nature structural & molecular biology

7SK-BAF axis controls pervasive transcription at enhancers

Ryan A Flynn[1,2], Brian T Do[1,2], Adam J Rubin[2], Eliezer Calo[3], Byron Lee[1,2], Hannes Kuchelmeister[4], Michael Rale[5], Ci Chu[1,2], Eric T Kool[4], Joanna Wysocka[3], Paul A Khavari[2] & Howard Y Chang[1,2]

RNA functions at enhancers remain mysterious. Here we show that the 7SK small nuclear RNA (snRNA) inhibits enhancer transcription by modulating nucleosome position. 7SK occupies enhancers and super enhancers genome wide in mouse and human cells, and it is required to limit enhancer-RNA initiation and synthesis in a manner distinct from promoter pausing. Clustered elements at super enhancers uniquely require 7SK to prevent convergent transcription and DNA-damage signaling. 7SK physically interacts with the BAF chromatin-remodeling complex, recruits BAF to enhancers and inhibits enhancer transcription by modulating chromatin structure. In turn, 7SK occupancy at enhancers coincides with that of Brd4 and is exquisitely sensitive to the bromodomain inhibitor JQ1. Thus, 7SK uses distinct mechanisms to counteract the diverse consequences of pervasive transcription that distinguish super enhancers, enhancers and promoters.

Eukaryotic genomes are extensively transcribed[1,2], but unfettered transcription alters gene expression and leads to genome damage through several means[3]. RNA polymerase (Pol) II transcribes functional regulatory elements, such as enhancers and super enhancers (SEs). SEs, also known as stretch enhancers, are distinguished from enhancers by their disproportionate concentration of bound transcriptional and chromatin-modification machinery, and they appear to be particularly important for regulating gene-expression networks and controlling cell state[4–6]. Nearly all active enhancers generate bidirectional transcripts termed enhancer RNAs (eRNAs)[2], and they are also marked by flanking, phased nucleosomes with specific histone modifications[7–9]. Genome-wide characterization of promoter and enhancer elements has revealed commonalities between promoters and enhancers, such as divergent transcription start site (TSS) pairs, well-phased +1 and −1 nucleosomes flanking the TSSs, and a central transcription factor (TF)-binding site[10]. Notably, whereas enhancer transcription is one of the earliest steps of gene activation[11], and some eRNAs are known to participate in gene regulation[12,13], far less is known about the control of eRNA transcription.

Most previous studies of how polymerases are regulated *in vivo* have focused on protein factors (reviewed in refs. 14,15). However, several long noncoding RNAs have enhancer-like activities that can recruit transcription-activating complexes and positively regulate genes *in cis* or *in trans*[16]. Short Alu-containing RNAs directly bind the preinitiation complex *in trans* and repress Pol II transcription[17]. Another *trans*-acting RNA, the 7SK snRNA, has been shown to uniquely control Pol II pausing at promoters, but its role genome wide has yet to be fully elucidated.

7SK is a highly abundant snRNA (with ~200,000 copies per cell[18]) that serves as a scaffold for Larp7, Hexim1, Mepce and Cdk9–CyclinT1 (also known as positive transcription elongation factor b (P-TEFb)), which together form the canonical 7SK small nuclear ribonucleoprotein (snRNP). A key step in RNA Pol II transcription is promoter-proximal pausing, which occurs bidirectionally ~25–60 nucleotides downstream of TSSs. Promoter-proximal pause release is gated by the P-TEFb–7SK snRNA pathway; release from 7SK allows for phosphorylation of Pol II by P-TEFb (as well as other kinases[19,20]) and subsequent elongation[21]. The 7SK snRNP is thought to reside in the nucleoplasm, but it has been suggested that 7SK may operate physically on chromatin[22–26]. Although previous studies have focused on the specific functions of the protein components of the 7SK snRNP, a comprehensive analysis of the RNA component of the 7SK snRNP at chromatin has not been determined.

RNA is ideally suited to control the flow of information in the nucleus for several reasons. RNA transcription or occupancy can mark unique allele or spatial positions in the nucleus[27]. Many long noncoding RNAs act as RNA scaffolds that bring multiple distinct protein complexes into physical proximity and consequently regulate gene expression or other functions[28]. Further, RNA can base-pair with itself, thus forming complex secondary and tertiary structures, and an individual RNA sequence can adopt multiple physical conformations with similar folding energies[29]. To understand how the 7SK snRNA mechanistically operates across the mammalian genome, we used an RNA-centric approach to identify new functions of 7SK. By combining recently described and new RNA-centric technologies, we characterized 7SK's *in vivo* binding partners, chromatin occupancy sites and RNP-specific conformations, and revealed a new role for 7SK as an important regulator of ATP-dependent chromatin remodeling at enhancers.

RESULTS

Conserved 7SK chromatin occupancy at super enhancers

We mapped the genomic occupancy of 7SK through chromatin isolation by RNA purification and subsequent deep sequencing (ChIRP-seq)

[1]Center for Personal Dynamic Regulomes, Stanford University School of Medicine, Stanford, California, USA. [2]Program in Epithelial Biology, Stanford University School of Medicine, Stanford, California, USA. [3]Department of Chemical and Systems Biology, Stanford University School of Medicine, Stanford, California, USA. [4]Department of Chemistry, Stanford University, Stanford, California, USA. [5]The Genome Institute, Washington University in St. Louis, St. Louis, Missouri, USA. Correspondence should be addressed to H.Y.C. (howchang@stanford.edu).

Received 19 December 2015; accepted 20 January 2016; published online 15 February 2016; doi:10.1038/nsmb.3176

in mouse embryonic stem (ES) cells. Using two orthogonal probe sets ('even' and 'odd') targeting 7SK, we recovered most of the cellular 7SK (**Supplementary Fig. 1a**,**b**). Genomic DNA was preferentially recovered over other abundant nuclear and cytoplasmic RNAs, and it was sensitive to RNase A treatment (**Supplementary Fig. 1b**). Unexpectedly, we observed extensive 7SK occupancy across the entire transcribed loci of mRNA genes, in a pattern that mirrored the binding profile of Pol II (**Fig. 1a** and **Supplementary Fig. 1c**) . Outside genic regions, 7SK also associated with SE and typical enhancer (TE) elements (**Fig. 1a**–**c**). At these elements, 7SK overlapped with regions bound by Pol II–pausing factors (Nelf-a and Supt5h) and by a reader of histone acetylation (Brd4), a Pol II initiation mark (TATA-binding protein (TBP)) as well as with open chromatin sites (as measured by assay of transposase-accessible chromatin using sequencing (ATAC-seq)) (**Fig. 1a**). The typical pattern of occupancy was a focal peak of open chromatin bound by Brd4 and other transcription factors, with 7SK extensively occupying the transcribed regions. SEs are recently reported designations for clustered enhancer elements[4]; whether SEs possess distinct biochemical or regulatory properties has been debated[30]. We found that the 7SK ChIRP-seq signal was strongly enriched at SEs, similarly to findings for other core transcriptional machinery (**Fig. 1b**,**c**). These data identified 7SK as a chromatin bound noncoding RNA occupying regulatory regions actively transcribed by Pol II.

ChIRP followed by quantitative PCR (ChIRP-qPCR) and other controls confirmed the specificity of these findings. We selectively recovered 7SK-bound regions in an RNA-dependent manner (**Supplementary Fig. 1d**). The 7SK ChIRP-seq signal was not simply a result of highly active transcription because we observed no enriched 7SK signal at the ribosomal DNA locus (transcribed by Pol I and Pol III but not Pol II), a result consistent with the exclusion of 7SK from the nucleolus (**Supplementary Fig. 1e**). 7SK is highly conserved at the primary sequence level between mice and humans. 7SK ChIRP-seq in human ES (H1) and HeLa cells each identified over 5,000 sites with similar distributions of peaks over genic and enhancer regions (**Supplementary Fig. 1f**–**i**). These results suggest that 7SK's role on chromatin is conserved across mammalian species and cell types.

We developed 'factor ratio analysis' to compare the co-occupancy of 7SK with transcriptional and chromatin regulators at promoters, SEs and TEs (**Supplementary Fig. 2a**). Because many factors involved in transcriptional initiation occupy both promoters and enhancers[10], regulatory features that may distinguish these classes of elements are actively being sought after. Globally, 7SK associated with actively transcribed chromatin and components of the Pol II transcriptional

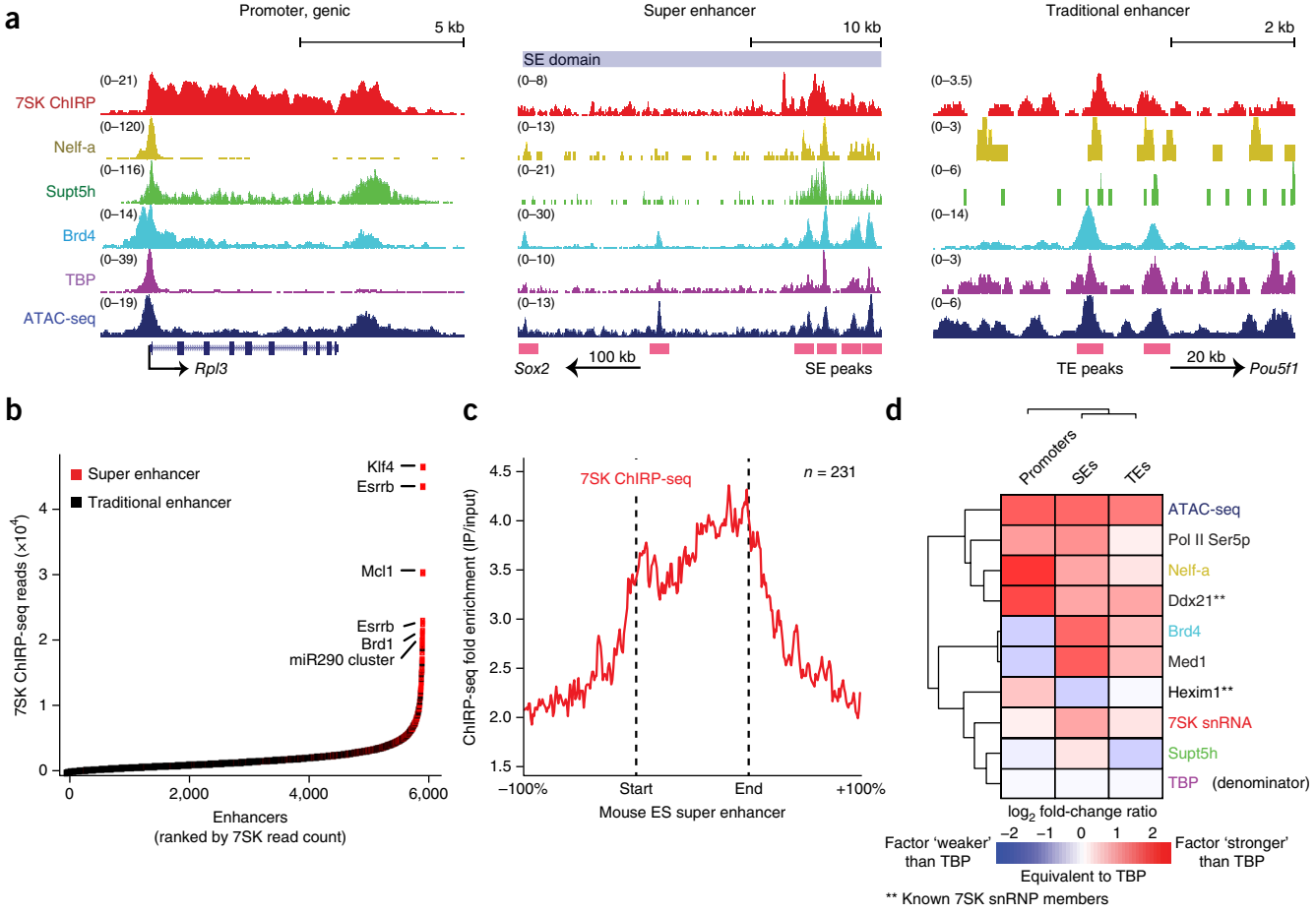

Figure 1 7SK binds transcribed regions and prefers SEs. (**a**) UCSC genome browser views of the *Rpl3* locus (left), *Sox2*-proximal SE (middle) and *Pou5f1*-distal typical enhancer (TE, right). 7SK snRNA (red), Nelf-a (yellow), Supt5h (green), Brd4 (turquoise), TBP (purple) and ATAC-seq of open chromatin (blue) are shown for each locus. Enhancer peaks are marked with pink bars at bottom. (**b**) Hockey plot of 7SK ChIRP-seq across active mouse ES-cell enhancers. Regions defined as SEs[4] are marked in red. (**c**) Metagene analysis of 7SK ChIRP-seq enrichment at 231 previously reported SE elements[4] (*n*, number of elements). The upstream and downstream regions of each SE element are scaled to display the same width of each element. (**d**) Heat map of factor ratio analysis at promoter, SE and TE elements. Colors indicate the $\log_2$ fold-change ratio relative to TBP.

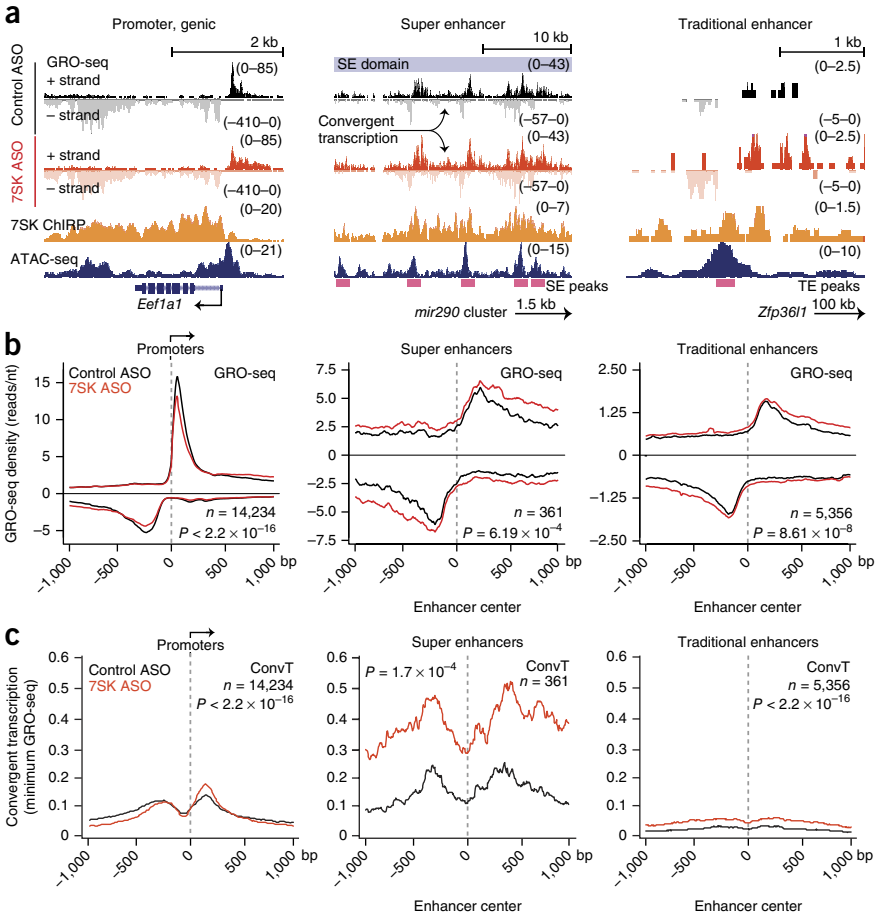

Figure 2 7SK differentially regulates transcription at enhancers and promoters, controlling convergent transcription and DNA-damage signaling. (**a**) UCSC genome browser views of the *Eef1a1* locus (left), the *mir290* family SEs (middle) and the *Zfp36l1* distal enhancer (right). Sense (black, control ASO; red, 7SK ASO) and antisense (gray, control ASO; light red, 7SK ASO) GRO-seq reads, 7SK ChIRP-seq (orange) and ATAC-seq of open chromatin (blue) are shown with enhancer peaks marked as in **Figure 1**. (**b**) Metagene analysis of GRO-seq signals at promoters (left), SEs (middle) and TEs (right), centered at each element ±1 kb. Control and 7SK ASO conditions are shown in black and red, respectively. *P* values are calculated with Hotelling's *t* test (*n*, number of elements). (**c**) Metagene analysis of convergent transcription (ConvT) as measured by GRO-seq, plotted as in **b** and scaled to the same ConvT value (*y* axis).

apparatus (**Supplementary Fig. 2b**). In the factor ratio analysis, the chromatin occupancy signal of each factor was normalized to TBP, thus allowing the proportion of different factors relative to Pol II initiation per base to be determined (Online Methods). We used start-site RNA sequencing (Start-seq)[31] data to precisely align experimentally measured TSSs at promoters, SEs and TEs for all analyses (**Supplementary Table 1** and Online Methods) Further, to directly compare SEs and TEs, we used ATAC-seq to identify individual peaks of open chromatin in SE domains (**Supplementary Table 1**), which can facilitate the analysis of functional regions within SEs[32].

Factor ratio analysis revealed that promoters, SEs and TEs each had distinct and characteristic ratios of factor occupancy, and components of the 7SK snRNP were key distinguishing features (**Fig. 1d** and **Supplementary Fig. 2**). Promoters were enriched for the pausing factor Nelf-a and Pol II Ser5p (initiated but paused Pol II). Additionally, canonical 7SK snRNP proteins such as Hexim1 and Ddx21 (ref. 33) were also biased for promoters. In contrast, SEs and TEs both had higher levels of mediator Med1 and histone acetylation reader Brd4, and these levels were greater for SEs than TEs. Finally, SEs had disproportionately more 7SK than promoters or TEs, and the 7SK at SEs was not associated with proportional levels of pausing (Nelf-a or Supt5h) or canonical 7SK snRNP (Hexim1 or Ddx21) factors as promoters (**Fig. 1d**). Factor ratio analysis using two independent TBP data sets or the basal transcription factor TAF1 as the denominator yielded the same findings, thus confirming the robustness of this approach (**Supplementary Fig. 2c**). The different ratios between 7SK and these chromatin factors at promoters and enhancers suggest that 7SK may mediate additional new functions at enhancer elements.

An orthogonal analysis interrogating binding patterns revealed that 7SK and Nelf-a associate with divergently transcribed polymerase complexes at promoters, tracking Pol II and TBP[10] (**Supplementary Fig. 2d**). We observed a similar trend at enhancer regions. However, at SEs the 7SK binding was broad, more similar to that of Pol II Ser5p (**Supplementary Fig. 2e** and **Fig. 1c**). To better understand how the colocalization of 7SK at promoters and enhancers relates to function, we extended the factor ratio analysis to 41 genome-wide data sets and performed principal component analysis. We found that the first two principal components of the factor ratio analysis differentiated promoters, SEs and TEs (**Supplementary Fig. 2f**). Interestingly, we found that SEs exhibited a pattern intermediate between those of promoters and TEs, but TE and SEs generally clustered together (**Supplementary Fig. 2f**). These data revealed that individual SE peaks are a special class of elements, which have features of both promoters and enhancers, and that SEs attract the most 7SK per Pol II initiation among all three elements. Additionally, whereas 7SK marked all three regions, pausing factor ratios suggested that Pol II pausing operates differently at enhancer elements.

7SK, ConvT, and DNA-damage signaling at SEs

The 7SK association with different factors at specific elements genome wide raised the possibility that 7SK has multiple functions. We therefore directly assayed the effects of 7SK on Pol II transcription by applying global run-on sequencing (GRO-seq) in mouse ES cells. We used antisense oligonucleotides (ASOs)[34] that robustly depleted 7SK (**Supplementary Fig. 3a**). Loss of 7SK globally altered transcription but with several distinct patterns. At promoters, loss of 7SK resulted in a reduction in promoter-proximal transcription and a gain in gene-body transcription, consistently with Pol II pause release (**Fig. 2a**,**b** and **Supplementary Fig. 3b**). In contrast, SEs and TEs exhibited consistent increases in both TSS-proximal and TSS-distal eRNA transcription (**Fig. 2a**,**b**). Quantification of individual elements showed over 1,600 enhancers with significantly increased transcription: 400 enhancers increased eRNA synthesis by 200% or more (131 enhancers at ≥300% and 44 enhancers at ≥400%). In contrast, fewer enhancers (243) showed significantly decreased eRNA production (**Supplementary Fig. 3c**). 7SK depletion also caused increased formation of trimethylated histone H3 K36 at enhancers,

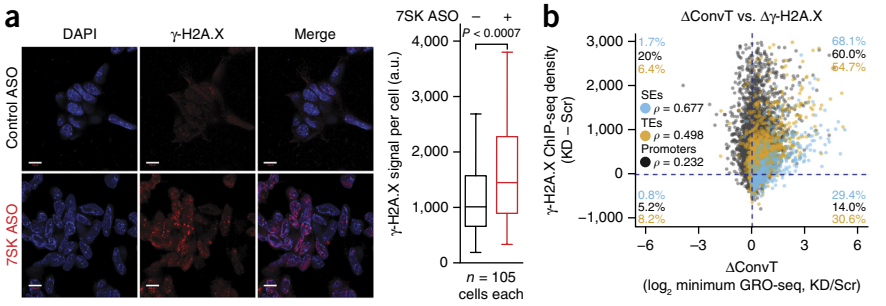

Figure 3 7SK prevents DNA damage at enhancers and SEs. (**a**) Left, representative immunofluorescence images (from a total of ten fields) of paraformaldehyde-fixed mouse ES cells. Scale bars, 10 μm. DAPI, 4′,6-diamidino-2-phenylindole. Right, quantification of 105 individual cells from control ASO– or 7SK ASO–treated mouse ES cells. Boxes, 25th and 75th percentile; center line, median values; whiskers, top and bottom quartiles. *P* values were calculated with the Kolmogorov-Smirnov test. a.u., arbitrary units. (**b**) Scatter-plot analysis of the change in 7SK ASO (KD)/control ASO (Scr) ratio in γ-H2AX ChIP-seq versus $\log_2$ change in ConvT at promoters, SEs and TEs. Spearman correlation values are noted for each element class.

a mark associated with active Pol II transcription (**Supplementary Fig. 3d**). Importantly, enhancers that gained GRO-seq and trimethylated histone H3 K36 signals after 7SK knockdown were enriched for similar ES-specific gene ontology terms above the background of previously defined ES enhancer elements (**Supplementary Fig. 3e**,**f** and Online Methods). Together, these data suggested that 7SK represses transcription at enhancers and promoters through distinct mechanisms.

One important consequence of unfettered transcription is convergent transcription, in which one region of the genome is transcribed on both the Watson and Crick strands (**Supplementary Fig. 4a**). Convergent transcription (ConvT) can lead to Pol II collisions and activation-induced cytidine deamination–mediated DNA damage[3]; thus, surveillance mechanisms that reduce this phenomenon are important for genome stability. Using an established metric of ConvT based on GRO-seq data (Online Methods), we found that transcribed regions experienced more ConvT than did insulator (Ctcf) or random regions, and SEs in particular were the most enriched (**Supplementary Fig. 4b**); these results mirror those from recent reports in B-cell lymphoma[3]. Loss of 7SK caused significant increases in ConvT at promoters, SEs and TEs (**Fig. 2c** and **Supplementary Fig. 4c**). Despite P-TEFb's previously reported activity both downstream and upstream of Pol II complexes[35], we observed a gain in ConvT signal in the downstream (mRNA) direction and loss in the upstream direction (**Fig. 2c**), These results strengthened the notion that promoters experience specific directional regulation and suggested that promoter pause release results in a more permissive environment for ConvT, possibly allowing downstream convergent Pol II complexes[36] to initiate and transcribe more readily.

At enhancers, elevated ConvT emanated in both directions. Notably, in control cells there was approximately eight-fold more ConvT at SEs than at TEs, and 7SK depletion resulted in broad increases in ConvT (**Fig. 2c**). Whereas SEs exhibited more transcription than TEs, they were particularly enriched in 7SK-dependent ConvT changes even when compared with promoters that have higher levels of transcription, as judged by GRO-seq (**Supplementary Fig. 4d**). This effect is probably a direct consequence of the clustered arrangement of open chromatin sites (each a TSS) within SEs; the peak-to-peak distance from one open chromatin site to the next was at least ten-fold shorter in SEs than TEs or promoters (**Supplementary Fig. 4e**). Finally, intragenic enhancers, which are located within the mRNA transcriptional unit, also experienced more ConvT than did intergenic TEs (median ConvT values of 0.067 and 0.025, respectively; **Supplementary Fig. 4f**). Therefore, 7SK has an active role in controlling the balance of transcription genome wide.

The observation that 7SK modulates ConvT, especially at SEs, raised the possibility that downstream signaling events, such as DNA-damage signaling might occur after 7SK loss. We monitored levels of S139-phosphorylated γ–histone 2A.X (γ-H2AX) by immunofluorescence and found elevated levels of γ-H2AX at the single-cell level after 7SK depletion (**Fig. 3a**). To gain a higher-resolution picture, we performed ChIP-seq of γ-H2AX in mouse ES cells with or without 7SK depletion. 7SK depletion caused significant increases of γ-H2AX at promoters, SEs and TEs (**Supplementary Fig. 4g**). Notably, there was a positive correlation between changes in ConvT and γ-H2AX at all three elements (**Fig. 3b**), and this correlation was strongest for SEs

Table 1 Mass spectrometry peptide counts

ChIRP-MS by 7SK and other nuclear noncoding RNAs

Protein	7SK	U1	U2	Xist	Protein complex
Larp7	1,027	0	0	0	7SK snRNA
Mepce	657	0	0	0	7SK snRNA
Ccnt1	655	0	0	0	7SK snRNA
Cdk9	451	0	0	0	7SK snRNA
Hexim1	342	0	0	0	7SK snRNA
Ddx21	116	77	29	9	7SK snRNA/snoRNP
Ctr9	69	24	29	0	Pol II elongation
Smarcc1/Baf155	69	0	0	0	BAF
Smarca4/Brg1	47	35	35	0	BAF
Brd4	42	0	0	0	7SK snRNA
Smarcd1/Baf60a	27	0	0	0	BAF
Actl6a/Baf53a	15	0	0	0	BAF
Smarcb1/Baf47	8	0	0	0	BAF

Coimmunoprecipitation with IgG control or BAF subunit Arid1a[a]

Protein	IgG	BAF (Arid1a)	Protein complex
Arid1a/Baf250a	0	253	BAF
Smarcc1/Baf155	0	208	BAF
Smarca4/Brg1	0	171	BAF
ActB	37	94	BAF
Smarcd1/Baf60a	0	60	BAF
Actl6a/Baf53a	0	45	BAF
Smarce1/Baf57	0	45	BAF
Smarcb1/Baf47	0	45	BAF
Smarcd2/Baf60b	0	15	BAF
Cdk9	0	2	Canonical 7SK
Ccnt1	0	0	Canonical 7SK
Hexim1	0	2	Canonical 7SK

[a]Full list of proteins is provided in **Supplementary Table 3**. SnoRNP, small nucleolar ribonucleoprotein. Slashes denote 'and'.

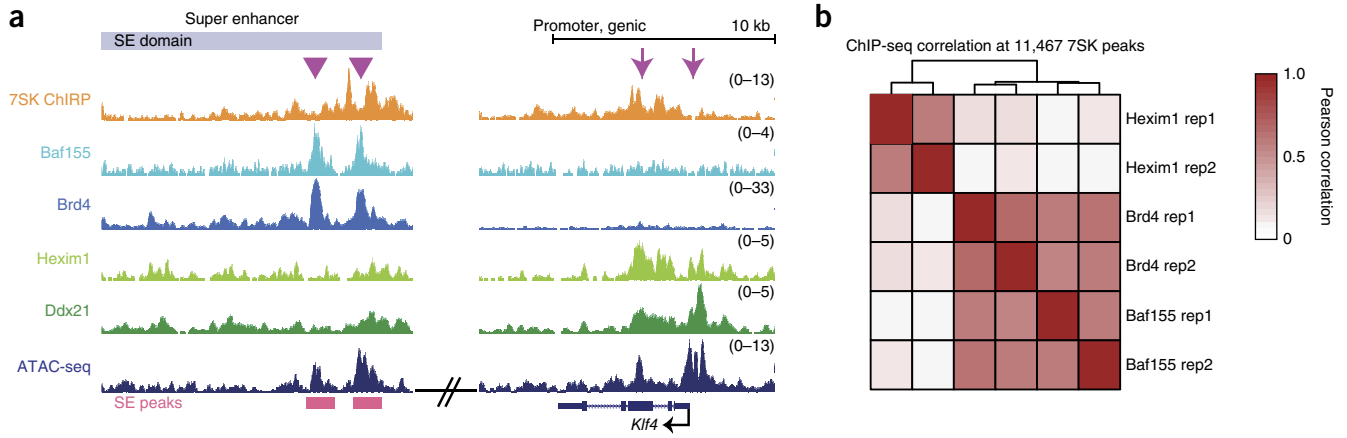

Figure 4 7SK interacts with BAF in a distinct snRNP. (**a**) UCSC genome browser view of the *Klf4* SEs (left) and genomic locus (right). Triangles (left) and arrows (right) denote regions of common binding for 7SK-SE– and 7SK-genic–associated complexes. (**b**) Pearson correlation analysis of Hexim1, Baf155 and Brd4 ChIP-seq density at 11,467 7SK ChIRP-seq peaks. Rep, replicate.

(Spearman correlation, $\rho = 0.677$). Together, these data suggested that 7SK is important for controlling the directional bias of transcription and preventing DNA damage at clustered SE elements.

7SK forms a new snRNP with the BAF complex

To discover how 7SK might functionally discriminate promoters from enhancers, we used ChIRP–mass spectrometry (ChIRP-MS) to identify new *in vivo* protein partners of the 7SK snRNA[37]. Canonical 7SK snRNP factors were highly enriched, including Larp7, Mepce, Ccnt1, Cdk9, Hexim1, Ddx21 and Brd4 (**Table 1**). Notably, 7SK retrieved five BAF-complex subunits, including the central ATPase subunit, Brg1, as well as Baf155, Baf60a, Baf53a and Baf47 (**Table 1**). The number of peptides of BAF subunits was comparable to that of a known 7SK snRNP partner (Brd4), and four of five BAF subunits were retrieved only by 7SK but not by three other nuclear noncoding RNAs (U1, U2 or Xist; **Table 1**). BAF is a mammalian ATP-dependent nucleosome-remodeling complex that has been linked to major human diseases, including in cancer and autism[38,39]. In mouse ES cells, BAF has recently been shown to regulate nucleosome occupancy and to inhibit eRNA synthesis at enhancers[40], thus making BAF an attractive partner for 7SK action. Moreover, recent reports have suggested that specific noncoding RNAs can impart locus-specific control of the BAF complex[41,42].

ChIP-seq analysis of Baf155 and other 7SK partners revealed that 7SK resides in two distinct complexes. Enhancer-bound 7SK coassociated with Baf155 and Brd4, whereas promoter-bound 7SK associated with Hexim1 and Ddx21, as exemplified by the *Klf4* locus (**Fig. 4a**). Global analysis of 7SK ChIRP-seq data revealed that ChIP-seq peaks of Baf155 and Brd4 clustered together and away from those of Hexim1 ChIP-seq (**Fig. 4b**). To further characterize these factors at the peak level, we analyzed the intersection of 7SK, Baf155 and Hexim1 peaks across the genome (**Supplementary Table 2**). We found that most enhancers had BAF occupancy (84%) but no HEXIM occupancy. Conversely, most promoters (73%) had HEXIM occupancy but no BAF occupancy. We next characterized these complexes biochemically: we separately immunoprecipitated BAF and Hexim1. By analyzing the BAF-enriched proteome by MS, we robustly recovered subunits of the BAF complex (range 14–253 peptides per protein; **Table 1** and **Supplementary Table 3**). However, we observed little Hexim1 or P-TEFb (**Table 1**). Western blot analysis of Hexim1-enriched proteins confirmed the lack of association between BAF and Hexim1 (**Supplementary Fig. 5a**). Despite the poor co-recovery between BAF and HEXIM, we found that both complexes separately pulled down 7SK. Native RNA immunoprecipitation (RIP) of Hexim1 or BAF subunit Arid1a recovered ~65% and ~6% of the cellular 7SK, respectively (**Supplementary Fig. 5b**). Given 7SK's high abundance, estimated at ~200,000 copies per cell[18], the 7SK–BAF complex may exist at more than 10,000 copies per cell and thus may potentially regulate thousands of loci across the genome.

We discovered that 7SK exists in two mutually exclusive RNA structures in association with BAF versus Hexim1. We determined the secondary structure of endogenous 7SK in association with either Hexim1 or BAF RNP complexes by using *in vivo* click selective 2′-hydroxyl

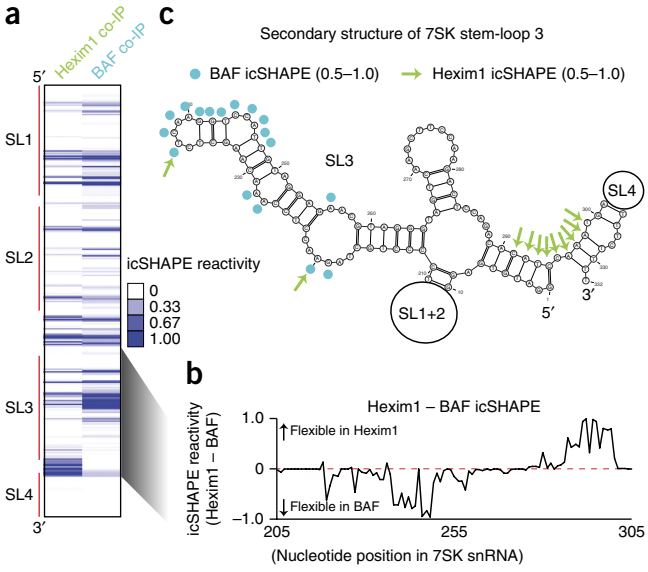

Figure 5 Two distinct 7SK RNA conformations in 7SK–BAF versus 7SK–Hexim complexes. (**a**) icSHAPE of Hexim1- or BAF-associated 7SK snRNA. The first 5 and last 15 nucleotides were not analyzed for structural reactivity, and highly reactive bases are shown in blue. (**b**) icSHAPE difference analysis between Hexim1- or BAF-associated 7SK from nucleotides 205 to 305. Positive values indicate more single-stranded in Hexim1, and negative values indicate bases more reactive in BAF. (**c**) Secondary-structure prediction of the 7SK snRNA with nucleotides differentially reactive in Hexim1- or BAF-associated structures (icSHAPE values of 0.5). Stem-loops (SL) 1, 2 and 4 are shown as loops.

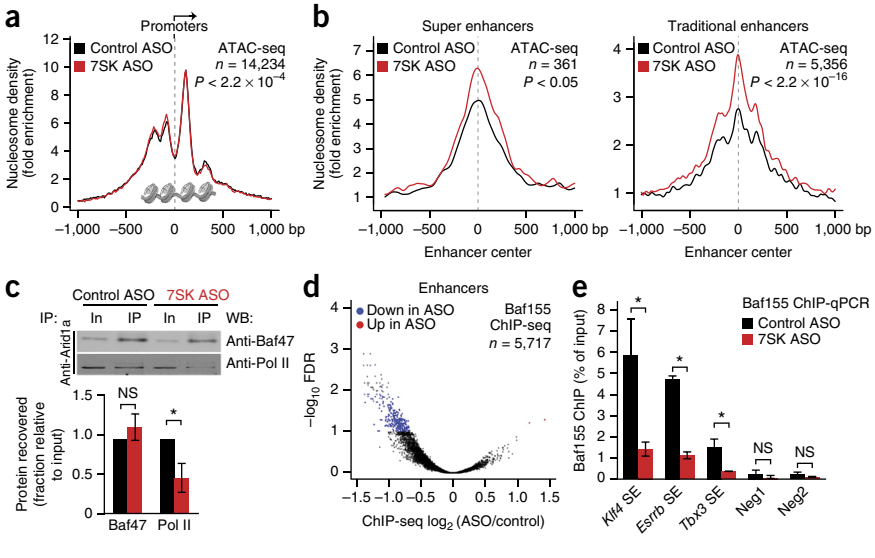

Figure 6 7SK recruits BAF and modulates nucleosome phasing at enhancers. (**a**,**b**) Metagene analysis of mononucleosome ATAC-seq signal at promoters (**a**), SEs (**b**, left) and TE (**b**, right) ±1 kb in mouse ES cells treated with control (black) or 7SK (red) ASOs. Nucleosomes are cartooned in gray. *P* values were calculated with Hotelling's *t* test (*n*, number of elements). (**c**) Co-IP of Arid1a followed by western blot (WB) analysis (top) and quantification (bottom). Quantification was performed from three independent cell cultures; mean values ± s.e.m. are shown. $*P < 0.005$, by two sided *t* test; NS, not significant. In, input. Uncropped images of gels are shown in **Supplementary Data Set 1**. (**d**) Volcano-plot analysis of all enhancers with significant (false discovery rate (FDR) < 0.1) increased (red) or decreased (blue) Baf155 ChIP-seq signal 7SK depletion. (**e**) Baf155 ChIP-qPCR analysis of mouse ES cells. Data are shown as mean and s.d. of two samples collected from separate cultures on different days. $*P < 0.05$ by two sided *t* test. Neg1 and Neg2, negative controls.

acylation and profiling experiment (icSHAPE)[43]. We treated immunoprecipitated Hexim1 or BAF complexes with the icSHAPE probe to obtain a snapshot of the 7SK structure in each RNP context. We found many single-nucleotide reactivity differences in stem-loops 1 and 2, as well as large contiguous differences between BAF- and Hexim1-associated structures in stem-loop 3 (**Fig. 5a** and **Supplementary Table 4**). By subtracting the icSHAPE values in BAF from Hexim1, we observed that nucleotides 219–250 of 7SK were substantially more reactive (single-stranded) in association with BAF, whereas nucleotides 284–300 were more reactive in association with Hexim1 (**Fig. 5b**). Mapping these reactivities onto the predicted secondary structure of 7SK's stem-loops 3 and 4 (**Fig. 5c**) further confirmed the distinct regional differences in BAF- and Hexim1-associated 7SK structural profiles. Altogether, ChIP-seq, biochemical purification and RNA structural data suggested that 7SK exists in two separable complexes: a 7SK–Hexim1 complex at promoters and a 7SK–BAF complex at enhancers.

7SK facilitates BAF action at enhancers

We next characterized the stepwise interplay of the 7SK–BAF complex and enhancer chromatin. Because transcription occurs on a nucleosomal template, and 7SK associates with the BAF complex at enhancers (**Fig. 6**), we hypothesized that 7SK might change chromatin organization. By using ATAC-seq, we inferred nucleosome positions in mouse ES cells (Online Methods) and obtained high-quality nucleosome positioning data at promoter and enhancer elements (**Supplementary Fig. 2c**,**d**). The organization and density of nucleosomes was nearly unchanged at promoters after the loss of 7SK (**Fig. 6a**). In contrast, at TE and SE elements, 7SK depletion caused stronger central nucleosome positioning (**Fig. 6b**). Histone H3 ChIP-seq revealed little change in histone occupancy at enhancers and promoters after the loss of 7SK (**Supplementary Fig. 5c**,**d**). Thus, 7SK affects relative nucleosome positioning but not occupancy.

Because 7SK is an RNA component of BAF, we hypothesized that it might bridge interactions important for the BAF complex's activity. Glycerol gradient analysis showed that 7SK depletion led to decomposition of the canonical 7SK snRNP (Cdk9, Ccnt1 and Hexim1) into fractions of lower molecular weight (**Supplementary Fig. 6a**), thus serving as positive control[44]. Importantly, 7SK depletion also shifted both Pol II and BAF complex subunits (Brg1, Arid1a and Baf155) toward lower-molecular-weight fractions (**Supplementary Fig. 6b**). These data suggested that 7SK functions to connect BAF to the transcriptional apparatus. In immunoprecipitation (IP) experiments with BAF subunit Arid1a, 7SK depletion

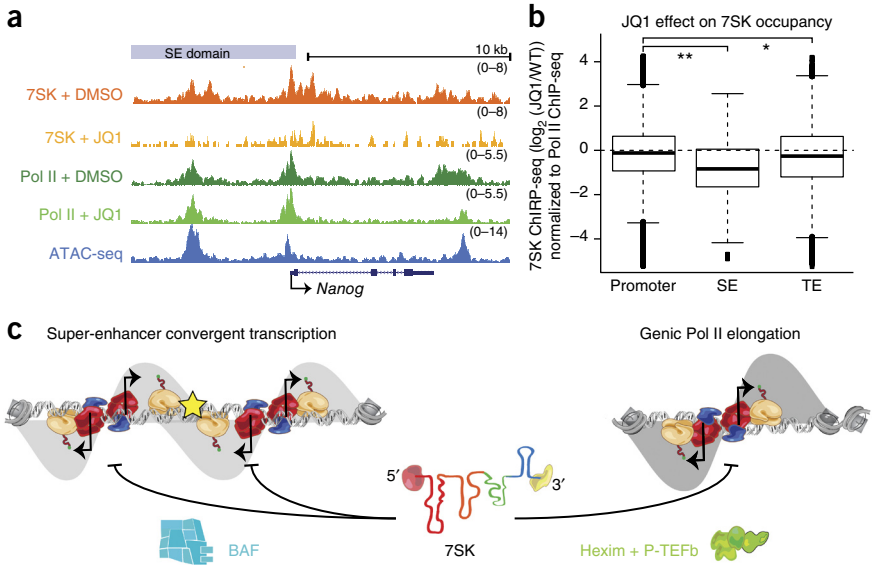

Figure 7 7SK is recruited to enhancers via bromodomain interaction. (**a**) UCSC genome browser view of the *Nanog* SEs and genomic locus. 7SK ChIRP and Pol II ChIP-seq in mouse ES cells treated for 1 h with DMSO or JQ1 are shown. (**b**) Quantitative analysis of 7SK sensitivity to JQ1 treatment at promoters, SEs and TEs, normalized to the levels of Pol II ChIP-seq. $*P < 6.5 \times 10^{-14}$; $**P < 2.2 \times 10^{-16}$ by Kolmogorov-Smirnov test. WT, wild type. Boxes, 25th to 75th percentiles; central lines, median; dashed vertical lines, 1.5× interquartile range. Data are from two cell cultures grown and collected on different days. *n* values are as in **Figure 2b**. (**c**) Model of 7SK's concerted action to repress transcription at enhancers with the BAF complex and at promoters via the Hexim1–P-TEFb snRNP.

reproducibly decreased Pol II co-IP with the BAF complex but did not affect the mutual co-IP of BAF subunits (for example, Baf47; **Fig. 6c**). These data supported 7SK's role in scaffolding BAF and Pol II, thereby facilitating a chromatin state compatible with transcribed loci.

Our biochemical data raised the possibility that 7SK might facilitate the *in vivo* recruitment to, or stabilization of, BAF on chromatin. To determine *in vivo* roles of 7SK and BAF, we depleted cells of 7SK and performed Baf155 ChIP-seq. Loss of 7SK resulted in global reduction of Baf155 occupancy at enhancers (**Fig. 6d** and **Supplementary Fig. 6c**,**d**). ChIP-qPCR validated Baf155 occupancy at multiple mouse ES SEs in a 7SK-dependent manner (**Fig. 6e**). Enhancers with 7SK-dependent BAF occupancy were enriched for selective activity during early embryogenesis, thus indicating relevance to ES-cell biology (**Supplementary Fig. 6e**). The identification of BAF, a known player in genome stability[45], as a 7SK-enhancer partner also adds to the mechanism of 7SK-mediated protection from DNA-damage signaling. Multivariate regression analysis showed that 7SK-dependent control of enhancer transcription, convergent transcription and BAF recruitment each independently and significantly diminished DNA damage (**Supplementary Fig. 6f**). The ConvT signal provided additional predictive power beyond the GRO-seq data, despite being derived from the GRO-seq data (**Supplementary Fig. 6f**), thus suggesting that overlapping read density is indeed important for γ-H2AX deposition.

Nucleosome positioning can modulate genomic TF binding[46], and ChIP-seq analysis of Oct4 (also known as Pou5f1) in mouse ES cells revealed that 7SK depletion also led to loss of Oct4 occupancy (**Supplementary Fig. 7a–e**). Analysis of TF footprints through ATAC-seq and the bioinformatics tool PIQ[47] (Online Methods) also showed that binding of Oct4 and Sox2, but not that of Ctcf, was dependent on 7SK (**Supplementary Fig. 7f–h**). To directly compare 7SK loss to BAF loss, we analyzed RNA-seq data collected after Brg1 depletion[40]. Loss of Brg1 or 7SK resulted in strongly concordant changes in eRNA synthesis at the same SE elements (281 of 349 SEs; $P = 1.85 \times 10^{-17}$ by Fisher's exact test) and TE elements (2,210 of 3,739 TEs; $P = 2.75 \times 10^{-15}$ by one-tailed Fisher's exact test). Thus, 7SK depletion mimics BAF loss. These results suggested that 7SK is important in the recruitment of BAF, a known inhibitor of eRNA transcription[40], to enhancers across the genome.

In turn, 7SK–BAF is targeted to enhancers via Brd4 bromodomain interaction. JQ1 is a small-molecule bromodomain inhibitor and a promising anticancer drug candidate[48]. We found that brief treatment (1 h) with JQ1 dissociated 7SK but not Pol II from enhancer chromatin genome wide, as measured by ChIRP-seq and ChIP-seq (**Supplementary Fig. 8**). We did not observe any compensatory expression of 7SK snRNP factors under these JQ1 treatment conditions (**Supplementary Fig. 8b**). At the *Nanog* promoter and associated SEs, Pol II's chromatin occupancy was largely stable, except for a loss at the 3′ end of the *Nanog* locus, thus suggesting that JQ1 indeed led to promoter-proximal pausing of Pol II (**Fig. 7a**). We observed globally increased Pol II pausing at genes (**Supplementary Fig. 8c**,**d**). Quantitative analysis of 7SK occupancy, normalized for changes of Pol II association after JQ1 treatment, revealed that 7SK occupancy was most sensitive to JQ1 at SEs compared to TEs and promoters (**Fig. 7b** and **Supplementary Fig. 8c–f**). This result is consistent with Brd4's relatively higher enrichment at SEs and TEs than promoters (**Fig. 1d**).

DISCUSSION

Our results revealed the distinct control mechanisms and consequences of pervasive transcription at enhancers versus promoters, notably distinguishing SEs. Enhancers are the earliest transcriptional responders to cellular differentiation and stress, and therefore are critical for orchestrated gene expression[11]. 7SK, a chromatin-bound noncoding RNA, represses transcription at enhancers via BAF chromatin remodeling and at promoters via Pol II pause release (**Fig. 7c**). Enhancer-specific linkage of 7SK to BAF allows cells to focally inhibit eRNA production and to resolve undesired consequences of enhancer transcription, such as convergent transcription at SEs. Thus, manipulation of 7SK may affect diseases including cancer and autism, in which BAF and enhancer dysfunction have important roles. Enhancers and promoters are often in proximity during gene expression, yet distinct mechanisms and small-molecule sensitivities (for example, to JQ1) control each element's transcription and other post-transcriptional features[49]. 7SK may potentially serve as a single integrator that globally regulates transcription. Future work will need to decipher how the cell discriminates among these elements and the roles of 7SK or other noncoding RNAs in this choice.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Sequencing data have been deposited in the Gene Expression Omnibus database under accession code GSE69143.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank members of the Chang laboratories for discussion and the following individuals for reagents and advice. BAF, D. Hargreaves and G. Crabtree (Stanford University); GRO-seq, I. Jonkers and J. Lis (Cornell University); JQ1, J. Bradner (Dana Farber Cancer Institute); human ES-cell culture, V. Sebastiano (Stanford University); mouse ES cells, P.A. Sharp (Massachusetts Institute of Technology); traveling ratio, P. Rahl (Syros Pharmaceutical) and C. Lin (Baylor College of Medicine); critical reading of the manuscript, P. Batista and R.C. Spitale (Stanford University). This work was supported by the Stanford Medical Scientist Program and US National Institutes of Health (NIH) grants 1F30CA189514-01 (R.A.F.); NIH grants GM068122 and GM110050 (E.T.K.); the Helen Hay Whitney Foundation (E.C.); and NIH grants P50-HG007735 and R01-HG004361, the California Institutes for Regenerative Medicine and the Howard Hughes Medical Institute (H.Y.C.).

AUTHOR CONTRIBUTIONS

R.A.F. and H.Y.C. conceived and designed the study. R.A.F. carried out the majority of the experiments and analysis. R.A.F. and M.R. performed ChIRP-seq, and R.A.F. and B.T.D. analyzed ChIRP-seq data. R.A.F. and C.C. analyzed ChIRP-MS data. R.A.F., E.C. and B.L. performed ChIP-seq and analyzed the data. R.A.F., H.K. and E.T.K. synthesized icSHAPE reagents and performed icSHAPE experiments. R.A.F. performed GRO-seq, and R.A.F., B.T.D., E.C. and J.W. analyzed GRO-seq data. R.A.F. performed ATAC-seq, and R.A.F., B.T.D., B.L., P.A.K. and A.J.R. analyzed the data. R.A.F. and H.Y.C. wrote the manuscript with input from all coauthors.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Andersson, R. *et al.* FANTOM Consortium: an atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
2. Djebali, S. *et al.* Landscape of transcription in human cells. *Nature* **489**, 101–108 (2012).
3. Meng, F.-L. *et al.* Convergent transcription at intragenic super-enhancers targets AID-initiated genomic instability. *Cell* **159**, 1538–1548 (2014).
4. Whyte, W.A. *et al.* Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).
5. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
6. Parker, S.C.J. *et al.* Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl. Acad. Sci. USA* **110**, 17921–17926 (2013).
7. Rada-Iglesias, A. *et al.* A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* **470**, 279–283 (2011).

8. Creyghton, M.P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA* **107**, 21931–21936 (2010).
9. He, H.H. *et al.* Nucleosome dynamics define transcriptional enhancers. *Nat. Genet.* **42**, 343–347 (2010).
10. Core, L.J. *et al.* Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* **46**, 1311–1320 (2014).
11. Arner, E. *et al.* FANTOM Consortium: transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science* **347**, 1010–1014 (2015).
12. Li, W. *et al.* Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* **498**, 516–520 (2013).
13. Lam, M.T.Y. *et al.* Rev-Erbs repress macrophage gene expression by inhibiting enhancer-directed transcription. *Nature* **498**, 511–515 (2013).
14. Kwak, H. & Lis, J.T. Control of transcriptional elongation. *Annu. Rev. Genet.* **47**, 483–508 (2013).
15. Campos, E.I. & Reinberg, D. Histones: annotating chromatin. *Annu. Rev. Genet.* **43**, 559–599 (2009).
16. Ørom, U.A. & Shiekhattar, R. Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell* **154**, 1190–1193 (2013).
17. Kassube, S.A. *et al.* Structural insights into transcriptional repression by noncoding RNAs that bind to human Pol II. *J. Mol. Biol.* **425**, 3639–3648 (2013).
18. Gurney, T. Jr. & Eliceiri, G.L. Intracellular distribution of low molecular weight RNA species in HeLa cells. *J. Cell Biol.* **87**, 398–403 (1980).
19. Bartkowiak, B. *et al.* CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes Dev.* **24**, 2303–2316 (2010).
20. Yu, M. *et al.* RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. *Science* **350**, 1383–1386 (2015).
21. Zhou, Q., Li, T. & Price, D.H. RNA polymerase II elongation control. *Annu. Rev. Biochem.* **81**, 119–143 (2012).
22. D'Orso, I. & Frankel, A.D. RNA-mediated displacement of an inhibitory snRNP complex activates transcription elongation. *Nat. Struct. Mol. Biol.* **17**, 815–821 (2010).
23. Ji, X. *et al.* SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* **153**, 855–868 (2013).
24. Liu, W. *et al.* Brd4 and JMJD6-associated anti-pause enhancers in regulation of transcriptional pause release. *Cell* **155**, 1581–1595 (2013).
25. McNamara, R.P., McCann, J.L., Gudipaty, S.A. & D'Orso, I. Transcription factors mediate the enzymatic disassembly of promoter-bound 7SK snRNP to locally recruit P-TEFb for transcription elongation. *Cell Rep.* **5**, 1256–1268 (2013).
26. McNamara, R.P., Reeder, J.E., McMillan, E.A. & Bacon, C.W. KAP1 Recruitment of the 7SK snRNP complex to promoters enables transcription elongation by RNA polymerase II. *Mol. Cell* **61**, 39–53 (2016).
27. Batista, P.J. & Chang, H.Y. Long noncoding RNAs: cellular address codes in development and disease. *Cell* **152**, 1298–1307 (2013).
28. Tsai, M.-C. *et al.* Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**, 689–693 (2010).
29. McGinnis, J.L., Dunkle, J.A., Cate, J.H.D. & Weeks, K.M. The mechanisms of RNA SHAPE chemistry. *J. Am. Chem. Soc.* **134**, 6617–6624 (2012).
30. Pott, S. & Lieb, J.D. What are super-enhancers? *Nat. Genet.* **47**, 8–12 (2015).
31. Williams, L.H. *et al.* Pausing of RNA polymerase II regulates mammalian developmental potential through control of signaling networks. *Mol. Cell* **58**, 311–322 (2015).
32. Hnisz, D. *et al.* Convergence of developmental and oncogenic signaling pathways at transcriptional super-enhancers. *Mol. Cell* **58**, 362–370 (2015).
33. Calo, E. *et al.* RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature* **518**, 249–253 (2015).
34. Castelo-Branco, G. *et al.* The non-coding snRNA 7SK controls transcriptional termination, poising, and bidirectionality in embryonic stem cells. *Genome Biol.* **14**, R98 (2013).
35. Flynn, R.A., Almada, A.E., Zamudio, J.R. & Sharp, P.A. Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. *Proc. Natl. Acad. Sci. USA* **108**, 10460–10465 (2011).
36. Mayer, A. *et al.* Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **161**, 541–554 (2015).
37. Chu, C. *et al.* Systematic discovery of Xist RNA binding proteins. *Cell* **161**, 404–416 (2015).
38. Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601 (2013).
39. Iossifov, I. *et al.* The contribution of *de novo* coding mutations to autism spectrum disorder. *Nature* **515**, 216–221 (2014).
40. Hainer, S.J. *et al.* Suppression of pervasive noncoding transcription in embryonic stem cells by esBAF. *Genes Dev.* **29**, 362–378 (2015).
41. Prensner, J.R. *et al.* The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. *Nat. Genet.* **45**, 1392–1398 (2013).
42. Han, P. *et al.* A long noncoding RNA protects the heart from pathological hypertrophy. *Nature* **514**, 102–106 (2014).
43. Spitale, R.C. *et al.* Structural imprints *in vivo* decode RNA regulatory mechanisms. *Nature* **519**, 486–490 (2015).
44. Krueger, B.J. *et al.* LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res.* **36**, 2219–2229 (2008).
45. Dykhuizen, E.C. *et al.* BAF complexes facilitate decatenation of DNA by topoisomerase IIα. *Nature* **497**, 624–627 (2013).
46. Spitz, F. & Furlong, E.E.M. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **13**, 613–626 (2012).
47. Sherwood, R.I. *et al.* Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat. Biotechnol.* **32**, 171–178 (2014).
48. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
49. Quinn, J.J. & Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* **17**, 47–62 (2015).

ONLINE METHODS

Cell culture and antisense-oligonucleotide knockdown of 7SK. V6.5 mouse ES cells (a kind gift from P.A. Sharp) were cultured on gelatinized plates with serum and LIF as previously described[35]. H1 human ES cells (a kind gift from V. Sebastiano) were cultured in mTeSR1 on Matrigel-coated dishes. All cells were negative for mycoplasma contamination. Cells were passaged as single cells with Accutase. HeLa cells were cultured in DMEM, 10% FBS, and 1% penicillin-streptomycin. To deplete mouse ES cells of 7SK, chimeric DNA–2′-O-methylated RNA oligonucleotides were used as previously described[34] (**Supplementary Table 5**). Briefly, cells were trypsinzed and nucleofected with an Amaxa mES cell kit (Lonza) with 1 nmol of ASO per 2 million mouse ES cells. After nucleofection, cells were grown for 2–24 h and used for downstream analysis.

Western blotting and antibodies. Cells were lysed in lysis buffer (50 mM HEPES, 200 mM NaCl, 1 mM EDTA, 10% glycerol, 0.1% NP-40, 0.2% Triton X-100, and 0.5% *N*-lauroylsarcosine, supplemented with protease-inhibitor cocktail (Roche)), briefly sonicated to solubilize chromatin, and spun for 10 min at 4 °C at 13,000 r.p.m. Clarified protein lysates were quantified with a BCA Protein Assay Reagent Kit (Pierce). Antibodies specific to the following proteins were used: Pol II N-term (Santa Cruz, sc-899), Ccnt1 (Santa Cruz, sc-10750), Cdk9 (Santa Cruz, sc-484), Brg1 (Santa Cruz, sc-17796), Hexim1 (Abcam, ab25388), Baf47 (Santa Cruz, sc-166165), Arid1a (Santa Cruz, sc-32761), and Baf155 (gift from G. Crabtree). Validation information for commercial antibodies is available on the manufacturers' websites.

Glycerol gradient analysis. Mouse ES cells were depleted of 7SK as described above, and biological duplicates (i.e., clonal cell cultures grown and collected on different days) of control and 7SK ASO were collected. After 12 h of knockdown, cells were lysed in glycerol gradient lysis buffer (10 mM HEPES, 2 mM $MgCl_2$, 10 mM KCl, 0.5% NP-40, 0.5 mM EDTA, and 150 mM NaCl, supplemented with protease-inhibitor cocktail (Roche) and SUPERaseIn (Life Technologies)) on ice for 15 min. Insoluble material was pelleted by centrifugation for 10 min at 4 °C at 13,200 r.p.m. Whole cell protein lysates were quantified as above, and 400 µg of lysate was used for each individual gradient. For each sample, 10 mL of a 10–30% glycerol gradient solution was mixed, and 400 µg of lysate (in 300 µL) was loaded and spun for 16 h at 4 °C at 40,000 r.p.m. in a Beckman SW41Ti rotor. After centrifugation and fractionation (300 µL per fraction, 30 fractions, collected from the top), samples were analyzed by western blotting.

Start-seq centering of promoters and enhancers. *Promoters.* 24,064 mm9 genes were obtained from RefSeq, and all microRNA and snoRNA genes were removed. The most upstream TSS was obtained, and a ±1-kb window was calculated around each. For each gene, the 2-kb window was first scanned along the sense strand to find the highest Start-seq[31] peak. The gene was discarded if it had no position with at least five reads on the sense strand. Next, the 2-kb window was scanned on the antisense strand to find the highest antisense peak upstream of the sense peak. A minimum threshold was not set for finding an antisense peak. If no antisense peak was found, the gene was labeled unidirectional. If this antisense peak was the sense peak for another gene, both genes were labeled bidirectional. Otherwise, the gene was labeled divergent. Only genes with unique sense peaks were retained, thus resulting in a final list of 14,234 genes with a Start-seq signal. Of these 14,234 genes, 2,060 were bidirectional, 1,443 were unidirectional, and 10,731 were divergent. We then calculated the distance (TSS-pair width) between sense and antisense peaks for all nonunidirectional genes.

Traditional enhancers. 8,563 mm9 enhancers were obtained from enhancer calls from White *et al.*[4], and ±1-kb windows were obtained around their centers. For each enhancer, the sense and antisense strands were simultaneously scanned to find the pair of sense and antisense nucleotides that maximized the sum of the Start-seq signal on both strands and where the sense peak was downstream of the antisense peak (consistently with divergent transcription). Enhancers were discarded if a Start-seq signal was not found on both strands, thus resulting in a final list of 5,356 enhancers. As with promoters, we then calculated the distance between sense and antisense peaks for all enhancers.

Super enhancers. 231 mm9 SE regions were obtained from White *et al.*[4] SEs calls and intersected with peak calls from ATAC-seq chromatin occupancy data, thus yielding 415 high-confidence ATAC-seq peaks within SEs regions. All peaks (known as SEs peaks) within 1 kb of a RefSeq TSS were removed, thus resulting in a list of 361 SEs peaks. The Start-seq centering algorithm that we ran with regular enhancers did not result in better-centered peaks, so we used these 361 SEs peaks in our final list.

ChIRP-seq assay. The ChIRP-seq assay was performed as described previously[50]. Mouse ES cells were cultured as above and treated with dimethyl sulfoxide (DMSO) or 500 nM JQ1 (ref. 48) (Gift from J. Bradner) for 1 h at 37 °C. Isolated RNA was used in qRT-PCR analysis (Stratagene) to quantify the enrichment of 7SK and depletion of other cellular RNAs. Isolated DNA was used for qPCR analysis or to make deep-sequencing libraries with an NEBNext DNA Library Prep Master Mix Set for Illumina (NEB). Library DNA was sequenced from a single end for 75 cycles on an Illumina HiSeq 2500 (summary of all sequencing experiments in **Supplementary Table 6**).

Sequencing reads were first trimmed of adaptors (FASTX Toolkit) and mapped with Bowtie to a custom index containing repetitive RNA elements (rRNA, snRNAs, and y-RNAs[51]). Subsequent reads were then mapped to mm9. Mapped reads were separately shifted toward the 3′ end with MACS and normalized to a total of 10 million reads. Even and odd replicates were merged as described previously[50] by taking the lower of the two read-density values at each nucleotide across the entire genome. The full pipeline is available at http://github.com/bdo311/chirpseq-analysis/.

ChIRP-seq data analysis. TSS regions were obtained from the RefSeq annotations of genes in the hg19 genome for H1/HeLa ChIRP-seq data. For each type of genomic feature (promoters, SEs and TEs), metagene plots were made by plotting the column means of an n × p matrix, in which the nth row represents a single example of the genomic feature, and the pth column represents the read density at position p of the genomic feature. The full pipeline for producing metagene plots is available at http://github.com/bdo311/metagene-maker/. 7SK ChIRP-seq peaks were called with MACS2 and defined as genic if the center of the peak was within 1 kb of an annotated gene (and were otherwise defined as intergenic).

Global run-on sequencing and data processing. Mouse ES cells were depleted of 7SK as described above, and biological duplicates of control and 7SK ASO-treated cells were collected. GRO-seq was performed as previously described[52] with the following modifications. After 5-bromouridine–labeled RNA was isolated, deep-sequencing libraries were constructed as described for FAST-iCLIP[51]. Libraries were sequenced from a single end for 75 cycles on an Illumina HiSeq 2500 or NextSeq 500 machine.

For each sample, 3′-adaptor trimming, quality filtering, PCR duplicate collapsing, and 5′-barcode trimming were performed as previously described[51]. Processed reads were then mapped with Bowtie to a custom index containing single-copy loci of repetitive RNA elements (rRNA, snRNAs, and y-RNAs[51]). Reads that did not map to the custom index were then mapped to mm9. These steps were similar to those in the ChIRP-seq data processing above. Metagenes were made with metagene-maker and with a custom Perl script after separation of sense from antisense reads. Statistics were calculated in a +500-bp window surrounding the center of each element.

To calculate convergent transcription (ConvT) in control and 7SK ASO-treated cells, we took the minimum of the sense signal and the antisense-normalized GRO-seq read density at each nucleotide, such that if a nucleotide had a GRO-seq signal on only one strand, ConvT = 0. To calculate the proportion of ConvT blocks before and after 7SK knockdown for different regulatory elements, we averaged the WT and ASO ConvT tracks at 100-bp intervals. For each treatment, we marked each interval as a ConvT interval if its mean ConvT signal exceeded the genome-wide mean. The observed versus expected fold change was calculated by division of the fraction of ConvT intervals mapping to each regulatory element type by the genomic fraction of that element type. Finally, metagene-maker was used to quantify average ConvT profiles in control and ASO for each regulatory element. Statistics were calculated in a ±1,000-bp window surrounding the center of each element. A summary of all changes observed at promoter and enhancers can be found in **Supplementary Table 7**.

To calculate enhancer peak-to-peak distances, we took the distance from an enhancer to its closest TE or individual SE peak. Peaks were centered with Start-seq signal (TEs) or ATAC-seq signal (SEs) as described previously.

ATAC sequencing and data processing. ATAC-seq was performed essentially as previously described[53]. Briefly, mouse ES cells were depleted of 7SK as described above. Transposition was performed for 30 min at 37 °C, after which cells were lysed, and isolated library DNA was sequenced from a single end for 75 cycles on an Illumina HiSeq 2500.

ATAC-seq reads were processed as previously described[53]. Once mapped reads were obtained, paired-end ATAC-seq fragments from two length ranges—fragments shorter than 100 bp, and fragments between 147 and 180 bp—were isolated for separate analysis. These ranges were previously used by Buenrostro *et al.*[53] to distinguish transposase insertion into nucleosome-free DNA and insertions flanking nucleosomes, respectively. At promoters, TEs and SEs, the number of fragment centers (single base pair) overlapping each bin was counted for each of the two fragment size ranges. We plotted the average fragment count in each bin normalized to the average fragment count in the first five bins to set the background signal to one. Statistics were calculated in an ±250-bp window surrounding the center of each element.

Transcription factor–footprint detection with PIQ. We used PIQ[47] to quantify DNA binding of TFs on the basis of ATAC-seq data. To maximize sensitivity, we merged ATAC-seq data from replicates. To eliminate detection biases based on sequencing depth, we sampled from the higher-depth merged library to achieve equal depth from control and 7SK-depleted libraries (resulting in ~89 million single-end reads for each condition). Purity thresholds for footprint calls in control samples were adjusted to identify 10,000 binding events for each TF.

Immunofluorescence (IF). IF was performed as described previously[33]. Briefly, mouse ES cells were treated with control ASO or 7SK ASO and seeded into 12-well plates containing 18-mm glass cover slips and cultured for 12 h, after which cells were fixed in 4% PFA for 10 min at room temperature and washed 3 × 5 min with PBS. Cells were permeabilized in PBS containing 0.3% (v/v) Triton X-100 for 5 min and blocked overnight at 4 °C in PBT buffer (PBS with 1% BSA, 0.1% Triton X-100 (v/v), and 0.05% sodium azide (w/v)). After blocking, cover slips were incubated in PBT with a 1:200 dilution of S139-phosphorylated γ-H2AX–specific antibody (Abcam, ab11174) at room temperature for 2 h. Cover slips were washed 3 × 5 min with PBT and incubated with Alexa-Fluor 568 secondary antibody (1:1,000; ThermoFisher Scientific, A-11036) for 1 h. Validation information for antibodies is available on the manufacturers' websites. Cells were washed 3 × 5 min with PBT and 2 × 5 min with PBS, rinsed briefly with water and mounted onto glass slides with VECTASHIELD (Vectorlabs, H-1200) mounting medium with DAPI. All images were acquired and processed with a Zeiss LSM700 confocal microscope.

***In vivo* click SHAPE (icSHAPE).** icSHAPE was performed largely as described previously[43] with the following modifications. Mouse ES cells were grown as above, cell lysates were made as for glycerol gradient analysis and quantified, and 400 μg of each sample was used for further processing. For Hexim1 enrichment, 30 μL of Protein A Dynabeads (Life Technologies) was prepared overnight at 4 °C with 8 μg of anti-Hexim1 antibody (Abcam, ab25388). For BAF (Arid1a) enrichment, 30 μL of Protein A Dynabeads (Life Technologies) was prepared overnight at 4 °C with 5 μg of rabbit anti-mouse secondary antibody (ThermoFisher Scientific, 31188) and 8 μg of anti-Arid1a antibody (Santa Cruz, sc-32761). Validation information for antibodies is available on the manufacturers' websites. After overnight incubation, beads were washed twice in 1 mL of NT2 buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1 mM EDTA, and 0.05% NP-40) and added to freshly prepared lysate and immunoprecipitated overnight at 4 °C. Samples were then washed three times in glycerol gradient lysis buffer each for 2 min on ice. After being washed, beads were resuspended in 30 μL of icSHAPE reaction buffer (100 mM HEPES, 6 mM $MgCl_2$, 100 mM NaCl, 1 U SUPERaseIn (Life Technologies), and 50 mM NAI-N_3 (or DMSO as mock)), and incubated at 37 °C for 12 min. After modification, RNA was immediately isolated with TRIzol (Life Technologies), and this was followed by RNA cleanup with RNeasy Mini columns (Qiagen). Subsequently, RNA was treated as standard DMSO or NAI-N_3 samples for standard icSHAPE library preparation[43]. DNA library samples were subjected to deep sequencing on Illumina NextSeq 500 machines for single-end 75-bp cycle runs.

Data-processing steps were performed as described previously[43] with the following modifications. We generated biological replicates of two IP samples (BAF and Hexim1) for DMSO and NAI-N3 modified experiments. Reads were collapsed to remove PCR duplicates, barcodes were removed, and reads were mapped specifically to the full-length mouse 7SK snRNA sequence (NCBI GeneID 19817). In NAI-N_3 libraries, we defined the −1 position of the 5′ mapped reads as RT (reverse transcription) stops, corresponding to structure-modified positions. We defined RT-stop coverage as the number of times that a base mapped as an RT stop. We constructed the background base density profile for 7SK as the sequencing depth of each base in the DMSO libraries. We defined the reactivity score as the value after subtraction of background RT stops (DMSO libraries) from RT stops of the modified NAI-N_3 libraries, adjusted by the background base density. The score was then scaled to the range of [0, 1], after removal of the outliers by 90% Winsorization (the top fifth percentile was set to 1, and the bottom fifth percentile was set to 0).

Native Immunoprecipitation (IP) and RNA-IP. Mouse ES-cell extract was prepared for analysis of native protein-protein and RNA-protein interaction analysis, as described previously[33]. Hexim1 (Abcam, ab25388) or BAF (Arid1a, Santa Cruz, sc-32761) was enriched with bead-coupled antibodies for 4 h at 4 °C. Validation information for antibodies is available on the manufacturers' websites. Enriched samples were washed three times on ice in glycerol gradient lysis buffer. Proteins or RNA pulled down with each IP were analyzed by western blotting, mass spectrometry, or qRT-PCR.

ChIP-seq. *Sample and library preparation.* ChIP-seq was performed as described previously[33]. DNA was purified for subsequent qPCR analysis or deep-sequencing library construction (as above for ChIRP-seq). Antibodies specific to the following proteins were used for ChIP studies: Pou5f1 (Santa Cruz, sc-8629), Hexim1 (Abcam, ab25388), Ddx21 (Novus, NB100-1781), Baf155 (Gift from G. Crabtree), γ-H2AX (Abcam, ab11174), and Brd4 (Bethyl, A301-985A). Validation information for commercial antibodies is available on the manufacturers' websites.

Factor ratio analysis. For all TE, SE and promoter regions (−1 kb to +1 kb around their centers), we calculated a mean read density for all ChIP-seq, ATAC-seq, and 7SK ChIRP-seq data sets by dividing total coverage by the number of nucleotides (2,000) and obtained enrichment values by dividing these values by the values of the corresponding input samples. Enrichment values for each factor at each region were all normalized by the enrichment of TBP[54], TBP[55], or TAF1 (ref. 55) at that region. To make heat maps, for all factors we averaged these TBP- or TAF1-normalized enrichment values across promoters, TEs and SEs, and displayed all factors or a subset of factors with R's pheatmap library. To show that promoters, TEs and SEs have different factor ratios as a class, we ran principal component analysis on the TBP[54]-normalized enrichment values for all regions with the prcomp function in R and plotted the first two principal components with R's ggplot2 library.

Data analysis. Raw ChIP-seq data were obtained from the Gene Expression Omnibus database (**Supplementary Table 8**). All ChIP-seq data were mapped to the genome and normalized to 10 million reads to facilitate comparison with ChIRP-seq data. The format of ChIP-seq data was identical to the format of ChIRP-seq data: for each factor of interest, each position in the genome had a value representing its binding level relative to other positions. Metagenes for ChIP-seq data were produced as described above, and statistics were calculated in a ±250-bp window surrounding the center of each element.

To assess cobinding of Hexim1, Baf155, and 7SK, we called peaks from two Hexim1 and two Baf155 ChIP-seq replicates and even and odd replicates of 7SK ChIRP-seq with MACS2 with the following parameters. 7SK parameters were –broad–broad-cutoff 0.3 –p 1e–3. We obtained 323,145 overlapping peaks, of which 85,230 had signal values within 50% of each other in both data sets. We took the top 50,000. Baf155 parameters were –broad–broad-cutoff 0.3 –p 1e–3. We obtained 1,006,367 overlapping peaks, of which 959,238 had signal values within 50% of each other in both data sets. We took the top 10,000. Hexim1 parameters were –p 1e-3. We obtained 346,242 overlapping narrow peaks, of which 19,634 had IDR <0.01. We took the top 10,000. These peaks were subsequently analyzed in two ways. First, each peak was assigned to either a promoter or enhancer element.

To derive a linear model to explain the γ-H2AX fold change after 7SK knockdown, we used $\log_2$ ASO/Scr fold change values at TEs and SEs for GRO-seq, Baf155 ChIP-seq, γ-H2AX ChIP-seq, and ConvT. We used the R library edgeR to

doi:10.1038/nsmb.3176

process the GRO-seq and ChIP-seq data and in-house Python scripts to process ConvT data. GRO-seq, Baf155, and ConvT values were then regressed onto the γ-H2AX fold change one factor at a time with R's lm function, and the adjusted R-squared value was calculated for each regression.

We used edgeR and displayed the data as volcano plots to quantitatively analyze the changes in ChIP-seq or GRO-seq signals. Specifically, we counted ChIP-seq or GRO-seq reads falling in specified windows around promoters and enhancer TSSs and used edgeR with two biological replicates for each condition (scrambled or 7SK-targeting ASO) to identify the subset of regions exhibiting significantly altered signals in 7SK-depleted cells. Read counts were normalized by library size for each sample. We used a threshold of FDR <0.1 to call significantly altered regions. Volcano plots represent the $\log_2$ fold change for the particular experiment (GRO-seq or ChIP-seq), plotted against $-\log_{10}$(FDR).

50. Chu, C., Qu, K., Zhong, F.L., Artandi, S.E. & Chang, H.Y. Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol. Cell* **44**, 667–678 (2011).
51. Flynn, R.A. *et al.* Dissecting noncoding and pathogen RNA-protein interactomes. *RNA* **21**, 135–143 (2014).
52. Jonkers, I., Kwak, H. & Lis, J.T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* **3**, e02407–e02407 (2014).
53. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y. & Greenleaf, W.J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218 (2013).
54. Rahl, P.B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
55. Liu, Z., Scannell, D.R., Eisen, M.B. & Tjian, R. Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* **146**, 720–731 (2011).