FACT (facilitates chromatin transcription) is a protein complex that allows RNA polymerase II (RNAPII) to overcome the nucleosome-induced barrier to transcription. While abundant in undifferentiated cells and many cancers, FACT is not abundant or is absent in most tissues. Therefore, we screened for additional proteins that might replace FACT upon differentiation. We identified two proteins, lens epithelium-derived growth factor (LEDGF) and hepatoma-derived growth factor 2 (HDGF2), each containing two high mobility group A (HMGA)–like AT-hooks and a methyl-lysine reading Pro-Trp-Trp-Pro (PWWP) domain that binds to H3K36me2 and H3K36me3. LEDGF and HDGF2 colocalize with H3K36me2/3 at genomic regions containing active genes. In myoblasts, LEDGF and HDGF2 are enriched on most active genes. Upon differentiation to myotubes, LEDGF levels decrease, while HDGF2 levels are maintained. Moreover, HDGF2 is required for their proper expression. HDGF2 knockout myoblasts exhibit an accumulation of paused RNAPII within the transcribed region of many HDGF2 target genes, indicating a defect in early elongation.

INTRODUCTION

RNA polymerase II (RNAPII) transcription is regulated at the level of initiation, pause release, promoter escape, +1 nucleosome release, and elongation (1–3). After promoter escape, RNAPII must overcome a nucleosome-induced barrier to transcription (4, 5). Two decades ago, we identified FACT (facilitates chromatin transcription) as a protein complex composed of suppressor of Ty elements 16 (SPT16), a suppressor of Ty elements 14 (SPT16), and structure specific recognition protein 1 (SSRP1) that alleviates this nucleosome-induced barrier to transcription (6, 7). More recently, we mapped the genomic binding of FACT in stem cells using chromatin immunoprecipitation sequencing (ChIP-seq) with an SPT16 antibody and found that FACT occupancy varies, being associated at high levels with only a subset of active genes (Fig. 1A). In addition, we and others have found that FACT is only expressed at high levels in some progenitor and transformed cells (fig. S1A) (8, 9). These observations imply the existence of alternative FACT-like chaperones in differentiated cells. We first hypothesized that the BET family proteins (Brd2, Brd3, and Brd4), which also have FACT-like activity in vitro, might replace FACT at these genes (10, 11). However, bromodomain and extra-terminal domain (BET) proteins predominately localize with acetylated nucleosomes proximal to the transcription start site (TSS) of genes, while a chaperone capable of fulfilling the function of FACT should localize in gene bodies (Fig. S1B) (11, 12).

RESULTS

Exploiting a similar biochemical strategy that first led us to identify FACT, we fractionated HeLa cell nuclear extract and identified a fraction that was depleted of FACT, BET proteins, and nucleolin (RNAPII-specific FACT-like chaperone), yet able to support transcription through nucleosomes (fig. S1, C and D) (6, 10, 11, 13). Through further chromatographic fractionation and mass spectrometry (MS), we identified LEDGF long isoform (p75) (also known as PSIP1) (Fig. 1B, fig. S1E, and table S1) (14, 15). Semiquantitative proteomics and RNA sequencing (RNA-seq) data suggest that this protein and its family member, HDGF2, are relatively abundant and expressed in most tissues, unlike the restricted expression of FACT. LEDGF and HDGF2 each contain a methyl-lysine reading Pro-Trp-Trp-Pro (PWWP) domain that has been shown to recognize H3K36me2/3 (histone H3 lysine 36 di- and trimethylation), two HMGA-like AT-hooks (similar to nucleolin) and an integrase binding domain (IBD) (Fig. 1C) (16–20). The short isoform of LEDGF (p52) and another PWWP-domain containing protein, HDGF, lack the IBD, and the latter contains a high mobility group box domain instead of high mobility group A (HMGA)–like AT-hooks (21). LEDGF and HDGF2 have generated considerable interest given their requirement for lentiviral integration, which favors integration into the body of transcribed genes by concurrently binding directly to lentiviral integrase and H3K36me2/3 in the host cell chromatin (17, 22). Retroviruses, which favor integration near the TSS of transcribed genes, use BET proteins in an analogous manner (23, 24). LEDGF and HDGF2 share another similarity with BET proteins in that they remain bound to mitotic chromosomes, suggesting that they might contribute to transcriptional memory (25–27).

To validate the FACT-like activity of LEDGF and its related proteins, we generated and tested purified recombinant versions of these proteins in a fully in vitro reconstituted chromatin transcription assay (Fig. 1D). As depicted, both isoforms of LEDGF (p75 and p52) and HDGF2 allow RNAPII to transcribe through nucleosomes in a manner similar to FACT. HDGF, which lacks the HMGA-like AT-hooks, did not substitute for FACT, suggesting that the FACT-like activity of these proteins probably lies within the region that contains the AT-hooks (21). Increasing amounts of FACT, LEDGF, or HDGF2 led to a linear increase in transcription activity such that at least two or more molecules of LEDGF and HDGF2 per nucleosome appeared to be required for efficient transcription in the assay (Fig. 1, E and F).
LEDGF and HDGF2 did not stimulate basal transcription from naked DNA templates, indicating that their activity is nucleosome specific (fig. S2A). In addition, as the IBD shares homology with the transcription factor IIIS (TFIIS) domain superfamily, we tested LEDGF and HDGF2 in an assay that detects TFIIS-RNAPII elongation activity on naked DNA templates. Neither LEDGF nor HDGF2 stimulated RNAPII elongation activity in this assay, implying that their function is nucleosome dependent (fig. S2B) (28).
To mechanistically uncover how LEDGF and HDGF2 function in our chromatin transcription assay, we tested their ability to function as a nucleosome chaperone. As these proteins are known to bind H3, unlike FACT, which binds to H2A/H2B, we suspected that they function similar to BET proteins, which also bind H3/H4 (16, 19, 25, 29). Therefore, we tested HDGF2 in a nucleosome transfer assay similar to the assay used to define chaperone activity for Brd2 (Fig. 1G) (10). This assay detects the transfer of histone octamers from oligonucleosome templates to a 32P-labeled 601 trapping DNA. This assay primarily detects the destabilization of nucleosomes and release of histones from the oligonucleosome template, as the free histones will preferentially bind to the 601 DNA sequence. As demonstrated in Fig. 1H, titrating HDGF2 in this assay lead to increasing amounts of nucleosomes formed on the 601 trapping DNA. As LEDGF and HDGF2 are known to bind H3 and not H2A/H2B, we also performed the assay using oligotetrasome donors (nucleosomes lacking H2A/H2B) to better define the nucleosome destabilizing properties of HDGF2 (Fig. S2C). We did not observe a transfer of tetrasomes, suggesting that HDGF2 requires the nucleosomal structure for this activity.

As the PWWP domains of these proteins are reported to bind to H3K36me2/3, we attempted to reconstitute the dependency on these modifications in vitro using the reconstituted transcription assay with synthetic modification-mimic histones but were not successful (fig. S2D) (16, 19, 27, 30). This is likely due to other domains (i.e., AT-hooks) of the LEDGF and HDGF2 proteins that exhibit DNA and nucleosome binding affinities independent of their PWWP domains and are sufficient to bind unmodified nucleosomes in a defined in vitro assay (16).

To investigate the functions of LEDGF and HDGF2 in RNAPII transcription in vivo, we first asked whether their binding correlates with regions having RNAPII occupancy and variable levels of FACT initially in 293T cells (Fig. 2A). Notably, all three factors, FACT, LEDGF, and HDGF2, are highly expressed in 293T cells, unlike more differentiated cells, where FACT expression is low (Fig. 2B and fig. S1A). Thus, we performed native ChIP-seq using FLAG antibody in 293T cells stably expressing FLAG-LEDGF and FLAG-HDGF2 by quantitative ChIP-MS. As the PWWP domains of these proteins are reported to bind to H3K36me2/3, we attempted to reconstitute the dependency on these modifications in vitro using the reconstituted transcription assay with synthetic modification-mimic histones but were not successful (fig. S2D) (16, 19, 27, 30). This is likely due to other domains (i.e., AT-hooks) of the LEDGF and HDGF2 proteins that exhibit DNA and nucleosome binding affinities independent of their PWWP domains and are sufficient to bind unmodified nucleosomes in a defined in vitro assay (16).

To investigate the functions of LEDGF and HDGF2 in RNAPII transcription in vivo, we first asked whether their binding correlates with regions having RNAPII occupancy and variable levels of FACT initially in 293T cells (Fig. 2A). Notably, all three factors, FACT, LEDGF, and HDGF2, are highly expressed in 293T cells, unlike more differentiated cells, where FACT expression is low (Fig. 2B and fig. S1A). Thus, we performed native ChIP-seq using FLAG antibody in 293T cells stably expressing FLAG-LEDGF and FLAG-HDGF2 by quantitative ChIP-MS. As the PWWP domains of these proteins are reported to bind to H3K36me2/3, we attempted to reconstitute the dependency on these modifications in vitro using the reconstituted transcription assay with synthetic modification-mimic histones but were not successful (fig. S2D) (16, 19, 27, 30). This is likely due to other domains (i.e., AT-hooks) of the LEDGF and HDGF2 proteins that exhibit DNA and nucleosome binding affinities independent of their PWWP domains and are sufficient to bind unmodified nucleosomes in a defined in vitro assay (16).

To investigate the functions of LEDGF and HDGF2 in RNAPII transcription in vivo, we first asked whether their binding correlates with regions having RNAPII occupancy and variable levels of FACT initially in 293T cells (Fig. 2A). Notably, all three factors, FACT, LEDGF, and HDGF2, are highly expressed in 293T cells, unlike more differentiated cells, where FACT expression is low (Fig. 2B and fig. S1A). Thus, we performed native ChIP-seq using FLAG antibody in 293T cells stably expressing FLAG-LEDGF and FLAG-HDGF2 by quantitative ChIP-MS. As the PWWP domains of these proteins are reported to bind to H3K36me2/3, we attempted to reconstitute the dependency on these modifications in vitro using the reconstituted transcription assay with synthetic modification-mimic histones but were not successful (fig. S2D) (16, 19, 27, 30). This is likely due to other domains (i.e., AT-hooks) of the LEDGF and HDGF2 proteins that exhibit DNA and nucleosome binding affinities independent of their PWWP domains and are sufficient to bind unmodified nucleosomes in a defined in vitro assay (16).

Our detailed bioinformatics analysis revealed that SPT16 binding is variable, not equally enriched on all clusters of active genes, whereas HDGF2 exhibits a binding pattern similar to LEDGF, being more restricted to the bodies of actively transcribed genes, whereas HDGF2 exhibits a binding pattern similar to LEDGF, being more restricted to the bodies of actively transcribed genes (Fig. 3, A and B). Upon differentiation (ESC to EB), approximately 881 genes showed a twofold (or more) increase in RNAPII binding, and HDGF2 was recruited to 40% of these genes (Fig. 3C). The ChIP-seq binding patterns of SPT16 revealed that it was also recruited to 58% of these genes and had a 23% overlap with HDGF2 (Fig. 3C). Moreover, the opposite trend holds true, as genes with reduced RNAPII (ESC to EB) usually exhibited a decrease in HDGF2 and SPT16 binding. Representative examples of specific genes showing these phenomena are shown in Fig. 3 (A and B). These results indicate that both HDGF2 and FACT follow RNAPII deposition at a subset of genes during mESC differentiation.

To validate that LEDGF and HDGF2 participate in transcription, we used CRISPR-Cas9 technology to knock out LEDGF alone or together with HDGF2 [LEDGF/HDGF2 double KO (dKO)] in mESCs. We were unable to knock out FACT in mESCs, and the dropout rank score of guide RNAs (gRNAs) from several CRISPR-CAS9 screens suggests that it is essential in mESCs (table S3). We did not observe a large number of genes whose expression (RNA-seq, twofold change cutoff, BH-corrected P < 0.05) was affected in either KO cell line compared to wild-type (WT) mESCs (Fig. 3D), suggesting that there is a redundancy likely due to high levels of FACT or other
unknown factors (Fig. 2B). When we differentiate these mESCs into EBs, we observed a substantial number of genes (1884 genes) affected in the LEDGF/HDGF2 dKO cells (Fig. 3D). However, the changes in the LEDGF KO cells were still minimal, implying a redundancy between LEDGF and HDGF2 at this stage of cellular differentiation.

Given that HDGF2 was recruited to approximately 40% of the genes that increase expression during differentiation of ESCs to EBs (Fig. 3C), we next analyzed the effect of the LEDGF/HDGF2 dKO on these up-regulated genes. Upon differentiation (ESC to EB), 2308 genes exhibited at least a twofold increase in expression in WT cells as measured by RNA-seq (Fig. 3E). Of these differentially expressed genes, ~27% failed to fully activate in the LEDGF/HDGF2 dKO cells (Fig. 3, E and F). Average ChIP-seq read density profiles showed a correlation between these genes and an increase in HDGF2 binding.
Fig. 3. LEDGF and HDGF2 are required for the induction of some genes in stem cells. (A and B) ChiP-seq tracks for LEDGF, HDGF2, SPT16, and RNAPII with corresponding RNA-seq tracks in ESCs and in EBs at the Tpm1 gene (A) and Esr sb gene (B). (C) Pie chart showing the percentage of genes exhibiting increased HDGF2 and/or FACT binding on genes with an increase in RNAPII during ESC differentiation into EB. (D) Mean average (MA) plots showing the number of differentially expressed genes (twofold, BH-corrected \( P < 0.05 \)) in LEDGF KO and LEDGF/HDGF2 double KO (dKO) in ESC to EB differentiation. (E) Venn diagram depicting the number of up-regulated genes (twofold) from ESC to EB in wild-type (WT) cells overlaid with the down-regulated genes in the LEDGF/HDGF2 dKO EBs. Of the 2308 genes normally up-regulated in the WT cells (ESC to EB), 618 were found to be down-regulated in the dKO EB cells. The overlap with dKO down-regulated genes is significant (\( P = 3.8 \times 10^{-3} \), hypergeometric test). (F) Boxplots with confidence intervals of expression of the 618 genes selected from (E) at WT ESC, WT EB, and dKO EB. Wilcoxon rank sum test: WT ESCs versus WT EBs, \( P = 4.5 \times 10^{-13} \); WT ESCs versus dKO EBs, \( P = 3.8 \times 10^{-3} \). (G) Top: Average density ChiP-seq profiles of HDGF2, SPT16, and LEDGF on the 618 genes that are up-regulated in WT EBs and down-regulated in dKO EBs. Bottom: Average density ChiP-seq profiles of HDGF2, SPT16, and LEDGF on the 4999 genes identified as not having any change in expression between ESCs and EBs.

(Esc to EB) (Fig. 3G, top left). Conversely, we did not observe this correlation on the genes that were activated to the same extent in dKO cells as in WT cells (Fig. 3G, bottom left). However, these genes did correlate with an increase in SPT16 binding, possibly indicating redundancy (Fig. 3G, bottom middle). SPT16 binding generally showed an overall increase in most genes that gained expression from ESC to EB, even those genes that failed to properly induce in the dKO cells (Fig. 3G). This redundancy could explain why many of the induction-compromised genes still showed partial activation in the dKO cells (Fig. 3F).

To control for the possible redundancy between LEDGF/HDGF2 and FACT, we used the myoblast (MB) to myotube (MT) cellular differentiation system (35, 36) as protein levels of LEDGF and the minimal levels of FACT (SPT16 and SSRP1) decrease, while HDGF2 remained constant during differentiation into MT (Fig. 4A). As expected, binding of both LEDGF and HDGF2 on the genome...
correlated with RNAPII levels as evidenced by average density profiles in MB (Fig. 4B, top, and fig. S6A). However, SPT16 was only detected at low levels on a few highly expressed genes in MB (fig. S7). Upon differentiation into MT, SPT16 was not detected on any genes and LEDGF binding decreased globally, while HDGF2 remained enriched on actively transcribed loci and accumulated on genes whose expression was induced (Fig. 4B, bottom, and figs. S8D and S9D).

As these data indicate that an MB to MT differentiation system is ideal for studying the function of HDGF2, we generated an HDGF2 KO in MB (fig. S8A). Upon differentiation to MT, 822 genes exhibited a twofold (or more) increase in expression in WT cells as measured by RNA-seq, while ~60% of these genes failed to properly induce in HDGF2 KO cells (Fig. 4C and fig. S8, B to D). These later genes exhibited increased HDGF2 binding upon differentiation in WT cells, as opposed to genes exhibiting normal induction in the HDGF2 KO cells (Fig. 4D). In contrast, LEDGF exhibited a decrease in its global binding upon differentiation (Fig. 4, B and D). An example of this phenomenon is observed at the myosin heavy chain (MHC) cluster that contains genes induced during differentiation to MT (Fig. 4E). Upon differentiation, HDGF2 levels increased in the induced genes within the cluster, whereas LEDGF binding decreased uniformly across the entire cluster. SPT16 binding was undetectable on this cluster in both the MB and MT. The induced expression of these genes was dependent on HDGF2 as their induction was not detected by RNA-seq in HDGF2 KO MT as compared to WT MT (Fig. 4E, bottom tracks). These data strongly suggest that in MTs, which contain undetectable levels of FACT and low levels of LEDGF, the binding of HDGF2 to chromatin is required for induction of its target genes.

We observed that most genes being expressed in both the MB and MT were enriched in HDGF2 at both stages. An example of one such highly expressed gene [NES (Nestin)] that is dependent on HDGF2 is shown in Fig. 4F. However, RNA-seq analysis revealed many such genes that are not dependent on HDGF2 (fig. S8, B and C). Therefore, the low levels of LEDGF in MTs and additional factors that have FACT-like activity probably exist for continued expression of these genes. One such additional factor may be nucleosome destabilizing factor (NDF), which was recently identified in Drosophila embryo extract through a screen for factors that stimulate acetylation of nucleosomes in vitro by the acetyltransferase p300 (37). Notably, NDF (CG4747) and the fly homolog of LEDGF/HDGF2 (CG7946) were originally identified together by ChiP-MS as proteins that enrich with actively transcribed chromatin containing H3K36me3 (33).

Similar to 293T cells, we observed an opposition between H3K27me3 with H3K36me2 and LEDGF and HDGF2 in MB (Fig. 2C and fig. S4, D and E to G). Upon differentiation to MT, a silent gene (C20orf166) in the MB located within an H3K27me3 domain is induced, and concurrently, we observed a loss in H3K27me3 with an accumulation of HDGF2 (Fig. 4G). The expression of this gene was dependent on HDGF2, suggesting an interplay between HDGF2 and H3K27me3 domains during differentiation. A more comprehensive bioinformatics analysis of all of our ChiP-seq data in 293T, mESC, MB, and MT cells is presented in the Supplemental Materials (figs. S3 and S9).

To better understand how HDGF2 affects transcription in cells, we used precision nuclear run-on sequencing (PRO-seq) in the WT and HDGF2 KO MB cells. PRO-seq precisely maps the level of engaged RNAPII on genes as mature transcripts are washed away during the nuclear isolation process (38). Metagene profiles of PRO-seq data revealed an increased RNAPII occupancy within the promoter proximal region of many HDGF2 target genes in the HDGF2 KO as compared to WT cells (Fig. 4H). Notably, the increased RNAPII peaks observed on genes in HDGF2 KO cells are located ~100-base pair (bp) upstream from the center of the +1 nucleosome, as mapped with assay for transspose-accessible chromatin sequencing data (Fig. 4I) (39, 40). These data illustrate that in the absence of HDGF2, RNAPII encounters a block to elongation due to the presence of a nucleosome (in this case, the +1) and indicates that HDGF2 does not affect initiation, pause release, or promoter escape but simply the release of RNAPII from the nucleosome-induced blockade to transcription (41). Notably, this +1 nucleosome–induced RNAPII pausing phenotype was observed in two independent HDGF2 KO MB cell lines, both of which could be rescued by stable lentiviral expression of HDGF2, demonstrating that the effect seen is due to HDGF2 elimination (Fig. 4, H and I, right, and fig. S8E).

**DISCUSSION**

In conclusion, we have identified two factors, LEDGF and HDGF2, as proteins that allow RNAPII to overcome the nucleosome-induced barrier to transcription elongation in differentiated cells that no longer express FACT (Fig. 4). During cellular differentiation, histone modifications reorganize, leading to distinct cell types with unique transcriptional profiles. In light of our findings, we propose that these proteins (LEDGF, HDGF2, and NDF) reorganize with histone modifications to maintain chromatin in a transcriptional competent state, hence overriding the requirement for FACT and helping to sustain the unique transcriptional profiles of particular cell types.

**MATERIALS AND METHODS**

**Purification and identification of LEDGF (PSIP1) from HeLa nuclear extract**

HeLa nuclear protein extract (4 g) was prepared as described in (42). Nuclear extract was dialyzed against BC100: Buffer C (BC) (pH 7.5) 100 mM KCl [20 mM tris-HCl, 20 mM β-mercaptoethanol, 0.2 mM phenylmethylsulfonyl fluoride (PMSF), 0.2 mM EDTA, 10% glycerol (v/v), and 100 mM KCl]. The number after BC denotes the salt concentration of the buffer. The extract was then loaded onto a phosphocellulose column and sequentially eluted with BC buffer containing 0.3, 0.5, and 1 M KCl. The novel FACT-like activity was eluted in the 1 M KCl (BC1000) fraction. This fraction was dialyzed against BC100 and loaded on a DEAE-cellulose column and sequentially loaded with BC buffer containing 0.3, 0.5, and 1 M KCl. The novel FACT-like activity did not bind the DEAE-cellulose column and was collected in the flow-through fraction (BC100). The flow-through fraction was dialyzed against BA100: Buffer A (BA) (pH 7.5) 100 mM NaCl (20 mM Hepes, 20 mM β-mercaptoethanol, 0.2 mM PMSF, 0.2 mM EDTA, 10% glycerol (v/v), and 100 mM NaCl) and loaded onto a heparin-agarose column. The column was washed with BA100 and eluted with BA600. The FACT–like activity was eluted in BA600 fraction, which was then dialyzed against BC100 and loaded onto a Q-Sepharose column. The column was eluted sequentially with BC buffer containing 0.25 and 0.5 M KCl. The 0.25 M (BC250) fraction was dialyzed against BA100 and loaded onto a heparin-agarose column. The heparin-agarose column was washed with BA100 and eluted sequentially with 0.3, 0.5, 0.7, and 1 M NaCl in BA buffer. The novel activity eluted in the 0.7 (BC700) fraction. A portion of this fraction was then loaded on a Superdex 200 (gel filtration column) that was equilibrated and run in BC100.
Fig. 4. LEDGF and HDGF2 substitute for FACT in differentiated cells. (A) Western blots of whole-cell protein extracts from MBs differentiated to MTs obtained from days 0 (D0), 3 (D3), and 6 (D6) using the antibodies indicated. (B) Average density ChIP-seq profiles for HDGF2, LEDGF, and SPT16 based on genes exhibiting RNAPII binding in MB (top) and MT (bottom). Levels based on genes with RNAPII binding: highest (purple; top 5%), high (blue; top 5 to 20%), medium (dark green; 20 to 80%), low (bright green; bottom 5 to 20%), and lowest (red; bottom 5%). (C) Venn diagram depicting the number of genes up-regulated twofold in MB to MT and fail to induce in the HDGF2 KO cells. (D) Average density ChIP-seq profiles for HDGF2 and LEDGF in MB and MT cells. Top: The 488 genes that are up-regulated in MB to MT and fail to induce in the HDGF2 KO cells. Bottom: The 334 genes that still induce in the HDGF2 KO. (E) Venn diagram depicting the number of genes up-regulated twofold in MB to MT and fail to induce in the HDGF2 KO cells. (F) Average density ChIP-seq profiles for HDGF2 binding to the MHC cluster (E), the NES locus (F), and the C20orf166 locus (G). (H) Metagene profiles of precision nuclear run-on sequencing (PRO-seq) RNA data plotted for the top 20% HDGF2-bound genes in WT and HDGF2 KO and HDGF2 KO rescue MB cell lines as indicated on the panels. (I) Plotted by 1 kb-pause site-1 kb-(scaled gene bodies)-1 kb-TES-1 kb. (J) Plotted with data centered on the TSS (+1 kb). First panel has assay for transposase-accessible chromatin sequencing (ATAC-seq) data overlayed with PRO-seq from WT cells showing the +1 nucleosome position (+1 nuc). (J) Schematic depiction showing the replacement of FACT by HDGF2 and/or LEDGF on chromatin during cellular differentiation, when FACT expression is reduced.
The activity was eluted from the gel filtration column with a mass range between 150 to 75 kDa.

**Chromatin assembly and chromatin transcription assays**

Chromatin assembly and chromatin transcription assays were performed as described in (10, 11).

**Naked DNA transcription assay and TFIIIS transcription assay**

Transcription on naked DNA template as shown in fig. S2A was performed with the same reagents [purified RNAPII and general transcription factors (GTFs)] as chromatin transcriptions except the DNA template used (PG5MPL-Gless) was not assembled with nucleosomes. For the TFIIIS assay in fig. S2B, the template used was 200 ng of a 563-bp polymerase chain reaction–amplified fragment from the puc18_MLP-601R (a gift from D. Luse). This template has a TATA box located 24-bp upstream of a well-defined TSS. There are no G's in the template strand until 21 bp following the TSS where there is a string of five G's. Transcription with our highly purified components (RNAPII and GTFs) was initiated with 2 mM adenosine 5′-triphosphate, cytidine 5′-triphosphate, and uridine 5′-triphosphate (UTP) and (5 μCi α32P UTP) in the absence of guanosine 5′-triphosphate (GTP), which allows the RNAPII to stall upon encountering the string of G’s at +21 bp. The elongation phase was then permitted for 30 min by adding a limiting concentration of GTP (0.1 μM). As there are many G’s in the transcribed strand, this limiting concentration of GTP causes RNAPII stalling along the template and allows TFIIIS activity to be detected. Elongation proceeds until RNAPII runs off the end of the template producing a 504-nucleotide RNA. Purified LEDGF, HDGF2, or TFIIIS was added during the elongation phase as indicated in the figure. The assay was stopped with 50 mM EDTA, and the products were separated on 7% polyacrylamide gel containing 7 M urea in tris-borate EDTA (TBE). The gels were visualized by autoradiography.

**Nucleosome transfer/chaperone assay**

The nucleosome transfer assay (Fig. 1H) was essentially performed as described in (10) and schematically depicted in Fig. 1G. The only minor difference was that we used the 601 nucleosome-positioning sequence instead of the 5S nucleosome-positioning sequence. The 190-bp fragment containing the 601 nucleosome-positioning sequence was isolated by digesting the p17-12-601 plasmid with the SCA1 restriction enzyme. The p17-12-601 plasmid contains 12 consecutive copies of the 601 sequence with an SCA1 site between each copy. The purified fragment was end-labeled with 32P as described in (10). For the assay, 15 ng of labeled 601 DNA fragment was incubated with 1 μg of purified oligonucleosomes in 10 mM Hepes (pH 7.5), 50 mM NaCl, 0.2 mM EDTA, 5 mM MgCl2, 5% glycerol, and recombinant HDGF2 as indicated. The assay was incubated for 3 hours at 37°C. At the end of the assay, the products were resolved on a 5% polyacrylamide gel cast and run in 0.25× TBE. The gels were dried and visualized by autoradiography. For Fig. S2C, the assay was performed identically to the assay in Fig. 1H, except that oligotetrasomes were used as the donor template. Oligotetrasomes were assembled on sonicated salmon sperm DNA (average size, 2 kb) by salt dialysis using recombinant H3/H4 dimers. The control nucleosome and tetrasome reference markers were assembled on the labeled 601 fragment with purified recombinant histones (H3/H4/H2A/H2B) for nucleosomes or (H3/H4) for tetrasomes by standard salt dialysis (2 M NaCl dialyzed down to 100 mM NaCl overnight).

**Recombinant LEDGF p75, LEDGF p55, HDGF2, and HDGF proteins**

The complementary DNAs (cDNAs) encoding the LEDGF p75, LEDGF p55, HDGF2, and HDGF proteins were cloned into the CBF vector (Addgene) at the SMA-1 site, which when expressed makes an N-terminal FLAG-tagged fusion protein. The LEDGF p75, LEDGF p55, HDGF2, and HDGF proteins were produced and purified with the same protocol as used for Brd2, Brd3, and Brd4 described in (10, 11).

**Native FLAG-LEDGF and FLAG-HDGF2 ChIP-seq and ChIP-MS**

Native FLAG-ChIPs for ChIP-seq and ChIP-MS were performed in 293T cells as described in (12), FLAG-LEDGF and FLAG-HDGF2 stable cell lines were produced using the pQCXIP Retroviral Vector (Clontech).

**Chromatin immunoprecipitation sequencing**

ChIP-seq experiments were performed as described before (31). Briefly, nuclei were isolated from cells fixed with 1% formaldehyde. Next, using a Diagenode Bioruptor, chromatin was fragmented into ~250 bp. ChIP was performed with the antibodies listed below. Chromatin from *Drosophila* (1:100 ratio to the experimental chromatin) and *Drosophila*-specific H2Av antibody was used as spike-in control in each sample. For ChIP-seq, libraries were prepared as described in (43) using 1 to 30 ng of immunoprecipitated DNA.

**RNA sequencing**

Total RNA from ESCs, EBs, 293T cells, MBs, and MTs was isolated with RNeasy (Qiagen) and reverse transcribed using SuperScript III and random hexamers (Life Technologies) to synthesize the first strand. Second strand was synthesized with deoxyuridine triphosphate to generate strand asymmetry using DNA polymerase I [M0209L, New England Biolabs (NEB)] and the *Escherichia coli* Ligase (L6090L, Enzymatics). RNA-seq libraries were constructed using the protocol described in (43).

**Antibodies**

The following antibodies are used in these studies:

- LEDGF (Proteintech) Rabbit Polyclonal, catalog number 25504-1-AP.
- HDGF2 (Proteintech) Rabbit Polyclonal, catalog number 15134-1-AP.
- SPT16 (Cell Signaling Technology) Rabbit Polyclonal D712K, catalog number 12191.
- SSRP1 (Abcam) Mouse Monoclonal 10D7, catalog number ab26212.
- Glyceraldehyde-3-phosphate dehydrogenase (Cell Signaling Technology) Rabbit Monoclonal 14C10, catalog number 2118.
- MHC/MyH1 (Developmental Studies Hybridoma Bank) Mouse Monoclonal MF 20.
- H3K27me3 (Cell Signaling Technology) Rabbit Monoclonal C36B11, catalog number 9733.
- H3K36me2 (Cell Signaling Technology) Rabbit Monoclonal C75H12, catalog number 2901.
- H3K36me3 (Abcam) Rabbit polyclonal, catalog number ab9050.
- RNAPII (Santa Cruz Biotechnology) Rabbit Polyclonal N-20, catalog number sc-899.
- Anti-FLAG (Sigma-Aldrich) M2 Agarose gel, Mouse Monoclonal M2, catalog number A2220.
- Histone H3 (Abcam) Rabbit polyclonal, catalog number ab1791.
- H2Av (Active Motif) Rabbit Polyclonal, catalog number 39715.

**mESC culture and differentiation**

E14Tga2 (CRL-1821, American Type Culture Collection) ESCs were grown in standard medium supplemented with leukemia inhibited factor (LIF), 1 μM mitogen-activated protein kinase 1/2 inhibitor.
(PD0325901, Stemgent), and 3 μM glycogen synthase kinase 3 inhibitor (CHIR99021, Stemgent). For EB differentiation, 600 k mESCs were plated in suspension plates with medium containing Dulbecco’s modified Eagle’s medium (DMEM) (Life Technologies), 20% fetal bovine serum (FBS), 1% non-essential amino acids (NEAA), 1% penicillin-streptomycin, 2 mM l-glutamine, and 100 mM ascorbic acid (Sigma-Aldrich). After 5 days, EB colonies were collected for downstream applications.

**MB cell culture and differentiation to MTs**

Human MBs were obtained as a gift from the Blau Laboratory (Stanford University). Cells were seeded on gelatinized plates and grown in Ham’s F-10 media (Gibco/Thermo Fisher Scientific) supplemented with 15% FBS (Atlanta Biologicals), 1 mM sodium pyruvate (Sigma-Aldrich), penicillin-streptomycin (100 μg/ml) (Gibco/Thermo Fisher Scientific), basic human fibroblast growth factor (2.5 ng/ml) (Promega), 1× GlutaMAX (Thermo Fisher Scientific), and 1 μM dexamethasone. Once 90% confluent, the MBs were differentiated to MTs by changing the medium to DMEM (Gibco/Thermo Fisher Scientific) supplemented with 2% horse serum (HyClone) and penicillin-streptomycin (100 μg/ml) (Gibco/Thermo Fisher Scientific). Differentiation proceeds for 6 days, and the efficiency can be monitored with microscopy as MTs are easily visualized detected.

**293T cell culture**

293T cells were cultured as described in (12).

**CRISPR genome editing**

gRNAs were designed using CRISPR design tool in https://benchling.com. All gRNAs below were cloned in pSpCas9(BB)-2A-GFP (plasmid 48138, Addgene) or to a lentiviral plasmid pLKO.1-puro U6 sgRNA (plasmid 50920, Addgene). The gRNAs were transfected into mESCs using Lipofectamine 2000 (Life Technologies) or infected into MBs together with lentiCas9–enhanced green fluorescent protein (plasmid 63592, Addgene). Single clones from green fluorescent protein (GFP)-positive and/or puromycin-resistant cells were genotyped and confirmed by sequencing.

**Guide RNAs**

The following gRNAs are used to knock out human HDGF2: hsHDGF2-gRNA-KO1, GCCACACGCCCTCAAGGCCG; hsHDGF2-gRNA-KO2, CACCGGCCCCATCTCCCG. The following gRNAs are used to knock out mouse HDGF2: msHDGF2-gRNA-KO1, TACCATCCAGTTACTTAGGG; msHDGF2-gRNA-KO2, GCTTACATTGGAACATCCTA. The following gRNAs are used to knock out mouse LEDGF:

msLEDGF-gRNA-KO1, TTTTCAATGGGACGTCTAGG; msLEDGF-gRNA-KO2, GGTCTCATTTGGACATCC.

**HDGF2 rescued MB cell lines**

Clonal HDGF2 KO MB cell lines were rescued with the lentivirus pLV-EF1a-IRESBlast that had the HDGF2 cDNA inserted within the multiple cloning sites (BamH1 and EcoR1). Following infection, cells were selected for blasticidin resistance. The HDGF2 cDNA was modified to make it resistant to the CRISPR gRNAs used to knock out HDGF2.

**Tissue expression**

The levels of expression of FACT subunits (SSRP1 and SPT16) shown in fig. S1A were obtained with the Expression Atlas (www.ebi.ac.uk/gxa/home) and ordered by hierarchical clustering.

**Data processing**

All samples were sequenced with either an Illumina HiSeq or NextSeq. The adapters were removed by the sequencing facility, and the quality of sequencing was assessed with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Reads having less than 80% of quality scores above 25 were removed with NGS QC Toolkit v2.3.3 (44) using the command IlluQC.pl. -se $file_path N A -l $percent -s $threshold -p $nb_cumu -o $out_path. Human hg19 and mouse mm10 from Illumina iGenomes University of California, Santa Cruz (UCSC) collection were used. ChIP-seq data were aligned with Bowtie v1.0.0 (45), allowing three mismatches and keeping uniquely aligned reads (bowtie -q -v 3 -p $nb_cumu -m 1 -k 1 -b -sam -se $1 bowtie_index_path "$file_path"). RNA-seq data were aligned with TopHat v2.0.9 (46), allowing three mismatches (tophat -N 3 --bowtie1 -o $out_path -p $nb_cumu $bowtie_index_path $file_path). Sam outputs were converted to Bam with SAMTools v1.0.6 (47) (samtools view -S -b $file_path -o $out_path/$file_name.bam) and sorted with Picard Tools v1.88 (http://broadinstitute.github.io/picard; java -jar SortSam.jar O = Sorted bam = input_bam). Data were further processed with Pasha (48) with the following parameters: WIGvs = TRUE, incrArtefactThr = every = 700000 for ChIP-seq and NA for RNA-seq, elongationSize = NA for ChIP-seq and 0 for RNA-seq. Input subtraction and scaling were performed with the function normAndSubtractWIG for 293TRex ChIP-seq. Data in mESC-EB and MB/MT differentiation system were spliced-in scaled with ChIPseqSplice (49). Fixed steps wiggle files were converted to bigwigs with the script wigToBigWig available on the UCSC Genome Browser website (http://hgdownload.soe.ucsc.edu/admin/exe/).

**Venn diagrams and bar graphs**

Signal was first detected with macs2 (50) (macs2 callpeak -t $bam_file_vector -c $input_file_vector -n $experiment_name --outdir $output_folder_nomodel_broad -f $format -g $genome_size -s $tag_size --nomodel --extsize $elongation_size --keep-dup $artefact_threshold --broad --broad-cutoff $qvalue). The parameters --extsize and --keep-dup were determined from the Pasha output log. Overlap was determined with the findOverlapsOfPeaks function of the ChIPpeakAnno (51) package (Figs. 1A, 2A, 3E, and 4C and figs. S8C and S9A).

**Heatmaps**

The heatmap of fig. S3A was obtained using a K-means clustering on SPT16 and LEDGF with six groups using the Bioconductor Package sepplots (52). Intervals are centered on the maximum polymerase II signal ≥10 kb and correspond to a macs2 peak detection (31,408 peaks) with broad mod and q < 0.03 (see “Venn diagrams and bar graph” section for details). Other marks are plotted on the above defined loci. The heatmaps in fig. S9 were generated with deepTools on galaxy centering on HDGF2 peaks for genes bound by RNAPII. The functions computeMatrix and plotHeatmap v 3.0.2.0 were used.

**Correlation plot**

Figure S4D is a Spearman’s correlation on the top 20% bound loci composed of the union of H3K27me3 and H3K36me2 (24,289) peaks. The peaks were obtained with macs2 in broad mode with q < 0.04.
Differential binding analysis
In mESC to EB differentiation system, the differential binding analysis (Fig. 3C) was performed with DiffBind (53) on RefSeq gene annotations. The percentages were calculated by computing the overlap of the different up-bound genes for each mark.

Differential expression
Differentially expressed genes were obtained with DESeq2 (Figs. 3, D and E, and 4C and fig. S8B) (54).

Metagene profiles
Metagene profiles (Figs. 3G and 4D and figs. S1B and S9E) were performed with the Bioconductor Package seqplots. Mean, 95% confidence intervals, and SE are indicated on the profiles. The meta-profiles of Fig. 4B and fig. S6 were obtained with in-house scripts that divide gene mean values into highest (purple; top 5%), high (blue; top 5 to 20%), medium (dark green; 20 to 80%), low (bright green; bottom 5 to 20%), and lowest (red; bottom 5%) bound genes.

RNA Pol II pausing
The violin plot of fig. S3C shows the pausing of polymerase II for the genes present in the different categories defined on the heatmap of fig. S3A. The pausing index was calculated as being the ratio of the mean values at TSS - 300 bp + 100 bp and 50% of gene body to transcription end site (TES). The number of genes for the different categories in fig. S3A (top to bottom) is as follows: 3159, 1492, 487, 2002, 4288, and 2013, respectively.

Housekeeping versus tissue specificity expression
In fig. S3D, the housekeeping feature of each gene of the different categories defined in fig. S3A was calculated using the tau metrics (55) using a threshold of 0.15 on the tissues as described in (56). Similar results were found using (57, 58) (data not shown).

Hidden Markov model
The Markov model of fig. S4A was calculated using ChromHMM v1.12 (59) with four states and a window size of 80 kb.

PRO-seq library preparation and sequence alignment
PRO-seq was performed according to the protocol (38) with minor modifications. In the nuclear run-on reaction using MB with spike-in Drosophila S2, all four biotinylated nucleotides were used at 25 μM each final concentration. RpPH (RNA 5’ pyrophosphohydrolase) (NEB) was used to remove the 5’ RNA cap. DNA Libraries were size-selected by AMPure PX beads (Beckman Coulter) and sequenced on a NextSeq 500. Adaptors were removed from raw reads by cutadapt 1.14 (60). Reads were trimmed from the 3’ end with removing low-quality bases using Trimmomatic 0.33 (61), requiring a read length of 16 to 36 bp. Reads derived from ribosomal RNA were filtered out by mapping reads on human and fly ribosomal DNA. The remaining reads were mapped on human hg19 or fly dm3 genome using Bowtie 1.1.2 (45) with options -m 1 -v 2. The 5’ ends of the reads were taken using bedtools genomcvc 2.25 (62) with options -strand -bg -s, and strands of reads were then swapped. Read counts were normalized by spike-in aligned reads. Normalized read counts were divided by total aligned million reads of WT sample. Metagene profiles of PRO-seq were generated by deepTools 3.0.0 computeMatrix and plotProfile (63).

Assay for transposase-accessible chromatin sequencing
The nucleosome profile shown in Fig. 41 was generated with NucleoATAC (40) using default parameters on SRR6652591 (39).

SUPPLEMENTARY MATERIALS
Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/10/eaay3068/DC1
Fig. S1. Identification of a novel FACT-like chromatin transcription chaperone.
Fig. S2. In vitro transcription assays.
Fig. S3. Bioinformatics analysis of genes within categories defined by levels of FACT, LEDGF, and HDGF2 binding.
Fig. S4. HDGF2/LEDGF colocalize with H3K36me2 and moderate levels of RNAPII excluding H3K27me3 domains.
Fig. S5. Validation of LEDGF and HDGF2 antibodies.
Fig. S6. Genes in MB and MT are categorized on the basis of RNAPII occupancy.
Fig. S7. Highly transcribed genes with SPT16 (FACT) occupancy in MBs.
Fig. S8. HDGF2 is required to activate MT-specific genes.
In the mESC to EB differentiation system, the differential binding analysis (Fig. 3C) was performed with DiffBind (53) on RefSeq gene annotations. The percentages were calculated by computing the overlap of the different up-bound genes for each mark.

REFERENCES AND NOTES
1. A. Mayer, H. M. Landry, L. S. Churchman, Pause & go: From the discovery of RNA polymerase pausing to its functional implications. Curr. Opin. Cell Biol. 46, 72–80 (2017).
2. J. Shandilya, S. G. Roberts, The transcription cycle in eukaryotes: From productive initiation to RNA polymerase II recycling. Biochem. Biophys. Acta 1819, 391–400 (2012).
3. A. C. Chiu, H. I. Suzuki, X. Wu, D. B. Mahat, A. J. Kriz, P. A. Sharp, Transcriptional pause sites delineate stable nucleosome-associated premature polyadenylation suppressed by U1 snRNP. Mol. Cell 69, 648–663.e7 (2018).
4. C. H. Chang, D. S. Luse, The H3/H4 tetramer blocks transcript elongation by RNA polymerase II in vitro. J. Biol. Chem. 272, 23427–23434 (1997).
5. G. LeRoy, G. Orphanides, W. S. Lane, D. Reinhberg. Requirement of RSF and FACT for transcription of chromatin templates in vitro. Science 282, 1990–1994 (1998).
6. G. Orphanides, G. LeRoy, C. H. Chang, D. S. Luse, D. Reinhberg. FACT, a factor that facilitates transcript elongation through nucleosomes. Cell 92, 105–116 (1998).
7. G. Orphanides, W. H. Wu, W. S. Lane, M. Hampsey, D. Reinhberg. The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSR1 proteins. Nature 400, 284–288 (1999).
8. H. Garcia, D. Fleyshman, K. Kolesnikova, A. Safina, M. Commanne, G. Paszkiewicz, A. Omelian, C. Morrison, K. Gurova, Expression of FACT in mammalian tissues suggests its role in maintaining of undifferentiated state of cells. Oncotarget 2, 793–796 (2011).
9. D. Fleyshman, L. Prendergast, A. Safina, G. Paszkiewicz, M. Commanne, K. Morgan, K. Attwood, K. Gurova, Level of FACT defines the transcriptional landscape and aggressive phenotype of breast cancer cells. Oncotarget 8, 20525–20542 (2017).
10. G. LeRoy, B. Rickards, S. J. Flint, The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. Mol. Cell 30, 51–60 (2008).
11. T. Kamno, Y. Kamno, G. LeRoy, E. Campos, H. W. Sun, S. R. Brooks, G. Vahedi, T. D. Heightman, B. A. Garcia, D. Reinhberg, U. Siebenlist, J. J. O'Shea, K. Ozato, BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. Nat. Struct. Mol. Biol. 21, 1047–1057 (2014).
12. G. LeRoy, I. Chepelov, P. A. DiMaggio, M. A. Blanco, B. M. Zee, K. Zhao, B. A. Garcia, Proteogenomic characterization and mapping of nucleosomes decoded by Brd and HP1 proteins. Genome Biol. 13, R68 (2012).
13. B. Rickards, S. J. Flint, M. D. Cole, G. LeRoy, Nucleolin is required for RNA polymerase I transcription in vivo. Mol. Cell. Biol. 27, 937–948 (2007).
14. H. Ge, Y. Si, R. G. Roeder, Isolation of CDNA encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. EMBO J. 17, 6723–6729 (1998).
15. D. P. Singh, A. Kimura, L. T. Chylack Jr., T. Shinohara, Lens epithelium-derived growth factor (LEDGF/p75) and p52 are derived from a single gene by alternative splicing. Genes 242, 265–273 (2000).
16. R. van Nuland, F. M. A. van Schaik, M. Simonis, S. van Heesch, E. Cuppen, R. Boelens, H. T. Timmers, H. van Ingen, Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. Epigenetics Chromatin 6, 12 (2013).
17. R. Schrijvers, S. Vets, J. de Rijck, N. Malani, F. D. Bushman, Z. Debyser, R. Gijbers, HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. Retrovirology 9, 84 (2012).
regulators serve as
A. and
Nat. Struct. Mol. Biol. modulation ChIP-mass spectrometry are linked to
B. Schambach, P. Dey, F. Cheropanov, M. Kvaratskhelia, S. H. Hughes, A. Engelman, HRP2 determines the efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor. Nucleic Acids Res. 40, 11518–11530 (2012).

De Rijck, J. Kogel, J. Demeulemeester, S. Vets, S. El Ashkar, N. Malani, F. D. Bushman, B. Landuyt, S. J. Husson, K. Busschots, R. Gijbers, J. Debyser, The BET family of proteins targets monoloye murine leukemia virus integration near transcription start sites. Cell Rep. 5, 886 (2013).

Gupta, T. Maetzig, G. N. Maertens, A. Sharif, M. Rothe, M. Weidner-Glunde, M. Galla, A. Schambach, P. Cheropanov, T. F. Schulz, Bromo- and extraterminal domain chromatin regulators serve as cofactors for murine leukemia virus integration. J. Virol. 87, 12721–12736 (2013).

De Rijck, J. Kogel, J. Demeulemeester, S. Vets, S. El Ashkar, N. Malani, F. D. Bushman, B. Landuyt, S. J. Husson, K. Busschots, R. Gijbers, J. Debyser, The BET family of proteins targets monoloye murine leukemia virus integration near transcription start sites. Cell Rep. 5, 886 (2013).

Gijbers, S. Vets, J. de Rijck, K. E. Ocevjea, K. Ronen, N. Malani, F. D. Bushman, Z. Debyeer, Role of the PWWP domain of lens epithelium-derived growth factor (LEDFG)/p75 cofactor in lentiviral integration targeting. J. Biol. Chem. 286, 41812–41825 (2011).

Guerman, V. B. Balan, A. J. Tackett, B. T. Chait, R. G. Roeder, Synergistic functions of SII and p300 in productive activator-dependent transcription of chromatin templates. Cell 125, 275–286 (2006).

J. K. McElber, L. L. McCullough, F. G. Whitby, T. Formosa, C. P. Hill, FACT disrupts nucleosome structure by binding H2A-H2B with conserved peptide motifs. Mol. Cell 20, 294–304 (2005).

M. D. Simon, F. Chu, L. R. Racki, C. C. de la Cruz, A. L. Burlinge, B. Panning, G. J. Narlikar, K. M. Shokat, The site-specific installation of methyl-lysine analogs into recombinant histones. Cell 128, 1003–1012 (2007).

O. Okusu, V. Narenda, C. H. Lee, N. Descotes, G. LeRoy, R. Raviram, L. Blumberg, K. Karch, P. P. Rocha, B. A. Garcia, J. A. Skok, D. Reinberg, Capturing the onset of PR2-mediated repressive domain formation. Mol. Cell 70, 1149–1162.e5 (2018).

Y. Tanaka, Z. Katagiri, K. Kawaihashi, D. Kioussis, S. Kitajima, Trithorax-group protein ASH1 methylates histone H3 lysine 36. Gene 397, 161–168 (2007).

C. I. Wang, A. A. Alekseyenko, G. LeRoy, A. E. H. Elia, A. A. Gorchakov, L. M. P. Britton, S. J. Elledge, P. V. Kharchenko, B. A. Garcia, M. I. Kuroda, Chromatin proteins captured by Chip-mass spectrometry are linked to dosage compensation in Drosophila. Nat. Struct. Mol. Biol. 20, 202–209 (2013).

D. A. Orlando, M. W. Chen, V. E. Brown, S. Solanki, Y. J. Choi, E. R. Olson, C. C. Fritz, J. E. Bradner, M. G. Guenther, Quantitative Chip-seq normalization reveals global modulation of the genome. Cell Rep. 9, 1163–1170 (2014).

H. M. Blau, C. Webster, Isolation and characterization of human muscle cells. Proc. Natl. Acad. Sci. U.S.A. 78, 5623–5627 (1981).

C. Webster, G. K. Pavlath, D. R. Parks, F. S. Walsh, H. M. Blau, Isolation of human myoblasts with the fluorescence-activated cell sorter. Exp. Cell Res. 174, 252–265 (1988).

J. Fei, H. Ishii, M. A. Hoeksema, F. Meitinger, G. A. Kassavetis, K. C. Glass, R. Ben, J. T. Kadonaga, NDF, a nucleosome-destabilizing factor that facilitates transcription through nucleosomes. Genes Dev. 32, 682–694 (2018).

D. B. Mahat, H. Kwak, G. T. Booth, I. H. Jonkers, C. G. Danko, R. K. Patel, C. T. Waters, K. Munson, L. J. Core, J. T. Lis, Base-pair-resolution genome-wide mapping of active RNA polymerases using precision needle run-on (pro-seq). Nat. Protoc. 11, 1455–1476 (2016).

H. A. Pinter, J. S. Packer, J. L. McFalone-Figueroa, D. A. Cusanovich, R. M. Daza, D. Aghamirizadeh, S. Srivastav, X. Qiu, D. Jackson, A. Minkina, A. C. Adey, J. Steeners, J. Shendure, C. Trapnell, Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. Mol. Cell 71, 858–871.e8 (2018).
Acknowledgments: We thank L. Vales for critical guidance and reading of the manuscript. We thank other members (past and present) of Reinberg Laboratory for discussion as the work was in progress. We are also grateful to D. Hernandez for technical assistance. Human MBs were obtained as a gift from the Blau Laboratory (Stanford University) and thanks to S.A. Marshall and E.J. Rendleman for help with PRO-seq NGS (Northwestern University).

Funding: This work was supported by grants to D.R. from NIH (R01CA199652) and HHMI. G.L. was partially supported by a grant from the Making Headway Foundation (189290). R.A.G. was supported by the Swedish Society for Medical Research. H.O.K. was supported by a fellowship from the NIH (F31HD090892). J.-R.Y. is supported by the American Cancer Society (PF-17-035-01). J.S. was a Simons Foundation's Junior Fellow and is also supported by an NIH (K99AA024837) grant. Author contributions: G.L., O.O., R.A.G., H.O.K., C.-H.L., J.-R.Y., J.S., and Y.A. conducted experiments. N.D. performed bioinformatics analyses. Y.A. and A.S. performed and analyzed PRO-seq experiments. G.L., O.O., N.D., and D.R. wrote the manuscript.

Competing interests: D.R. is a cofounder of Constellation Pharmaceuticals and Fulcrum Therapeutics. The authors declare that they have no other competing interests.

Data and materials availability: Genome sequencing results are available on GEO under accession number GSE117155 and will be available publicly upon acceptance. Additional data related to this paper may be requested from the authors.

Submitted 6 June 2019
Accepted 5 September 2019
Published 2 October 2019
10.1126/sciadv.aay3068

Citation: G. LeRoy, O. Oksuz, N. Descostes, Y. Aoi, R. A. Ganai, H. O. Kara, J.-R. Yu, C.-H. Lee, J. Stafford, A. Shilatifard, D. Reinberg, LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells. Sci. Adv. 5, eaay3068 (2019).
LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells

Gary LeRoy, Ozgur Oksuz, Nicolas Descostes, Yuki Aoi, Rais A. Ganai, Havva Ortabozkoyun Kara, Jia-Ray Yu, Chul-Hwan Lee, James Stafford, Ali Shilatifard and Danny Reinberg

Sci Adv 5 (10), eaay3068.
DOI: 10.1126/sciadv.aay3068