PNAS PLUS

Genome-wide targeting of the epigenetic regulatory protein CTCF to gene promoters by the transcription factor TFII-I

Rodrigo Peña-Hernández[a,1], Maud Marques[a,1], Khalid Hilmi[a,1], Teijun Zhao[a], Amine Saad[a], Moulay A. Alaoui-Jamali[a], Sonia V. del Rincon[a], Todd Ashworth[b], Ananda L. Roy[c], Beverly M. Emerson[d], and Michael Witcher[a,2]

[a]Departments of Oncology and Experimental Medicine, Lady Davis Institute, Jewish General Hospital, McGill University, Montreal, QC H3T 1E2, Canada; [b]Immunoscience Research Unit, Pfizer Inc., Cambridge, MA 02139; [c]Department of Developmental, Molecular and Chemical Biology, Tufts University School of Medicine, Boston, MA 02111; and [d]Department of Regulatory Biology, Salk Institute, La Jolla, CA 92037

Edited by Mark Groudine, Fred Hutchinson Cancer Research Center, Seattle, WA, and approved January 9, 2015 (received for review August 29, 2014)

CCCTC-binding factor (CTCF) is a key regulator of nuclear chromatin structure and gene regulation. The impact of CTCF on transcriptional output is highly varied, ranging from repression to transcriptional pausing and transactivation. The multifunctional nature of CTCF may be directed solely through remodeling chromatin architecture. However, another hypothesis is that the multifunctional nature of CTCF is mediated, in part, through differential association with protein partners having unique functions. Consistent with this hypothesis, our mass spectrometry analyses of CTCF interacting partners reveal a previously undefined association with the transcription factor general transcription factor II-I (TFII-I). Biochemical fractionation of CTCF indicates that a distinct CTCF complex incorporating TFII-I is assembled on DNA. Unexpectedly, we found that the interaction between CTCF and TFII-I is essential for directing CTCF to the promoter proximal regulatory regions of target genes across the genome, particularly at genes involved in metabolism. At genes coregulated by CTCF and TFII-I, we find knockdown of TFII-I results in diminished CTCF binding, lack of cyclin-dependent kinase 8 (CDK8) recruitment, and an attenuation of RNA polymerase II phosphorylation at serine 5. Phenotypically, knockdown of TFII-I alters the cellular response to metabolic stress. Our data indicate that TFII-I directs CTCF binding to target genes, and in turn the two proteins cooperate to recruit CDK8 and enhance transcription initiation.

CTCF | TFII-I | CDK8

Proper epigenetic programming is essential for cell and system homeostasis. Disruption of this code through aberrant histone modification and DNA methylation can result in a transition toward numerous disease states. The epigenetic regulatory protein CCCTC-binding factor (CTCF) has been shown to coordinate epigenetic processes on multiple levels, including nucleosome arrangement, histone modification, and DNA methylation, as well as mediating chromosomal interactions *in cis* and *in trans* (1–3). Additionally, CTCF has a more direct influence on transcription; 7–14% of CTCF sites are found within promoter proximal regions (4). Functionally, there is evidence that CTCF acts primarily as a transcriptional activator in these cases, and it is unclear that its transactivating capabilities are dependent on epigenetic processes (5–8); when bound within exonic regions, it promotes polymerase II (Pol II) pausing (9), adding another level of transcriptional regulation to its repertoire.

How can one protein participate in such disparate cellular activities? One theory is that through modulating epigenetic marks and mediating chromosomal loops, CTCF may act upon all of these processes (10, 11). Another idea is that CTCF forms multiple, distinct protein interactions within the nucleus, and these protein interactions mediate the disparate biological outputs (12, 13). Further, it is possible that proteins associated with CTCF mediate many of the epigenetic and chromosomal organization responsibilities of CTCF (14–16). For example, *trans*-activating factors, including the transcriptional repressor protein YY1 (Yy1) and Y box-binding protein 1 (Ybx1), have been demonstrated to act cooperatively with CTCF to regulate transcription (17, 18). CTCF also cooperates with cohesins and core transcription factors, such as TAF3, to mediate long-range chromatin interactions (19), and with nucleophosmin to spatially organize CTCF within the nucleus (20). In *Drosophila*, the intriguing observation was made that depletion of the CTCF partner centrosome-associated zinc finger protein (CP190) resulted in a reduced pool of CTCF associated with chromatin (21). An elegant study mapping the sequence specificity of the 11 zinc finger proteins of CTCF found that flanking sequences could influence CTCF binding through an unknown mechanism (22). These flanking sequences do not target CP190 because there is no human homolog. Together, these studies strongly suggest partner proteins may not only modulate CTCF function, but also cooperate with CTCF to define its binding sites. However, to date, proteins modulating the site-specific interaction of CTCF with chromatin in mammalian cells has been lacking.

Significance

CCCTC-binding factor (CTCF) is an epigenetic regulatory protein that is not only functionally diverse, but is also targeted to highly diverse DNA binding sites. CTCF cooperates with accessory proteins to achieve various functional outputs. Further evidence in *Drosophila* shows that CTCF may also be targeted to chromatin via accessory proteins. The identity of such mammalian proteins remains elusive. Herein, we describe evidence that the transcription factor general transcription factor II-I (TFII-I) targets CTCF binding to metabolism-related genes across the genome. We find that TFII-I regulates the transcription of genes within this network on the level of initiation via RNA polymerase II phosphorylation. These results provide a starting point for understanding a biological network communicating information between chromatin architecture, transcription, and metabolism.

Author contributions: R.P.-H., M.M., K.H., and M.W. designed research; R.P.-H., M.M., K.H., A.S., S.V.d.R., and M.W. performed research; T.Z., M.A.A.-J., T.A., A.L.R., and B.M.E. contributed new reagents/analytic tools; R.P.-H., M.M., K.H., and M.W. analyzed data; and R.P.-H., M.M., K.H., and M.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession nos. GSE60915, GSE60917, and GSE60918).

[1]R.P.-H., M.M., and K.H. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: michael.witcher@mcgill.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1416674112/-/DCSupplemental.

Here, CTCF immunopurification and biochemical fractionation were used to identify CTCF interacting partners. Among these, we focused the current study on defining the relevance of the interaction between CTCF and general transcription factor II-I (TFII-I). TFII-I is a basal transcriptional factor that has the ability to bind to core promoter elements such as pyrimidine-rich initiator (Inr) element sequences, as well as upstream regulatory elements (E-box) (23). It has been proposed that TFII-I can function both as a basal transcriptional factor and an activator that can interact with complexes assembled at upstream regulatory sites (24). In some contexts, the *trans*-activating activity of TFII-I is regulated by tyrosine phosphorylation in response to mitogenic signals (25, 26). Thus, TFII-I activity may provide a link between the extracellular environment and genes necessary to respond to the cellular milieu.

Using TFII-I knockdown (TFII-I KD) cells we show that TFII-I is responsible for targeting CTCF to promoter regions genome wide. This targeting is enriched at genes involved in metabolism. We conclude that coordinated gene regulation by TFII-I and CTCF is important for mediating transcription initiation at target genes, and phenotypically for mediating the antiproliferative response imposed by nutrient deprivation.

Results and Discussion

We previously observed strong binding between CTCF and known protein partners in MDA-MB-435 cells (27). Therefore, we decided to look for previously unidentified CTCF interacting proteins from these cell lysates, spiked with recombinant CTCF, using an affinity purification approach followed by mass spectrometry. For these experiments, conditions were sufficiently stringent that low background binding was observed in IgG controls by Coomassie blue staining (Fig. S1). However, the detection limit of Coomassie is 100 ng (28), so the data does not exclude the capture of contaminants by IgG controls below this threshold. Liquid chromatography/MS (LC/MS) of the purified proteins revealed several previously identified proteins (including YY1, Ybx1, TopoII, Npm, and Ubtf) and uncovered TFII-I as a novel CTCF binding partner (Fig. 1*A*). Binding to TFII-I was validated by coimmunoprecipitation (co-IP) and reverse co-IPs (Fig. 1 *B*–*D*) using lysates from cells of three tissue types (MDA-MB-435, HCT116, and WEHI-231). We observed a robust interaction between CTCF and TFII-I in all cell types tested, even in the presence of nucleases. Exposure of lysates to RNase did mitigate the interaction between CTCF and DEAD box protein 5 (DDX5) as previously reported (Fig. 1*C*) (16). CTCF bound the two most abundant TFII-I isoforms of TFII-I (β and Δ) with comparable affinity (Fig. S2). Deletion mutations of CTCF revealed multiple domains were capable of mediating the interaction between CTCF and TFII-I (Fig. S2). The zinc finger domain of CTCF does not interact with TFII-I, indicating this interaction is not solely dependent on binding to DNA, consistent with our Co-IP data incorporating nucleases. Similarly, a multidomain interaction profile was observed between CTCF and the transcription factor YY1 (17).

The interaction of CTCF with TFII-I was also supported using confocal microscopy to show colocalization of TFII-I and CTCF within MDA-MB-435 cells (Fig. 1*E*). Volocity software was used to quantify pixel colocalization within the two channels used to detect CTCF (green; Alexa Fluor 488) and TFII-I (red; Alexa Fluor 594) for Z-series confocal data (29). The Volocity-calculated Pearson correlation coefficient of 0.85 indicates a significant degree of 3D colocalization (1.0 = 100% overlap of signal).

We previously showed the CTCF binding partner TopoIIβ localized to only a subset of CTCF targets, suggesting the existence of distinct CTCF complexes (27). Consistent with a previous report (30), we find size-exclusion chromatography cannot separate CTCF into distinct complexes because CTCF-containing elutions are constrained to sizes above 1 MDa (Fig. S3); this suggests CTCF is integrated into at least one large multiprotein complex. To probe for biochemically distinct CTCF complexes, we fractionated CTCF from cell extracts using column chromatography, with the final step in our purification scheme using DNA cellulose as a binding matrix (Fig. 1*F*). This approach showed CTCF is indeed integrated into biochemically distinct complexes that have unique affinities for DNA. This biochemical separation also served to validate our mass spectrometry data, as TFII-I copurified with CTCF, primarily contained within a single CTCF complex.

Knockdown studies show CTCF primarily acts as a positive regulator of target genes (6), and we surmised that CTCF may cooperate with TFII-I to regulate the production of target transcripts. TFII-I was originally found as a component of the basal transcription machinery (24), but its binding at target genes may also be mediated by growth signals (26, 31), as is commonly seen with other *trans* activators. Thus, the interaction between TFII-I and CTCF represents a novel link between the extracellular environment and epigenetic organization.

To test the possibility that CTCF and TFII-I may cooperatively regulate expression of common target genes, we first identified genes whose expression are modulated by TFII-I through microarray analysis of mRNA from TFII-I KD cells (32) (Fig. 2*A*). A panel of the top hits from our microarray data was validated by conventional RT-quantitative PCR (RT-qPCR; Fig. 2*B*). We identified 500 genes differentially regulated between the control and the TFII-I KD cells (TOP500: fold change KD vs. CT > 1.6 and $P < 0.05$). Attesting to the specificity of this data, the TFII-I coding gene (*Gtf2i*) was the top gene down-regulated. To functionally categorize these genes, we performed functional annotation, Gene Ontology (GO) and pathway (KEGG) analysis. Of the genes regulated by TFII-I, over 75% are predicted to be involved in metabolism (Fig. 2*C*). KEGG pathway analysis also showed a significant enrichment for transcripts involved in metabolic processes, with 46 metabolic genes regulated by TFII-I being predicted to function in the same pathway (Fig. 2*D*; $P = 0.0002$).

One of the genes identified by microarray and down-regulated in TFII-I–depleted cells was the cyclin-dependent kinase inhibitor 2A ($Cdkn2a^{Arf}$) (Fig. 2*B*). This tumor suppressor gene was previously shown to be transcriptionally regulated by CTCF (27) and therefore might prove a suitable model to study cooperativity between TFII-I and CTCF. ChIP shows both proteins colocalize to the $Cdkn2a^{Arf}$ promoter region (Fig. 3*A*). As expected, TFII-I binding was lost in knockdown cells (Fig. 3*A*), but surprisingly, CTCF binding was significantly diminished as well ($P < 0.05$; Fig. 3*A*). This loss was not generalized to all CTCF binding sites in tumor suppressor genes, because binding at the zinc finger MYND domain-containing protein (*Zmynd10*) was undisturbed (Fig. 3*B*).

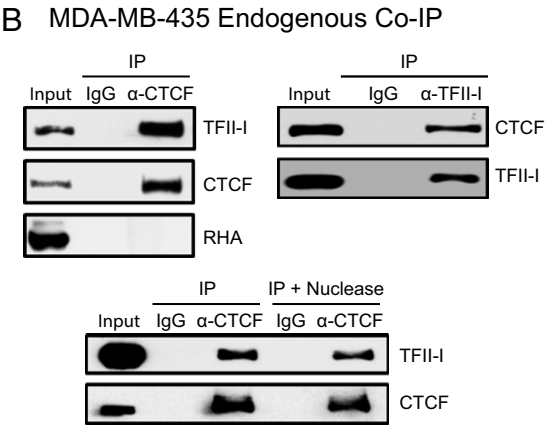
To ensure that depletion of TFII-I was directly responsible for the loss CTCF binding at $Cdkn2a^{Arf}$, we reconstituted our TFII-I knockdown cells with exogenous TFII-I (β and Δ isoforms) (26) and probed for CTCF binding by ChIP. CTCF binding was indeed restored at the $Cdkn2a^{Arf}$ promoter in the TFII-I knockdown complemented with exogenous TFII-I (Fig. 3*C*).

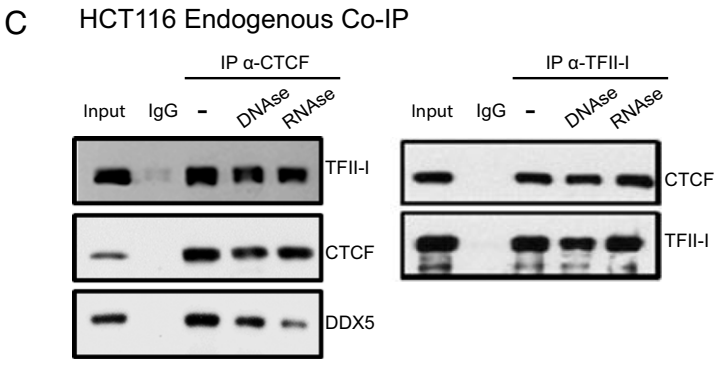
In Fig. 3*D*, we show that depletion of CTCF also regulates $Cdkn2a^{Arf}$ and other TFII-I target genes, including ketohexokinase (*Khk*), pyruvate dehydrogenase (*Pdhb*), and glutamate dehydrogenase 1 (*Glud1*), in a similar manner to what is observed post–TFII-I knockdown (Fig. 2*B*). Our data indicate CTCF binding is targeted by TFII-I and further support a cooperative role between CTCF and TFII-I to transcriptionally regulate a subset of genes.

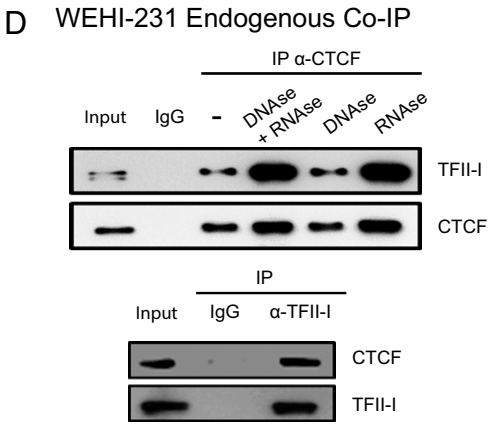
It was not technically feasible to quantify TFII-I binding to DNA upon CTCF knockdown because the loss of CTCF compromised the cellular pools of TFII-I (Fig. S4). This result indicates there is a feedback loop whereby CTCF and TFII-I regulate

A

Identity	Status	# Peptides	emPAI score	MASCOT score
TFII-I	Novel interacting protein	55	1.10	1176
NPM	Known interacting protein	56	2.76	718
UBTF	Known interacting protein	11	0.11	233
YB1	Known interacting protein	19	1.45	663
YY1	Known interacting protein	16	0.53	285
CKII	Known interacting protein	7	0.27	145
Topo II	Known interacting protein	10	0.09	198
CTCF	Bait protein	26	0.55	404







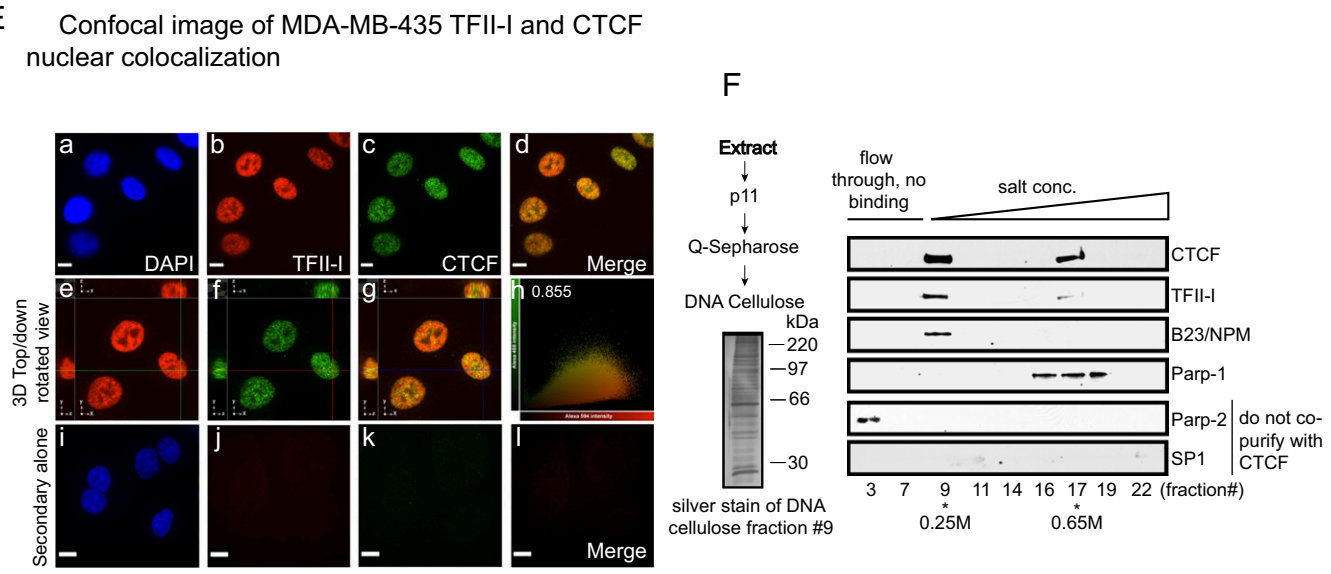


Fig. 1. CTCF interacts with transcription factor TFII-I. (*A*) Mass spectrometry analysis of CTCF interacting proteins from MDA-MB-435 cells reveals association with known cofactors and a novel partner, TFII-I. Peptide number, Mascot, and emPAI scores were calculated using Matrix Science software. (*B*) Forward and reverse co-IP of MDA-MB-435 extracts using anti-CTCF and TFII-I antibodies. Inputs represent 2% of immunoprecipitated material. Treatment of extracts with nucleases (DNase and RNase) showed the interaction between the two proteins is independent of nucleic acids. (*C*) Interaction between CTCF and TFII-I was assessed by co-IP (forward and reverse) using HCT116 lysates. DDX5 interacts with CTCF in an RNase-dependent manner. (*D*) Interaction between CTCF and TFII-I was assessed by co-IP (forward and reverse) using WEHI-231 lysates. (*E*) (*a–d*) MDA-435 cells were grown on coverslips, fixed, and stained for confocal microscopy. Antibodies against TFII-I (secondary, Alexa 594, red) and CTCF (secondary, Alexa 488, green) were used to bind endogenous proteins, followed by visualization using secondary antibodies. (Scale bar: 10 μm.) (*e–g*) Top-down and rotated view of stained cells. Large panel (*y–x*) is top-down, small panels (*y–z*, *Left*) and (*z–x*, *Top*) show view rotated along designated axis. The intersecting lines point to an example of colocalizing clusters. (Scale bar: axis, 3 μm.) (*h*) Volocity-generated scatter plot of CTCF and TFI-II intensities with calculated Pearson correlation coefficient of 0.855. (*i–l*) Same experiment as in *a–d* except with the omission of the primary antibodies. (Scale bar: 10 μm.) (*F*) Biochemical separation of CTCF-containing complexes by column chromatography. Immunoblotting for CTCF eluents demonstrate integration of CTCF into multiple protein complexes on DNA. (*Left*) Purification scheme and a representative silver-stained gel of the DNA cellulose purified fraction 9. (*Right*) Western blotting for proteins copurifying with CTCF after separation on a DNA cellulose matrix, as well as noneluted proteins (Parp-2 and SP1). Proteins copurifying with CTCF were eluted in fractions 9 and 17.

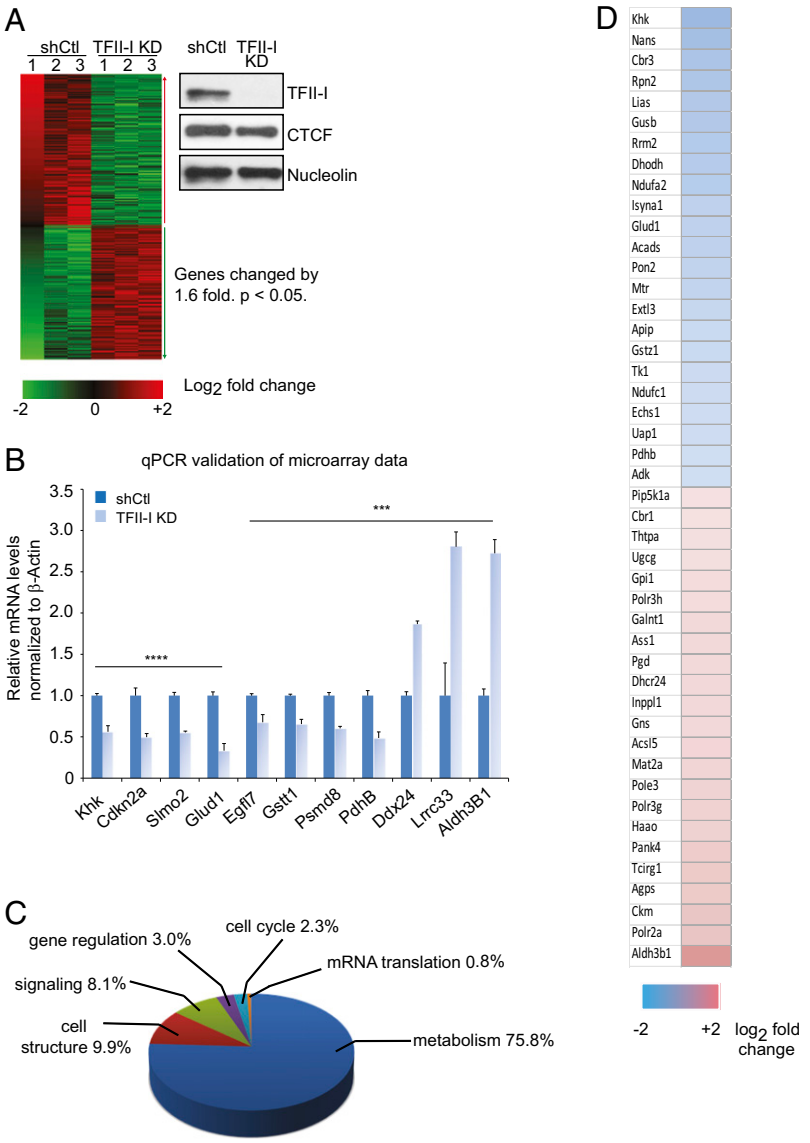


Fig. 2. Microarray profiling of genes regulated by TFII-I. (*A*, *Left*) Heat map of three independent WEHI-231 RNA samples from scrambled control (shctl) and TFII-I KD cells. Only those genes showing minimally a 1.6-fold change between the control and the TFII-I KD samples and $P < 0.05$ are represented. (*Right*) Western blot of TFII-I levels from ctl and TFII-I KD. (*B*) RT-qPCR validation of a subset of the genes identified differentially expressed between ctl and TFII-I–depleted cells. Relative mRNA levels were calculated by $2^{-\Delta CT}$ method and normalized to β-actin. Error bars represent SEM from at least three independent experiments (two-tailed Student *t* test, $^{***}P \leq 0.05$, $^{****}P \leq 0.01$). (*C*) Pie chart of GO analysis performed for the TOP500 genes (fold change >1.6 and $P < 0.05$). (*D*) Pathway analysis using KEGG was performed on the TOP500 genes differentially expressed. List of 46 genes found in a single metabolic pathway ($P = 0.00017$).

each other through complementary mechanisms. Loss of TFII-I results in altered targeting of CTCF to chromatin. Comparably, when CTCF expression is lost, TFII-I protein levels diminish. Thus, each protein is dependent on the other for full activity.

CTCF has epigenetic insulator activity, being capable of preventing the spread of repressive histone modifications such as H3K27me^3 (4, 33) and enabling the local accumulation of activating marks such as histone acetylation and the histone variant H2A.Z (27, 33, 34). Therefore, we investigated whether the diminished *Cdkn2a*^{Arf} expression and decreased CTCF binding we observed after TFII-I KD was concomitant with changes to the epigenetic landscape. No significant changes of the repressive mark H3K27me^3 or the activating marks H2A.Z and H3K27Ac were observed (Fig. 3*E*).

Eukaryotic transcription initiation is a dynamic and complex process requiring the association of general transcription factors to form a preinitiation complex (PIC) at promoter regions as an initial step (35). Once this complex has been established, and RNA Pol II is recruited, RNA Pol II is able to clear the proximal promoter and initiate mRNA synthesis after it is phosphorylated at serine 5 residue of the C-terminal domain (CTD) on the largest RNA Pol II subunit (36, 37). This phosphorylation event is actively removed and replaced by serine 2 phosphorylation during the elongation phase of transcription. CTCF has been implicated in the recruitment of RNA Pol II to target genes (5), and TFII-I is known as a transcription initiation factor (24), so we next looked at RNA Pol II recruitment to the *Cdkn2a*^{Arf} promoter. Total RNA Pol II association with the *Cdkn2a*^{Arf} promoter was consistent between the control and knockdown cells (Fig. 3*E*), suggesting the transcriptional regulation lies downstream of PIC formation and RNA Pol II recruitment. Analysis of serine 5 phosphorylation of RNA Pol II revealed a substantial decrease of this modification upon TFII-I KD (Fig. 3*E*). RNA Pol II phosphorylation on serine 5 is required for

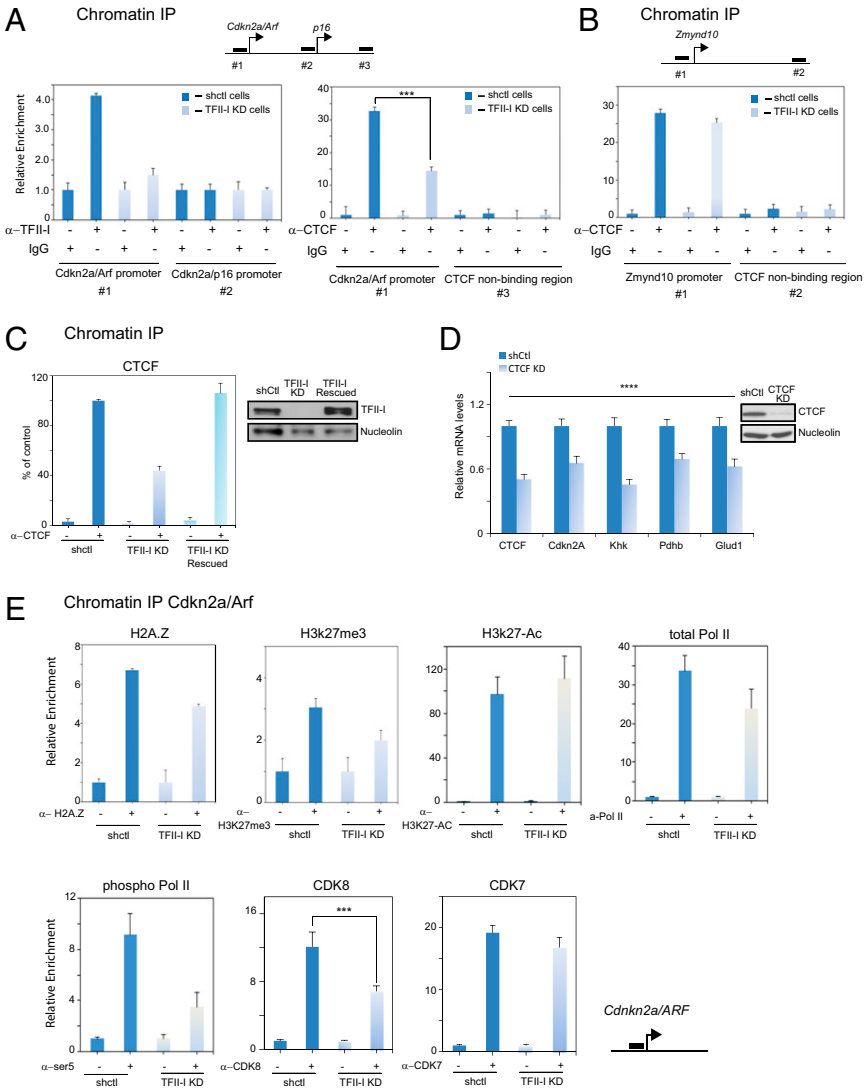


Fig. 3. TFII-I knockdown impairs CTCF binding at *Cdkn2a/Arf* promoter. ChIP analysis of TFII-I and CTCF at *Cdkn2a/Arf-p16* locus ($n = 3$, two-tailed Student *t* test, $***P \leq 0.05$; *A*) or *Zmynd10* gene ($n = 3$; *B*) in control (shctl) and (TFII-I KD) cells. (*C*) ChIP of CTCF at the *Cdkn2a/ARF* locus in control, TFII-I KD, and TFII-I KD + exogenous TFII-I. Western blot of TFII-I in the control, TFII-I KD, and TFII-I–complemented cells ($n = 3$). (*D*) RT-qPCR of TFII-I identified target genes in control and CTCF knockdown cells from at least three independent experiments. Relative mRNA levels were calculated by $2^{-\Delta CT}$ method and normalized to *Ftl1* (two-tailed Student *t* test, $****P \leq 0.01$). Western blot of CTCF in control and knockdown cells with and without exogenous TFII-I. (*E*) ChIP–qPCR of H2A.Z, H3K27me3, H3K27Ac, total RNA polymerase, Phospho-Ser5-RNA polymerase II, CDK7, and CDK8 ($n = 3$, two-tailed Student *t* test, $***P \leq 0.05$) at the *Cdkn2a/ARF* locus in control and TFII-I KD cells. Relative enrichment was calculated fold Ab over no Ab. For all graphs, error bars represent SEM.

transcription initiation, and its loss can explain the reduction in transcriptional output of the $Cdkn2a^{Arf}$ transcript.

Next, we wanted to identify the kinase responsible for RNA Pol II modification when $Cdkn2a^{Arf}$ is bound by TFII-I and CTCF. Serine 5 of the RNA Pol II CTD heptad repeat is primarily targeted by the cyclin-dependent kinase 7 (CDK7) and cyclin-dependent kinase 8 (CDK8) (38–42). Though we did not observe any change in association of CDK7 at the $Cdkn2a^{Arf}$ proximal promoter after TFII-I KD, CDK8 binding was clearly disrupted (Fig. 3*E*). We propose that CTCF and TFII-I cooperate to form a scaffolding complex required for the efficient recruitment of the CDK8 complex to the $Cdkn2a^{Arf}$ promoter. Because CTCF binds other core transcription factors, such as Taf3 (43), it is possible that CTCF and TFII-I integrate into a larger scaffolding complex at core promoter regions enabling the recruitment of CDK8. This may be similar, or identical, to the scaffolding complex previously shown to promote reinitiation (44).

To explore whether TFII-I might be involved in directing CTCF to binding sites genome wide, we carried out ChIP sequencing (ChIP-seq) experiments to evaluate CTCF binding to genomic DNA in control (ctl) and TFII-I KD cells. Of the 24,169 CTCF peaks identified in these experiments, 6,978 were lost in the absence of TFII-I ($P = 0.03$), consistent with the data we collected using the $Cdkn2a^{Arf}$ promoter as a model (Fig. 4 *A* and *B*). Distribution analysis revealed that CTCF was primarily displaced from promoter and proximal upstream regulatory regions (Fig. 4 *C* and *D*). Of the 3,986 CTCF sites located within ±3 kb of a transcription start site, 777 were lost, using a stringent cutoff of 3.7-fold loss of sequence tags at a given site (Fig. 4*E*). Visualization of our ChIP-seq data at the *Znf219* locus using the University of California Santa Cruz (UCSC) genome browser highlighted the specificity of CTCF at promoter regions. Here, CTCF sites were found at the 5′ regulatory region, the proximal promoter and within multiple exons. Of these, only CTCF binding at the proximal promoter was dependent on TFII-I (Fig. 4*F*).

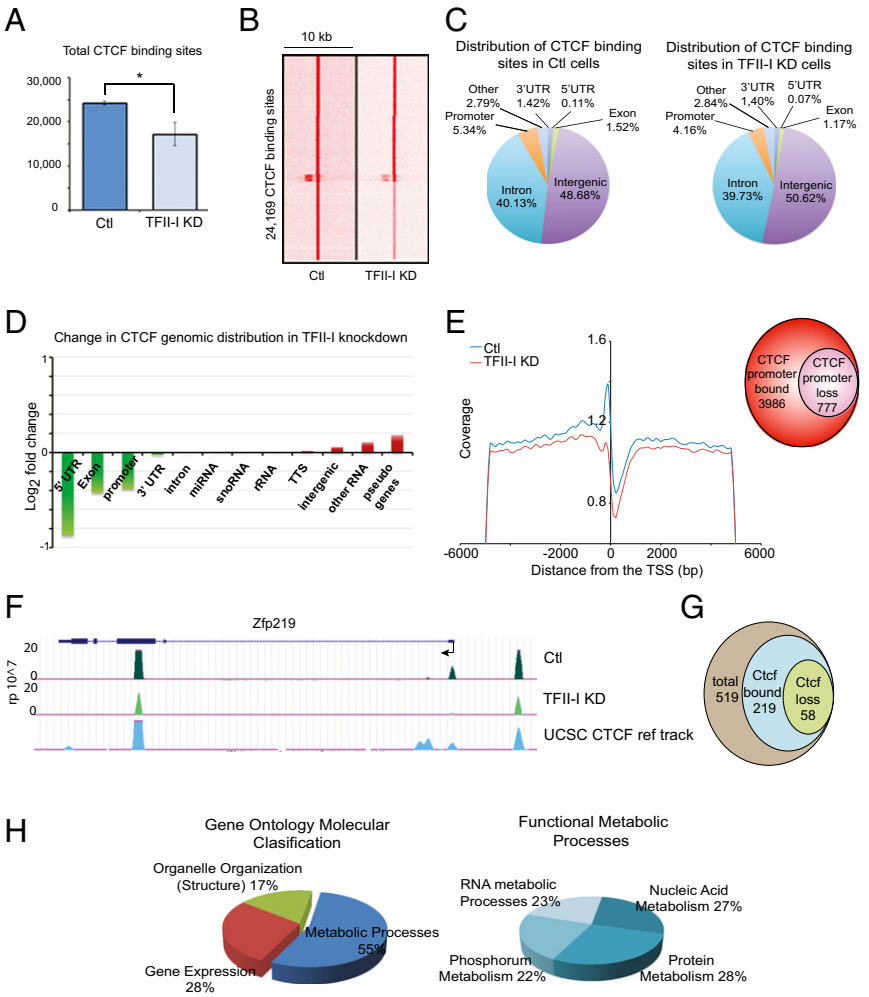


Fig. 4. CTCF binding is disrupted at promoter regions genome wide upon TFII-I knockdown. (*A*) Histogram of the number of CTCF binding sites identified in ctl and TFII-I KD cells by MACS [$*P = 0.03$; ChIP-seq data represent results from two ($n = 2$) independent experiments]. (*B*) Heat map of CTCF binding sites in control and TFII-I KD cells within a ±5-kb window surrounding the binding sites. (*C*) Pie chart of CTCF peak distribution at various genomic loci. (*D*) Histogram of the $\log_2$ fold change in CTCF peak distribution between ctl and TFII-I KD cells. (*E*) Smooth-lined scatter of CTCF tag density in ±5-kb window around the transcription start site (TSS) and its coverage represented by number of tags per base pair. (*F*) Visual representation of CTCF binding in gene *Znf219* in control and knockdown cells using University of Southern California genome browser. UCSC CTCF reference track was used for comparison. (*G*) Of the 519 genes revealed by microarray analysis to undergo significant change upon TFII-I KD, 219 harbor CTCF binding sites; of these, 59 show a loss of CTCF binding upon TFII-I KD. (*H*) GO analysis of the 777 genes where CTCF binding is lost in promoter regions. Approximately 55% of these genes are related to metabolic processes such as RNA, nucleic acid, phosphorum, and protein metabolism.

Recently, a genome-wide screen of TFII-I binding sites was carried out using the K562 cell line as a model (23). We aligned these sites with CTCF binding sites from the same cell line using publicly available Encode data. We find 20% of TFII-I sites localized near transcription start sites are cooccupied by CTCF (Fig. S5), whereas CTCF can be found at less than 10% of TFII-I sites bound outside promoter proximal regions. Again, this overlap underscores the potential importance of cooperativity between CTCF and TFII-I.

Next, we compared overlap between genes regulated by TFII-I from our microarray data with gene promoters where CTCF binding was occluded after TFII-I KD. This analysis revealed that of the 519 genes significantly changed upon TFII-I KD (fold change ≥1.6, $P < 0.05$), ~219 genes have a CTCF binding site, and of these, 58 genes have a loss of CTCF binding in TFII-I KD cells (Fig. 4*G*). Overall, our data predict that TFII-I directs CTCF binding primarily to promoter regions, concordant with its own role as a general transcription factor. CTCF has previously been shown to facilitate the binding of UBTF to ribosomal DNA (30), but to our knowledge TFII-I is the first mammalian factor shown to enhance the association of CTCF to chromatin. In *Drosophila*, the introduction of deletion mutations to the insulator protein Cp190 virtually abolished CTCF binding to polytene chromosomes (21). Thus, there is a precedent for zinc finger proteins directing CTCF to target sites. Recently it was demonstrated that sequences flanking core CTCF binding elements influence CTCF binding in vivo (22). Cooperative loading of CTCF and protein binding partners onto chromatin may explain this phenomenon.

Our GO analysis of the 777 genes where CTCF binding was displaced after TFII-I KD showed enrichment for genes involved in metabolic processes (Fig. 4*H*); this is highly consistent with our GO and KEGG pathway analysis of genes regulated by TFII-I (Fig. 2 *C* and *D*). As a model, we showed that CTCF binding to the *Cdkn2a*^{Arf} locus is lost after TFII-I KD (Fig. 3), and our ChIP-seq data reveals that this model is extended to other CTCF targets, including the metabolic genes *Khk* and *PdhB* (Fig. 5*A*). Mechanistically, TFII-I controlled metabolic genes in a similar fashion, as we saw at the *Cdkn2a*^{Arf} locus. At the *Khk* gene, where TFII-I targets

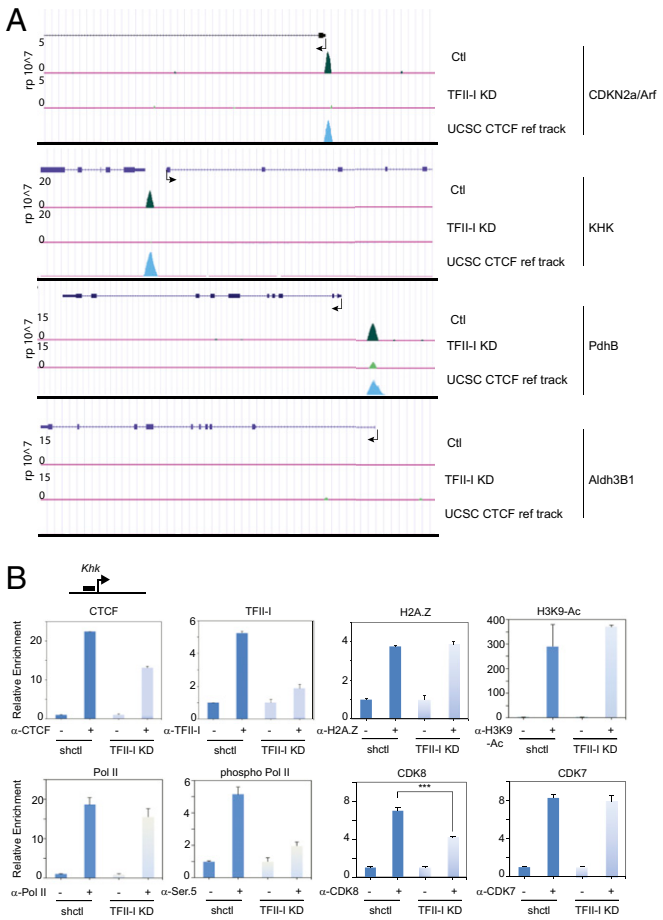


Fig. 5. TFII-I targets CTCF and CDK8 to target genes. (*A*) Visual representation of CTCF binding in select genes in ctl and TFII-I KD using the UCSC genome browser. A UCSC CTCF reference track was added for comparison. Aldh3B1 was use as a negative control because it does not show a CTCF binding site. (*B*) ChIP analysis of the promoter region of *Khk*. A loss of CTCF and TFII-I binding upon TFII-I knockdown, as well loss of Phospho-Ser5-RNA Pol II and CDK8 ($n = 3$) are observed. Relative enrichment was calculated fold Ab over no Ab. For all graphs, error bars represent SEM. ***$P \leq 0.05$ (two-tailed Student *t* test).

CTCF to the proximal promoter (Fig. 5), we see that TFII-I KD has little impact on total RNA Pol II recruitment or epigenetic modification. However, similar to our data at the *Cdkn2a*^{Arf} gene, both CDK8 recruitment and serine 5 phosphorylation of RNA Pol II are compromised in TFII-I knockdown cells (Fig. 5*B*). This data also suggests that the position of CTCF binding relative to the transcription start site may greatly influence its role in the transcriptional process. Our data indicates proximal promoter bound CTCF acts as an activator of initiation, whereas other reports have demonstrated CTCF bound within exonic regions acts as a negative regulator of transcription through enhanced pausing (9, 45).

Because of the link between TFII-I and genes involved in metabolism, we wanted to test TFII-I KD cells for an altered response to nutrient deprivation. Glucose represents one of the principle sources of cellular energy, and we see the altered expression of several genes that might alter the utilization of glucose as a fuel source after TFII-I KD, including *PdhB*, a key mediator of glycolytic flux, and other genes that may impact flux within the tricarboxylic acid cycle (TCA) cycle, such as *Me2*. This finding prompted us to compare the growth of control and TFII-I KD cells under conditions of low glucose. TFII-I KD cells display a survival advantage over control cells, when grown in low glucose (Fig. 6*A*). TFII-I KD cells underwent less cell death in this nutrient-depleted environment as evidenced by a strikingly lower sub-G1 population (Fig. 6 *B* and *C*). Cell death was further confirmed through the use of Annexin V staining. Here, we observed that 2 d in low glucose led to more control cells positive for Annexin V staining than TFII-I knockdown cells (Fig. 6 *D* and *E*), again indicating TFII-I knockdowns are more adaptive to metabolic stress.

Another key energy source is the amino acid glutamine. There are multiple pathways through which glutamine can be metabolized for energy production. One key pathway is the conversion of glutamine to α-ketoglutarate by Glud1, which provides α-ketoglutarate for subsequent utilization by the TCA cycle. Because we see lower levels of the *Glud1* transcript in TFII-I KD cells we examined the response to glutamine deprivation. Again, we see TFII-I KD cells are adaptive to these conditions and show much less sub-G1 content (Fig. S6*A*). Even under conditions of low glucose and glutamine deprivation, TFII-I KD cells are resistant to cell death (Fig. S6*B*). Recently, it has been shown that cancer-specific mutations of TFII-I increase cell proliferation (46). Though it remains to be proven, it is also possible that loss of function *GTF2i* mutations may disrupt metabolic homeostasis

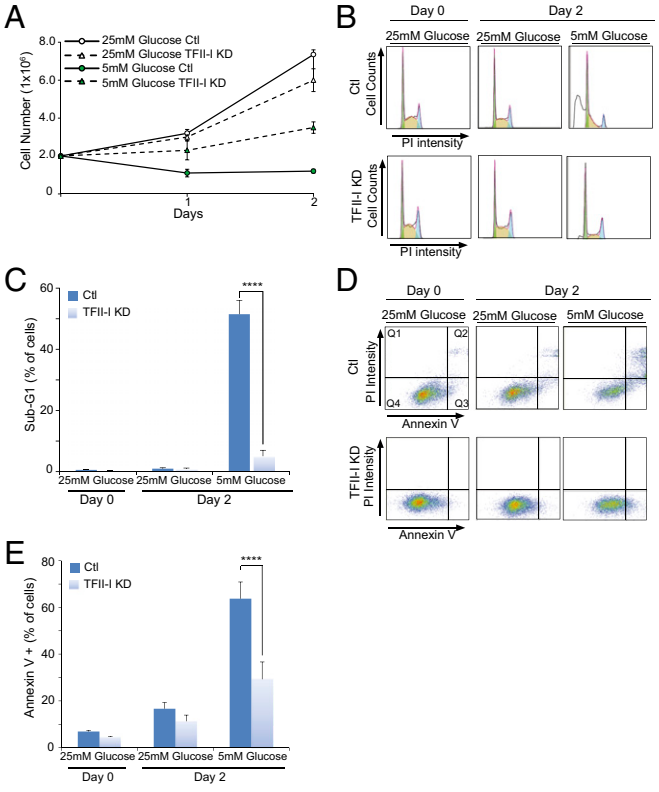


Fig. 6. TFII-I knockdown cells show a survival advantage in response to nutrient deprivation. (*A*) Ctl and TFII-I KD cells were cultured in normal glucose (25 mM) and low glucose (5 mM) conditions, both supplemented with 5% FBS for 2 d. TFII-I KD show a significant survival advantage under low glucose conditions compared with ctl cells. (*B*) Representative cell cycle analysis by propidium iodide staining after 2 d of culture in normal and low glucose. Ctl cells display a higher sub-G1 population than TFII-I KD in low glucose. (*C*) Histogram shows the frequency of cells in sub-G1 population culture under low glucose conditions. (*D*) Representative Annexin V and propidium iodide staining density plots after 48 h of low glucose treatment. Ctl cells have more than 30% more Annexin V positive cells compared with TFII-I KD cells. (*E*) Quantification of Annexin V positive ctl and TFII-I KD cells under normal or low glucose conditions. The resulting histogram shows a significant difference between ctl and TFII-I KD cells. For all graphs, error bars indicate SEM (****$P < 0.01$, two-tailed Student *t* test).

to provide a survival advantage that is distinct from its previously described role in proliferation (47, 48).

Cancer cells undergo substantial metabolic reprogramming to enable high rates of proliferation and survival in a less-than-ideal environment. Among the many changes already characterized is decreased expression of *Khk* in renal carcinomas (49), a gene which we show here is dependent on TFII-I for proper expression. In the future, we will explore whether TFII-I or CTCF binding to such genes is compromised leading to diminished expression in cancers.

Herein, we show that TFII-I is a key regulator of metabolic gene expression, and we propose that this effect is mediated, at least in part, by promoting CTCF binding at the promoter region of a subset of these genes. Our data indicate TFII-I and CTCF cooperate to promote CDK8 recruitment and Pol II phosphorylation on serine 5. Intriguingly, CDK8 recruitment to early response genes has been documented in response to various extracellular signals, including hypoxia, serum, and hormones (50–52). It will be important for future studies to probe the possibility that extracellular signals known to modify TFII-I activity impact the transcription of metabolic regulatory genes through the TFII-I/CTCF/CDK8 axis described herein.

Materials and Methods

Cell Culture and Growth Curves. WEHI-231 control and WEHI-231 TFII-I knockdown cell lines were a generous gift of A.L.R., and were cultured as previously described (32). MDA-MB-435 and HCT-116 cells were cultured in RPMI and DMEM, respectively, supplemented with 10% (vol/vol) FBS. For proliferation assays, 2×10^5 WEHI-231 control and KD cells were cultured in 2.5 mL of DMEM with 25 mM (high) or 5 mM (low) glucose ± glutamine and 5% (vol/vol) FBS. Cells were counted at 24 and 48 h, and cell viability was evaluated by trypan blue exclusion assay.

Generation of CTCF Knockdowns. shRNA against mouse CTCF was purchased from Sigma (SHCLNG-NM_181322). Briefly, virus was packaged in Hek293T cells. The 5.5×10^6 cultured cells were transfected with 5 μg of packaging vector Pax2, 2 μg of envelope vector MD2G, and 7 μg of shRNA against CTCF using 42 μg of polyethylenimine (PEI; 1 mg/mL). The mix of PEI and plasmid DNA was prepared in DMEM without FBS and incubated at room temperature (RT) for 15 min. After this time, the transfection mix was added to the cells and viruses were collected 72 h after transfection. For infection of CTCF shRNA viral supernatants into WEHI-231 cells, 0.3×10^6 cells/mL were seeded in six-well dishes and 500 μL of virus were added along with 8 μg of hexadimethrine bromide (Polybrene; 4 mg/mL). At 24 h post initial infection, a second round of infection was carried out. After 72 h of initial infection, cells were selected with 0.5 μg of puromycin for 48 h and then collected.

Antibodies. Primary antibodies used for co-IP and ChIP were as follows: rabbit polyclonal anti-CTCF catalog no. 07-729 (Millipore); mouse anti-TFII-I catalog no. 610942 (BD Biosciences); rabbit polyclonal anti-H2A.Z catalog no. 07-594 (Millipore); rabbit polyclonal anti-RNA Pol II catalog no. sc-9001 (Santa Cruz); mouse monoclonal anti-RNA Pol II phosphor-Ser5 H14 catalog no. MMS-143R (Covance); rabbit polyclonal anti-H3K27ac catalog no. 07-360 and H3K9ac catalog no. 07-352, and rabbit polyclonal anti-H3K27me3 catalog no. 07-449 (Millipore); goat polyclonal anti-CDK8 catalog no. sc-1521 (Santa Cruz Biotechnology); anti-Flag M2 Affinity gel catalog no. A2220 (Sigma). For Western blot, mouse monoclonal anti-Flag catalog no.1804 (Sigma); rabbit polyclonal anti-DDX5 catalog no. ab21696 (Abcam); mouse anti-TFII-I catalog no. 610942 (BD Biosciences); mouse monoclonal anti-CTCF catalog no. 612149 (BD Biosciences); mouse monoclonal anti-Actin (H-6) catalog no. sc-376421; rabbit polyclonal anti-Sp1 (H-225) catalog no. sc-14027, mouse monoclonal nucleolin (c23) catalog no. sc-8031, and mouse monoclonal anti-B23 Nucleophosmin catalog no. sc-47725 (Santa Cruz Biotechnology). Rabbit polyclonal anti-parp1 catalog no. 9542S was from Cell Signaling Technology and rabbit polyclonal anti-Parp 2 catalog no. ab93841 was from Abcam.

Cell Cycle and Flow Cytometry Analysis. Briefly, cells cultured under high- and low-glucose conditions were washed in 1× PBS with 3% FBS and spun down for 5 min at $635 \times g$. Cells were fixed in ethanol at 75% in PBS and stored at 4 °C. Before propidium iodide staining, cells were centrifuged to remove fixation media and washed in 1× PBS with 3% FBS followed by overnight staining. Before FACS analysis, cells were resuspended in 300 μL of fresh propidium iodide and analyzed in BD FACSCalibur. For apoptosis analysis, Annexin V-CF633 (Biotium) staining protocol was followed per the manufacturer's instructions, and analysis was carried out using FlowJo software.

Mass Spectrometry and Protein Sequencing. MDA-MB-435 whole-cell extracts spiked with 1 μg recombinant CTCF were subject to immunoprecipitation (IP) with rabbit polyclonal anti-CTCF. Immunoprecipitated proteins were resolved by SDS/PAGE and visualized with Coomassie blue staining. Coomassie blue-stained protein bands were excised and subjected to disulfide reduction and alkylation. Trypsin-digested samples were injected onto nano-HPLC system with a C18 capillary column. Peptides were eluted with a linear gradient from 5% to 45% acetonitrile and effluent was electrosprayed into the LTQ mass spectrometer. The mass spectrometer was operated in data-dependent mode to automatically switch between survey MS and MS/MS acquisitions. The conventional MS spectra (survey scan) were acquired in the Orbitrap at a resolution of 60,000 followed by three MS/MS scanning, peptide charged over 10,000 were kept for sequencing. MS/MS spectra were searched against an IPI human database (IPI version 3.24) using Mascot (version 2.3.2; Matrix Science); the threshold of parental and ionic fragment were set at ±0.4 and 0.1 Da, respectively. Manual inspection of all MS/MS spectra for modified peptides was performed to validate assignments. emPAI scores are dependent on the number of observed peptides divided by the number of observable peptides (53) and calculated using Matrix Science software ($\text{emPAI} = 10^{\text{PAI}} - 1$, where PAI is the number of observed peptides divided by the number of observable peptides). Mascot scores were calculated using Matrix Science software.

Column Chromatography. All procedures were performed at 4 °C, and all chromatography matrices were purchased from GE Healthcare Bio-Sciences. Flow rate was 0.5 mL/min except for DNA Cellulose, which was set at 0.1 mL/min. HeLa whole-cell extracts were sonicated and diluted 10-fold in binding buffer [25 mM Hepes (pH 7.6), 50 mM KCl, 10% glycerol, 12.5 mM MgCl2, and 1 mM DTT]. Phosphocellulose p11 column (cation exchange) was prepared according to the manufacturer's instructions and equilibrated with 2× column volume with binding buffer. The diluted cell extract was applied to column, and washes were performed with 2× column volumes of binding buffer. The retained proteins were eluted with increasing concentrations of KCL. The eluted fractions were combined and salt concentration and pH were adjusted before loading to Q-Sepharose (anion exchange). After washes, a gradient of linearly increasing salt concentration was applied to elute the sample components from the column. The eluted protein fractions were pooled and loaded on DNA cellulose column (native double-stranded calf thymus DNA cellulose). DNA cellulose column was prewashed before use in 20 mM Tris·HCl (pH 8.0), 50 mM KCl, 10% glycerol, 1 mM MgCl2, 1 μM ZnCl2, and 1 mM DTT. After washing the column extensively, the bound proteins were eluted with the above buffer containing increasing concentrations of KCL.

Confocal Microscopy. MDA-MB-435 cell monolayers grown on glass coverslips in six-well plates were washed twice with 1× PBS and fixed at RT in 1% PFA containing 0.25% Triton X-100 for 15 min. The cells were then washed twice with PBS and blocked in blocking buffer (1× PBS, 1% BSA, 2% normal goat serum) for 60 min. Antibodies targeting CTCF (BD Biosciences catalog no. 612149; 1:50), TFII-I (Cell Signaling Technologies; catalog no. 4562; 1:200) were then added directly to the blocking buffer and incubated overnight. The cells were then washed three times in 1× PBS and incubated in blocking buffer containing respective Alexa-conjugated secondary antibodies (Invitrogen; 1:1,000) and incubated for 1 h at RT. The cells were then washed three times in 1× PBS and mounted on slides with VectaShield containing DAPI. Images were captured using a 60× oil-immersion objective on the Wave FX spinning disk confocal microscopy system (Quorum Technologies) and analyzed using Volocity. Intensity scatter plot and corresponding Pearson correlation coefficient for CTCF and TFII-I double-stained cells were generated using Volocity.

Western Blot. Cell were lysed with whole-cell lysis buffer [20 mM Tris (pH 7.5), 420 mM NaCl, 2 mM $MgCl_2$, 1 mM EDTA, 10% glycerol, 0.5% Nonidet P-40, 0.5% Triton] supplemented with fresh 1 mM DTT and PMSF and protease and phosphatase inhibitor mixture, BGP, P8340, and NAF. Whole-cell lysis buffer was added 2× the volume of the cell pellet and left on ice for 25 min. Lysates were centrifuge at top speed at 4 °C and supernatant was collected and transferred to a new tube. Protein concentration was measured by Bradford assay. A total of 35 μg of protein were loaded into 8% acrylamide gels and transfer was done overnight at 30 V at 4 °C. Membranes were washed three times with Tris-buffered saline with Tween (TBST; Tris base 20 mM, NaCl 137 mM, and 0.1% Tween 20) for 5, 10, and 15 min. Membrane was blocked with 5% milk in TBST and incubated with primary antibody

overnight at 4 °C. After primary antibody incubation, membrane was washed as previously described and incubated with secondary antibody for 1 h at room temperature. Western blot was revealed using a Clarity Western Enhanced Chemiluminescence Kit (BioRad).

Gene Expression Analysis. Gene Elute Mammalian Total RNA kit (Sigma) was used for RNA extraction following manufacturer instructions. The 1 μg at 50 ng/μL of total WEHI-231 cell RNA were sent to Innovation Génome Québec for microarray analysis. Illumina Beads technology was used (MouseWG-6_V2). Microarray analysis was performed using Bioconductor package limma (54). Validation of target genes identified was realized as followed: total RNA were extracted with Gene Elute Mammalian Total RNA kit following manufacturer instructions. cDNA was generated by reverse transcription using 1 μg of RNA. Easy Script kit from ABM was used following manufacturer's protocol. qPCR reactions were carried out using SYBR Green dye from Promega with specific primers at 50 μM concentration. Nascent mRNA was examined for CDKN2B/Arf after on-column DNase digestion and using primers spanning the first exon and intron. qPCR primer sequences are listed in Table S1. Microarray data sets are available from the Gene Expression Omnibus (GEO) database under accession no. GSE60915 (www.ncbi.nlm.nih.gov/geo).

Co-IP. Cells were collected, washed with 1× PBS, and lysed as described above. Protein concentration was quantified by Bradford assay, and 1 mg of protein was used for IP. Protein lysates were diluted five times in IP buffer [20 mM Tris (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 2 mM EDTA, 0.5% Triton X-100], and phenylmethylsulfonyl fluoride (PMSF) was added to avoid protein degradation. Protocol for samples treated with nucleases (DNase and RNase) was carried out as previously described (55). Lysates were precleared for 2 h with Protein G agarose beads. After the preclearing stage, beads were pelleted and supernatant was collected and transferred to a new tube where antibodies were added and nutated overnight at 4 °C. Antibodies were collected by adding fresh Protein G agarose beads and nutated at 4° for 2 h. Beads were pelleted and the supernatant was discarded. The collected beads were washed three times by adding 1 mL with IP buffer (100 mM NaCl) and centrifuged at 1,700 × *g* to pellet the beads. The final wash was done with IP buffer containing 0.1% Triton X-100 and spun down at 2,700 × *g*. Proteins were eluted from the beads by adding 25 μL of 2× SDS loading buffer [3M Tris (pH 6.8), 20% SDS, 100% glycerol, H_2O, and bromophenol blue] and heated at 100 °C for 10 min. Beads were again pelleted and the resulting supernatant was loaded to acrylamide gel and blotted as described above.

ChIP. WEHI-231 ctl and TFII-I KD cells were collected and cross-linked with 1% formaldehyde in PBS for 10 min at room temperature. The cross-linking reaction was stopped with 125 mM glycine, and the cells washed with 1× PBS and stored at −80 °C until the assay was carried out. Cells were lysed and DNA sheered by sonication in cell lysis/ChIP buffer [0.25% Nonidet P-40, 0.25% Triton X-100, 0.25% sodium deoxycholate, 0.1% SDS, 50 mM Tris (pH 8.0), 50 mM NaCl, 5 mM EDTA] 15 times for 15 s each. Lysates were centrifuged for 15 min at 20,000 × *g* at 4 °C, and supernatant was collected. A total of 1 mg of protein was precleared for 2 h with Protein G agarose beads (50% slurry blocked with salmon sperm DNA) at 4 °C. IP was carried out by adding 2 μg of antibody and 30 μL of agarose G beads, nutating overnight at 4 °C. After IP, beads were pelleted by centrifugation, followed by four washes to remove unspecific binding using a variety of buffers with varying concentrations of salt. Buffers 1–3 contained 0.1% SDS, 1% Triton-X, 2 mM EDTA, 20 mM Tris (pH 8.0), and 150 mM NaCl, 300 mM NaCl, 500 mM NaCl, respectively. Buffer 4 contained 0.25 M LiCl, 1% Nonidet P-40, 1% sodium deoxycholate, 1 mM EDTA, and 10 mM Tris (pH 8.0). Two additional washes with TE were done to remove any residual buffers from the beads. Complexes bound to the beads were eluted with 500 μL of elution buffer [1% SDS, 1 mM EDTA, 50 mM Tris (pH 8.0)] at 65 °C for 25 min. Beads were pelleted by centrifugation and supernatant was collected. Reverse cross-linking was done by adding 0.2 mM NaCl at 65 °C overnight followed by treatment with Proteinase K at 45 °C for 1 h, and a second incubation of 15 min at 65 °C. DNA recovery was carried out with separation using 500 μL of phenol–chloroform. The aqueous layer was recovered and a second recovery was done with chloroform alone to ensure the complete removal of phenol. DNA precipitation was done overnight by adding 2 μL of yeast transfer RNA (ytRNA) as carrier, 17 μL of sodium acetate, and 900 μL of 95% ethanol. DNA was pelleted by centrifugation at top speed for 15 min at 4 °C, and pellets were washed with 70% ethanol and dried by vacuum centrifugation. DNA was resuspended in 100 μL of H_2O and stored at −20 °C. qPCR primers used are listed in Table S1.

ChIP-seq and Bioinformatics. The ChIP protocol described above was used for ChIP-seq with the exception of the DNA recovery step; in this case, DNA was retrieved using a PCR purification kit (Qiagen) following manufacturer's instruction. CTCF ChIP was performed in duplicate using lysates from WEHI-231-shctl and WEHI-231-TFII-I KD cells. Recovered DNA was sent to Institut de Recherche en Immunologie et Cancérologie sequencing facility where library construction and sequencing (100 bases, paired end, HiSeq; Illumina) were performed. DNA sequences obtained were trimmed to 45 bases (quality score >30) and were aligned to the mouse genome (National Center for Biotechnology Information Build 37, July 2007, mm9) using the BWA algorithm (56). After alignment, duplicates were removed and only the sequences with MAPQ score ≥30 were kept for further analysis. The model-based analysis of ChIP-SEq (57) peak-finding algorithm (MACS) was used to identify peaks in ctl and TFII-I KD conditions using the default settings. The loss of CTCF binding sites in TFII-I KD cells was quantified using MACS with TFII-I KD data set as background; during this process, MACS software normalized the total tags count between the two samples. The intersect function of BEDTools with an overlap window of 50 bp was used to identify region of colocalization between TFII-I and CTCF in K562 cells in Fig. S4 (GEO accession no. GSE51065). HOMER was used to annotate CTCF peaks and determine their genomic distribution (58). seqMiner software was used to generate all cluster and heat map data (59). Data sets are available on the GEO database (accession no. GSE60917).

Gene Ontology and KEGG Analysis. R (version 3.1.0) and Bioconductor (version 2.14) package biomaRt (60), GOstat (61), and KEGGprofile (62) were used to carry out all GO and pathway analysis on expression and ChIP-seq data.

1. Marshall AD, Bailey CG, Rasko JEJ (2014) CTCF and BORIS in genome regulation and cancer. *Curr Opin Genet Dev* 24(0):8–15.
2. Ong C-T, Corces VG (2014) CTCF: An architectural protein bridging genome topology and function. *Nat Rev Genet* 15(4):234–246.
3. Fiorentino FP, Giordano A (2012) The tumor suppressor role of CTCF. *J Cell Physiol* 227(2):479–492.
4. Cuddapah S, et al. (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* 19(1):24–32.
5. Chernukhin I, et al. (2007) CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. *Mol Cell Biol* 27(5):1631–1648.
6. Soshnikova N, Montavon T, Leleu M, Galjart N, Duboule D (2010) Functional analysis of CTCF during mammalian limb development. *Dev Cell* 19(6):819–830.
7. Vostrov AA, Taheny MJ, Quitschke WW (2002) A region to the N-terminal side of the CTCF zinc finger domain is essential for activating transcription from the amyloid precursor protein promoter. *J Biol Chem* 277(2):1619–1627.
8. Saldaña-Meyer R, et al. (2014) CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev* 28(7):723–734.
9. Shukla S, et al. (2011) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479(7371):74–79.
10. Dixon JR, et al. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485(7398):376–380.
11. Handoko L, et al. (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat Genet* 43(7):630–638.
12. Weth O, Renkawitz R (2011) CTCF function is modulated by neighboring DNA binding factors. *Biochem Cell Biol* 89(5):459–468.
13. Wallace JA, Felsenfeld G (2007) We gather together: Insulators and genome organization. *Curr Opin Genet Dev* 17(5):400–407.
14. Ishihara K, Oshimura M, Nakao M (2006) CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol Cell* 23(5):733–742.
15. Xiao T, Wallace J, Felsenfeld G (2011) Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol Cell Biol* 31(11):2174–2183.
16. Yao H, et al. (2010) Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev* 24(22):2543–2555.
17. Donohoe ME, Zhang L-F, Xu N, Shi Y, Lee JT (2007) Identification of a CTCF cofactor, Yy1, for the X chromosome binary switch. *Mol Cell* 25(1):43–56.
18. Chernukhin IV, et al. (2000) Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. *J Biol Chem* 275(38):29915–29921.
19. Merkenschlager M, Odom DT (2013) CTCF and cohesin: Linking gene regulatory elements with their targets. *Cell* 152(6):1285–1297.
20. Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G (2004) CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* 13(2):291–298.
21. Mohan M, et al. (2007) The *Drosophila* insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J* 26(19):4203–4214.

22. Nakahashi H, et al. (2013) A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Reports* 3(5):1678–1689.
23. Fan AX, et al. (2014) Genomic and proteomic analysis of transcription factor TFII-I reveals insight into the response to cellular stress. *Nucleic Acids Res* 42(12):7625–7641.
24. Roy AL, Meisterernst M, Pognonec P, Roeder RG (1991) Cooperative interaction of an initiator-binding transcription initiation factor and the helix-loop-helix activator USF. *Nature* 354(6350):245–248.
25. Novina CD, Cheriyath V, Roy AL (1998) Regulation of TFII-I activity by phosphorylation. *J Biol Chem* 273(50):33443–33448.
26. Hakre S, et al. (2006) Opposing functions of TFII-I spliced isoforms in growth factor-induced gene expression. *Mol Cell* 24(2):301–308.
27. Witcher M, Emerson BM (2009) Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol Cell* 34(3):271–284.
28. Weiss W, Görg A (2009) *High-Resolution Two-Dimensional Electrophoresis. Proteomics, Methods in Molecular Biology*, eds Reinders J, Sickmann A (Humana, New York), Vol 564, pp 13–32.
29. Dunn KW, Kamocka MM, McDonald JH (2011) A practical guide to evaluating colocalization in biological microscopy. *Am J Physiol Cell Physiol* 300(4):C723–C742.
30. van de Nobelen S, et al. (2010) CTCF regulates the local epigenetic state of ribosomal DNA repeats. *Epigenetics Chromatin* 3(1):19.
31. Kim D-W, Cochran BH (2000) Extracellular signal-regulated kinase binds to TFII-I and regulates its activation of the c-fos promoter. *Mol Cell Biol* 20(4):1140–1148.
32. Ashworth T, Roy AL (2007) Cutting edge: TFII-I controls B cell proliferation via regulating NF-kappaB. *J Immunol* 178(5):2631–2635.
33. Soto-Reyes E, Recillas-Targa F (2010) Epigenetic regulation of the human p53 gene promoter by the CTCF transcription factor in transformed cell lines. *Oncogene* 29(15): 2217–2227.
34. Splinter E, et al. (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev* 20(17):2349–2354.
35. Li X-Y, Virbasius A, Zhu X, Green MR (1999) Enhancement of TBP binding by activators and general transcription factors. *Nature* 399(6736):605–609.
36. Egloff S, Murphy S (2008) Cracking the RNA polymerase II CTD code. *Trends Genet* 24(6):280–288.
37. Napolitano G, Lania L, Majello B (2014) RNA polymerase II CTD modifications: How many tales from a single tail. *J Cell Physiol* 229(5):538–544.
38. Helenius K, et al. (2011) Requirement of TFIIH kinase subunit Mat1 for RNA Pol II C-terminal domain Ser5 phosphorylation, transcription and mRNA turnover. *Nucleic Acids Res* 39(12):5025–5035.
39. Rickert P, Corden JL, Lees E (1999) Cyclin C/CDK8 and cyclin H/CDK7/p36 are biochemically distinct CTD kinases. *Oncogene* 18(4):1093–1102.
40. Ramanathan Y, et al. (2001) Three RNA polymerase II carboxyl-terminal domain kinases display distinct substrate preferences. *J Biol Chem* 276(14):10913–10920.
41. Egloff S, Dienstbier M, Murphy S (2012) Updating the RNA polymerase CTD code: Adding gene-specific layers. *Trends Genet* 28(7):333–341.
42. Galbraith MD, Donner AJ, Espinosa JM (2010) CDK8: A positive regulator of transcription. *Transcription* 1(1):4–12.
43. Liu Z, Scannell DR, Eisen MB, Tjian R (2011) Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* 146(5):720–731.
44. Liu Y, et al. (2004) Two cyclin-dependent kinases promote RNA polymerase II transcription and formation of the scaffold complex. *Mol Cell Biol* 24(4):1721–1735.
45. Gomes NP, Espinosa JM (2010) Gene-specific repression of the p53 target gene PUMA via intragenic CTCF-Cohesin binding. *Genes Dev* 24(10):1022–1034.
46. Petrini I, et al. (2014) A specific missense mutation in GTF2I occurs at high frequency in thymic epithelial tumors. *Nat Genet* 46(8):844–849.
47. Desgranges ZP, et al. (2005) Inhibition of TFII-I-dependent cell cycle regulation by p53. *Mol Cell Biol* 25(24):10940–10952.
48. Ashworth T, Roy AL (2009) Phase specific functions of the transcription factor TFII-I during cell cycle. *Cell Cycle* 8(4):596–605.
49. Hwa JS, et al. (2006) The expression of ketohexokinase is diminished in human clear cell type of renal cell carcinoma. *Proteomics* 6(3):1077–1084.
50. Donner AJ, Ebmeier CC, Taatjes DJ, Espinosa JM (2010) CDK8 is a positive regulator of transcriptional elongation within the serum response network. *Nat Struct Mol Biol* 17(2):194–201.
51. Galbraith MD, et al. (2013) HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. *Cell* 153(6):1327–1339.
52. Belakavadi M, Fondell JD (2010) Cyclin-dependent kinase 8 positively cooperates with Mediator to promote thyroid hormone receptor-dependent transcriptional activation. *Mol Cell Biol* 30(10):2437–2448.
53. Ishihama Y, et al. (2005) Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 4(9):1265–1272.
54. Smyth GK (2005) limma: Linear models for microarray data. *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Statistics for Biology and Health, eds Gentleman R, et al (Springer, New York), pp 397–420.
55. Zhang N, et al. (2008) A handcuff model for the cohesin complex. *J Cell Biol* 183(6): 1019–1031.
56. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
57. Zhang Y, et al. (2008) Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9(9): R137.
58. Heinz S, et al. (2010) Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol Cell* 38(4):576–589.
59. Ye T, et al. (2011) seqMINER: An integrated ChIP-seq data interpretation platform. *Nucleic Acids Res* 39(6):e35.
60. Durinck S, et al. (2005) BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* 21(16):3439–3440.
61. Falcon S, Gentleman R (2007) Using GOstats to test gene lists for GO term association. *Bioinformatics* 23(2):257–258.
62. Shao S, Shyr Y (2012) KEGGprofile: An annotation and visualization package for multi-types and multi-groups expression data in KEGG pathway. R package version 1.6.1. Available at www.bioconductor.org/packages/release/bioc/html/KEGGprofile.html.