gurova,K., Chang,H.W., Valieva,M.E., Sandlesh,P. and Studitsky,V.M. (2018) Structure and function of the histone chaperone FACT—resolving FACTual issues. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1861**, 892–904.
63. Cheng,J., Yang,H., Fang,J., Ma,L., Gong,R., Wang,P., Li,Z. and Xu,Y. (2015) Molecular mechanism for USP7-mediated DNMT1 stabilization by acetylation. *Nat. Commun.*, **6**, 7023.
64. Felle,M., Joppien,S., Nemeth,A., Diermeier,S., Thalhammer,V., Dobner,T., Kremmer,E., Kappler,R. and Langst,G. (2011) The USP7/Dnmt1 complex stimulates the DNA methylation activity of Dnmt1 and regulates the stability of UHRF1. *Nucleic Acids Res.*, **39**, 8355–8365.
65. Qin,W., Leonhardt,H. and Spada,F. (2011) Usp7 and Uhrf1 control ubiquitination and stability of the maintenance DNA methyltransferase Dnmt1. *J. Cell. Biochem.*, **112**, 439–444.
66. Meng,H., Harrison,D.J. and Meehan,R.R. (2015) MBD4 interacts with and recruits USP7 to heterochromatic foci. *J. Cell. Biochem.*, **116**, 476–485.
67. Zemp,I. and Lingner,J. (2014) The shelterin component TPP1 is a binding partner and substrate for the deubiquitinating enzyme USP7. *J. Biol. Chem.*, **289**, 28595–28606.
68. Holla,S., Dhakshnamoorthy,J., Folco,H.D., Balachandran,V., Xiao,H., Sun,L.L., Wheeler,D., Zofall,M. and Grewal,S.I.S. (2020) Positioning heterochromatin at the nuclear periphery suppresses histone turnover to promote epigenetic inheritance. *Cell*, **180**, 150–164.
69. Cheung,V., Chua,G., Batada,N.N., Landry,C.R., Michnick,S.W., Hughes,T.R. and Winston,F. (2008) Chromatin- and transcription-related factors repress transcription from within coding regions throughout the Saccharomyces cerevisiae genome. *PLoS Biol.*, **6**, e277.
70. Martin,B.J.E., Chruscicki,A.T. and Howe,L.J. (2018) Transcription promotes the interaction of the facilitates chromatin transactions (FACT) complex with nucleosomes in Saccharomyces cerevisiae. *Genetics*, **210**, 869–881.
71. Pathak,R., Singh,P., Ananthakrishnan,S., Adamczyk,S., Schimmel,O. and Govind,C.K. (2018) Acetylation-dependent recruitment of the FACT complex and its role in regulating Pol II occupancy genome-wide in Saccharomyces cerevisiae. *Genetics*, **209**, 743–756.
72. Vinayachandran,V., Reja,R., Rossi,M.J., Park,B., Rieber,L., Mittal,C., Mahony,S. and Pugh,B.F. (2018) Widespread and precise reprogramming of yeast protein-genome interactions in response to heat shock. *Genome Res.*, **28**, 357–366.
73. True,J.D., Muldoon,J.J., Carver,M.N., Poorey,K., Shetty,S.J., Bekiranov,S. and Auble,D.T. (2016) The Modifier of Transcription 1 (Mot1) ATPase and Spt16 histone chaperone co-regulate transcription through preinitiation complex assembly and nucleosome organization. *J. Biol. Chem.*, **291**, 15307–15319.
74. Jeronimo,C., Watanabe,S., Kaplan,C.D., Peterson,C.L. and Robert,F. (2015) The histone chaperones FACT and Spt6 restrict H2A.Z from intragenic locations. *Mol. Cell*, **58**, 1113–1123.
75. Deyter,G.M. and Biggins,S. (2014) The FACT complex interacts with the E3 ubiquitin ligase Psh1 to prevent ectopic localization of CENP-A. *Genes Dev.*, **28**, 1815–1826.
76. Heo,K., Kim,H., Choi,S.H., Choi,J., Kim,K., Gu,J., Lieber,M.R., Yang,A.S. and An,W. (2008) FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. *Mol. Cell*, **30**, 86–97.
77. Fleyshman,D., Prendergast,L., Safina,A., Paszkiewicz,G., Commane,M., Morgan,K., Attwood,K. and Gurova,K. (2017) Level of FACT defines the transcriptional landscape and aggressive phenotype of breast cancer cells. *Oncotarget*, **8**, 20525–20542.
78. Nune,M., Morgan,M.T., Connell,Z., McCullough,L., Jbara,M., Sun,H., Brik,A., Formosa,T. and Wolberger,C. (2019) FACT and Ubp10 collaborate to modulate H2B deubiquitination and nucleosome dynamics. *Elife*, **8**, e40988.
79. Pavri,R., Zhu,B., Li,G., Trojer,P., Mandal,S., Shilatifard,A. and Reinberg,D. (2006) Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell*, **125**, 703–717.
80. Tettey,T.T., Gao,X., Shao,W., Li,H., Story,B.A., Chitsazan,A.D., Glaser,R.L., Goode,Z.H., Seidel,C.W., Conaway,R.C. *et al.* (2019) A role for FACT in RNA Polymerase II Promoter-Proximal pausing. *Cell Rep.*, **27**, 3770–3779.
81. Shao,W. and Zeitlinger,J. (2017) Paused RNA polymerase II inhibits new transcriptional initiation. *Nat. Genet.*, **49**, 1045–1051.
82. Kolundzic,E., Ofenbauer,A., Bulut,S.I., Uyar,B., Baytek,G., Sommermeier,A., Seelk,S., He,M., Hirsekorn,A., Vucicevic,D. *et al.* (2018) FACT sets a barrier for cell fate reprogramming in Caenorhabditis elegans and human cells. *Dev. Cell*, **46**, 611–626.
83. Shen,Z., Formosa,T. and Tantin,D. (2018) FACT inhibition blocks induction but not maintenance of pluripotency. *Stem Cells Dev.*, **27**, 1693–1701.
84. Mele,M., Ferreira,P.G., Reverter,F., DeLuca,D.S., Monlong,J., Sammeth,M., Young,T.R., Goldmann,J.M., Pervouchine,D.D., Sullivan,T.J. *et al.* (2015) Human genomics. The human transcriptome across tissues and individuals. *Science*, **348**, 660–665.
85. Riso,V., Cammisa,M., Kukreja,H., Anvar,Z., Verde,G., Sparago,A., Acurzio,B., Lad,S., Lonardo,E., Sankar,A. *et al.* (2016) ZFP57 maintains the parent-of-origin-specific expression of the imprinted genes and differentially affects non-imprinted targets in mouse embryonic stem cells. *Nucleic Acids Res.*, **44**, 8165–8178.
86. Liu,N., Zhang,Z., Wu,H., Jiang,Y., Meng,L., Xiong,J., Zhao,Z., Zhou,X., Li,J., Li,H. *et al.* (2015) Recognition of H3K9 methylation by GLP is required for efficient establishment of H3K9 methylation, rapid target gene repression, and mouse viability. *Genes Dev.*, **29**, 379–393.
87. Cao,K., Collings,C.K., Morgan,M.A., Marshall,S.A., Rendleman,E.J., Ozark,P.A., Smith,E.R. and Shilatifard,A. (2018) An Mll4/COMPASS-Lsd1 epigenetic axis governs enhancer function and pluripotency transition in embryonic stem cells. *Sci. Adv.*, **4**, eaap8747.
88. Ostapcuk,V., Mohn,F., Carl,S.H., Basters,A., Hess,D., Iesmantavicius,V., Lampersberger,L., Flemr,M., Pandey,A., Thoma,N.H. *et al.* (2018) Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. *Nature*, **557**, 739–743.
89. Chronis,C., Fiziev,P., Papp,B., Butz,S., Bonora,G., Sabri,S., Ernst,J. and Plath,K. (2017) Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell*, **168**, 442–459.
90. Deutsch,E.W., Csordas,A., Sun,Z., Jarnuczak,A., Perez-Riverol,Y., Ternent,T., Campbell,D.S., Bernal-Llinares,M., Okuda,S., Kawano,S. *et al.* (2017) The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res.*, **45**, D1100–D1106.
91. Perez-Riverol,Y., Csordas,A., Bai,J., Bernal-Llinares,M., Hewapathirana,S., Kundu,D.J., Inuganti,A., Griss,J., Mayer,G., Eisenacher,M. *et al.* (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.*, **47**, D442–D450.